Technology *Report*

Issue 8 May 2000

Drug Treatments for Alzheimer's Disease. I. A Comparative Analysis of Clinical Trials

Publications can be requested from:

CCOHTA
110-955 Green Valley Crescent
Ottawa, Ontario, Canada, K2C 3V4
Tel. (613) 226-2553
Fax. (613) 226-5392
Email: pubs@ccohta.ca

or download from CCOHTA's web site:
http://www.ccohta.ca

Cite as: Wolfson C, Moride Y, Perrault A, Momoli F, Demers L, Oremus M. **Drug treatments for Alzheimer's disease. I. A comparative analysis of clinical trials**. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA); 2000.

Legal Deposit - 2000
National Library of Canada
ISBN 1-895561-78-7 (Set)
ISBN 1-895561-82-5 (Part 1)

Publications Mail Agreement Number: 1633171

Canadian Coordinating Office for Health Technology Assessment

Drug Treatments for Alzheimer's Disease. I. A Comparative Analysis of Clinical Trials

Christina Wolfson, PhD[1,2]
Yola Moride, PhD[1,3]
Anne Perrault, MSc[1]
Franco Momoli, MSc[1,2]
Louise Demers, PhD[1]
Mark Oremus, MA[1,2]

May 2000

Technical assistance provided by:
Carole Bohbot, Irina Lazariciu, Alexander Tsertsvadze

1. Centre for Clinical Epidemiology and Community Studies, S.M.B.D. Jewish General Hospital, Montréal, Québec
2. Joint Departments of Epidemiology & Biostatistics and Occupational Health, McGill University, Montréal, Québec
3. Faculté de Pharmacie, Université de Montréal, , Montréal, Québec

TABLE OF CONTENTS

REVIEWERS

These individuals kindly provided comments on this report.
CCOHTA takes sole responsibility for the final form and content.

CCOHTA Scientific Advisory Panel

Dr. Gina Bravo
Sherbrooke University
Geriatric Institute
Sherbrooke, Quebec

Dr. David Persaud
Dalhousie University
Halifax, Nova Scotia

External Reviewers

Dr. Richard Gladstone
University of Toronto
Toronto, Ontario

Dr. Willie Malloy
McMaster University
Geriatric Research Group
Hamilton, Ontario

ACKNOWLEDGMENTS

The Canadian Coordinating Office for Health Technology Assessment (CCOHTA) provided the funding for this study. Partial salary support for Carole Bohbot and Alexander Tsertsvadze was provided by the "Fonds de la Recherche en Santé du Québec" (FRSQ), through an Equipe Prioritaire "Épidémiologie de la démence chez les personnes âgées".

EXECUTIVE SUMMARY

Alzheimer's disease is a devastating disorder of the elderly, which is expected to become a major public health problem in Canada over the next two decades. Recent projections indicate that the number of cases of Alzheimer's disease in Canada will increase from 160,000 to more than 380,000 by the year 2020. These projections are based on the assumption that the incidence of Alzheimer's disease will remain unchanged.

Significant research efforts in Canada and around the world are seeking to determine the cause of AD and to find a cure. Until this has been achieved, however, there remain thousands of Canadian elderly suffering from AD. A major research focus in AD, at the present time, is the development of therapies that could alter the course of the disease in those already afflicted.

In this report we review those published clinical trials that we believe have the methodological integrity to provide the best evidence on the efficacy of donepezil, metrifonate, rivastigmine, selegiline, vitamin E, lecithin, linopirdine, propentofylline and gingko biloba for the treatment of Alzheimer's disease.

Twenty-seven randomized clinical trials were retrieved from the literature and found to meet appropriate methodological standards.

We conclude that for selegiline, vitamin E, lecithin, linopirdine, and propentofylline the published data do not provide support for efficacy.

Based on the evidence we reviewed, it is our conclusion that donepezil, metrifonate and rivastigmine, however, all provide statistically significant modest benefit on cognitive performance and global functioning to the elderly with probable AD who are eligible for inclusion in clinical trials. The magnitude of the effect is similar for all of the medications. The results from the trials of gingko biloba are promising but the effects are smaller than those from the above mentioned therapies.

Although all of these medications appear to be well tolerated, in terms of the occurrence of adverse events, dropout rates are sometimes high and may have resulted in overestimation of apparent treatment effects.

It is important to note that this report is based on the results of published trials only and as a result may be subject to publication bias. In particular, if a bias exists such that trials showing no effect are less likely to be published then our findings may overestimate efficacy through the selective inclusion of positive trials. The exclusion of trials that were not published due to poor methodology, however, would not have resulted in a bias as it is unlikely that such trials would have met our standards of methodological rigour.

In a companion report (*Drug Treatments for Alzheimer's Disease. II. A Review of Outcome Measures in Clinical Trials*), we review the psychometric properties of the primary and secondary outcome measures used in these trials and we remain concerned about the wide variety of scales used that do not have adequate psychometric assessment. Although a great number of the scales used have evidence of reliability and validity, responsiveness to change has generally not been adequately assessed. For this reason, the clinical significance of the treatment: placebo differences remain unclear.

The ability to carry out a comparative analysis of therapies for AD depends not only on the comparability of design, duration, and outcome measures used but also on the methods of reporting the results of the trials. There was no consistent method of reporting the results of the AD trials even when emanating from the same group of investigators. Although journals do have different guidelines for authors, we recommend that reports of clinical trials of therapies for AD follow a standardized format for presenting results, which would enhance the ability to compare results across studies.

1 INTRODUCTION

1.1 Introduction

Alzheimer's disease (AD) is a devastating disorder resulting in loss of components of short term memory and immediate recall, plus a decline in other higher cognitive functions such as attention. Eventually memory loss is so severe that patients lose the ability to care for themselves. The most recent information on the prevalence of AD in Canada comes from the Canadian Study of Health and Aging (CSHA) which was completed in 1992 (1). In this population based survey, randomly selected groups of people aged 65 and over were screened for the presence of cognitive impairment using the modified mini mental state examination (3MS) (2). Those who screened positive for cognitive impairment and a randomly selected sample of those screening negative were asked to undergo a clinical assessment to determine the presence of dementia and to provide a diagnosis. The DSM-IIIR (3) criteria were used for the diagnosis of dementia and the NINCDS/ADRDA criteria (4) for the diagnosis of AD. A total of 9008 community dwelling elderly participated and the prevalence of AD was found to be 1% for those 65-74 years of age, 7% for those 75-84 years of age and 26% for those over the age of 85.

The results from the CSHA are in agreement with prevalence figures from studies done elsewhere. It was estimated that at the time of the study there were more than 78,000 elderly Canadians living in the community suffering from AD. Projections from the same study, based on the assumption of a stable incidence and survival indicate that by the year 2020 the total number of cases of AD in Canada (living in the community or institutionalized) will have increased from approximately 161,000 in 1992 to more than 387,000 making this disease a major public health issue in the next millennium.

Up until the last decade there were no acceptable drug therapies for Alzheimer's disease (AD). In September of 1993, tacrine (Cognex®) was approved for use as a treatment for AD by the United Stated Food and Drug Administration (FDA). This was followed in November of 1996 by the approval of donepezil hydrochloride (ARICEPT^{TM}). Donepezil was approved in Canada in August of 1997. Tacrine has not been approved in Canada. A number of other potential medications are presently being

tested and some show promise. In this report we carry out a review of the best evidence for the efficacy of individual drug therapies for AD. Evaluation of therapies for AD is complex due to the varied methodologies applied in trials (e.g., trial design, inclusion criteria, outcome measures, dosages, length of the trial, size of study, etc.), as well as the variation in the presentation of methods and detailed results in the published literature. In a previous review of tacrine (5), our group concluded that a pooling of trials to carry out a standard meta-analysis was inappropriate due to the disparate methods used.

1.2 Objectives

The project has three main objectives:

1: To assess the evidence for the clinical efficacy of individual therapies for AD.

2: To compare the clinical efficacy of individual therapies for AD.

3: To review, in a companion report, the reliability and validity data on all primary and secondary outcome measures used in the trials selected in Objective 1.

1.3 Efficacy/Effectiveness

The efficacy of a pharmacological or medical treatment is defined as the extent to which the specific intervention produces a beneficial result under ideal conditions. For the most part, these ideal conditions are the aim of randomized clinical trials. Efficacy is a prerequisite for effectiveness, which is defined as the measure of the extent to which the specific intervention produces a beneficial result when employed under circumstances that mimic the “real world”. The extrapolation of efficacy data to evidence of effectiveness is complex and requires not only the consideration of a number of trial related features (e.g., exclusion criteria, compliance with the treatment regimen, withdrawals from the study, monitoring of compliance, etc.) but also information on the acceptability of the treatment to both patients, caregivers, and treating physicians.

1.4 Best Evidence Synthesis

The limitations of traditional narrative reviews to assess the efficacy of treatments have long been recognized. Meta-analysis was introduced in the mid-1970s as a means of synthesizing the findings of multiple studies of the same topic in order to improve the

quality of evidence assessment. Although there has been an extraordinary growth in the use of meta-analyses to assess the efficacy of intervention studies, the methodology has introduced a new set of limitations and problems. Interested readers are referred to an excellent review by Slavin (6) for more details.

As a response to concerns about misleading conclusions resulting from meta-analyses, Slavin proposed an alternative procedure that is called "best evidence synthesis", which incorporates many of the positive features of meta-analysis but also retains important aspects of "intelligent and insightful" narrative reviews. A best evidence synthesis includes the careful selection of studies based on methodological quality and a synthesis of the selected literature, but does not result in the presentation of a summary statistic based on a quantitative analysis.

1.5 Research Group

The study was carried out at the Centre for Clinical Epidemiology and Community Studies (CCECS), Sir Mortimer B. Davis-Jewish General Hospital, Montréal, Québec, Canada. The four investigators were: Dr. Christina Wolfson, a neuroepidemiologist and associate professor in the Departments of Epidemiology and Biostatistics and of Medicine at McGill University; Dr. Yola Moride, a pharmacoepidemiologist and assistant professor in the Faculty of Pharmacy at the University of Montréal; Anne Perrault, an epidemiologist and research associate at the CCECS; and Franco Momoli, a PhD candidate in the Department of Epidemiology and Biostatistics at McGill University. After the study started, Louise Demers, an occupational therapist and post-doctoral fellow at the CCECS and an expert in functional scales, was recruited. Mark Oremus, a MSc candidate in the Department of Epidemiology and Biostatistics at McGill University was the project coordinator. Two research assistants, Irina Lazariciu, a MSc candidate in the Department of Mathematics and Statistics at McGill University, and Dr. Alexander Tsertsvadze, a Diploma candidate in the Department of Epidemiology and Biostatistics at McGill University, provided technical support. Carole Bohbot provided secretarial assistance.

2 METHODS

2.1 Medications for Alzheimer's Disease

The World Wide Web, general medical textbooks and textbooks specific to AD and dementia were first consulted to identify possible drug therapies for AD. Medications were selected without regard for their mechanism of action. In that way, 86 medications were identified. In consultation with CCOHTA, the search strategy for published trial reports was restricted to therapies that were "on the market" or in Phase III pharmaceutical company clinical trials. "On the market" meant being available to consumers for purchase either by prescription or over the counter. Approval by Health Canada's Drug Protection Branch or inclusion on a provincial government's drug formulary were not considered prerequisites for meeting the "on the market" criterion. As a result of this consultation, the list of medications was reduced to 14 (table 2.1).

Table 2.1: Drug Therapies for Alzheimer's Disease – On the Market or in Phase III Clinical Trials

Donepezil	Vitamin E	Indomethacin
Metrifonate	Selegiline	Prednisone
Rivastigmine	Ginkgo Biloba	Linopirdine
Galanthamine	Estrogen	Propentofylline
Lecithin	Aspirin	

2.2 Search Strategy

Four specific search areas were targeted: (i) PERUSE - McGill University; (ii) Cochrane Library; (iii) review articles; (iv) multi-media. No clinical trials published prior to 1986 were included in the review. The NINCDS/ADRDA research diagnostic criteria (4) for AD were published in 1984, and it was felt that the validity of AD diagnoses in trials conducted earlier would be difficult to establish. Unpublished material and studies reported only as abstracts were also excluded from the review.

2.2.1 PERUSE - McGill University

PERUSE is a McGill University Libraries search engine that provides access to several databases related to social sciences, life sciences and physical sciences. PERUSE was used to search the following databases: Applied Science and Technology, CINAHL (Nursing and Allied Health), Core Biomedical Collection, Core Biomedical Collection III, HealthSTAR, MEDLINE and PsycINFO.

The 14 medications in table 2.1 were used as keywords in the search. Each medication name was combined with the keywords "Alzheimer's" and "randomized". The general MeSH term "Alzheimer's Disease/drug-therapy" was also tried for each database to capture review articles and additional randomized controlled trials (RCTs).

2.2.2 Cochrane Library

The Cochrane Library is a regularly updated database collection that contains systematic, up-to-date reviews of RCTs. The Cochrane Controlled Trial Register (CCTR), a bibliography of clinical trials within the larger database, was used to search for RCTs and review articles that may not have been identified through PERUSE. The search term employed was "Alzheimer-disease-drug-therapy".

2.2.3 Review Articles

PERUSE and the CCTR yielded 40 review articles that were hand searched for other RCTs. Clinical trials published before 1990 were primarily identified from review articles.

2.2.4 Multi-Media

Miscellaneous newspapers, press clippings and popular magazines that had been gathered by the principal investigator were also searched for any references to clinical trials or review articles.

2.3 Results

The literature search yielded 65 clinical trials (table 2.2).

2.3.1 Selection of Trials for Review

The 65 trials were subjected to a blinded quality assessment and an evaluation of diagnostic criteria before being retained for review. Prior to the review process, three reviewers (CW, YM and FM) carried out an independent blinded review of a treatment for AD that was not part of this study. A clinical trial of thiamine (7) was blinded by a research assistant who removed the abstract, year, journal, authors, results and discussion. The methodological quality of the trial was assessed using the Jadad scale (8). At a consensus meeting following the thiamine "training" review, it was decided that the portion of the results section pertaining to losses to follow-up, withdrawals and adverse effects should be included with the material for blinded review. The training session was also useful to clarify the scoring system of the Jadad scale, which has been described as "a quantitative index of the likelihood that the [trial's] reported methodology and results are free of bias." (9) Rating the quality of a clinical trial is one means of helping to ensure that only the most methodologically sound trials are included in a systematic review (10).

The Jadad scale was chosen because it appeared both easy to use and well constructed. Moher et al. (9) examined 24 quality instruments and found that the Jadad scale was the only instrument that met their criteria for rigorous development. These criteria were: clear reporting of how the scale items were initially selected and why the final items were chosen, and discussion of how the instrument discriminated between RCTs of varying quality. The Jadad scale can be used as either a 3-item or 6-item version. The 3-item scale contains the three questions (pertaining to randomization, double-blind, and patient withdrawals) shown in table 1 of appendix A. In this study, we used the 6-item scale (consisting of all the questions in tables 1 and 2 of appendix A), which ranges in score from 0 to 8.

Jadad et al. (8) found fair to good interrater reliability for their scale. In their study reliability was assessed with seven unblinded and eight blinded reviewers (blinded as to author, journal and results) in an assessment of 36 pain research trials. Intraclass correlation coefficients (ICCs) ranged from 0.56 to 0.76 for analyses of 3-item and 6-item versions of the instrument. A recent study (11) of 76 trials in the area of blood transfusions during surgery found low interrater reliability (kappas from 0.37 to 0.39) for the Jadad scale when either two or four raters reviewed the published articles in a blinded

or unblinded manner. Interrater reliability improved (kappas from 0.53 to 0.59) when the item asking about adequate descriptions of patient withdrawals was removed from the scale.

Although interrater reliability differed between the two aforementioned studies, the Jadad scale was still chosen because of its rigorous development and apparent ease of use. Also, the scale item pertaining to withdrawals was retained because examinations of drug efficacy should consider the impact of dropouts. The withdrawal question could be included without necessarily threatening the reliability of the Jadad scale because interrater reliability fluctuates across different research domains. Poor performance by a scale item in one area is not necessarily indicative of poor performance in another area. Interrater reliability was examined after all of the trials had been scored to see if there was any ex post facto cause for concern.

After the training session with the thiamine trial, a research assistant distributed blinded copies of the methods section from each of the 65 trials to the raters. Portions of the results that detailed withdrawals and losses to follow-up were also distributed. Blinding involved removing all references to author, journal and location from the material provided to the raters. Each rater independently assessed all 65 blinded trials. A consensus meeting was held to identify and resolve any major discrepancies in scoring for particular studies. No scoring problems were found.

The raters decided at the consensus meeting to exclude those trials that had a quality score of less than 5. This cut-off ensured, due to the Jadad scoring mechanism, that acceptable trial reports would almost certainly be randomized and double-blind in nature. Additionally, reports would include descriptions of at least some other features judged important (by the raters) for evaluating drug treatments (e.g., withdrawals, adverse effects, inclusion/exclusion criteria and statistical analyses).

In a separate assessment, one of the reviewers (CW) reviewed the diagnostic criteria used in each trial. In addition to scoring 5 or more on the Jadad scale, trials for inclusion in the review were required to have used the NINCDS/ADRDA diagnostic criteria (4) to identify AD. Given that AD is a syndrome, not an illness with definite biological markers before death, it was important to ensure that a uniform set of diagnostic criteria was selected to maximize the comparability of patients across trials.

Several of the ginkgo biloba studies (12-14) posed particular problems as they did not clearly indicate whether or not subjects satisfied the NINCDS/ADRDA criteria. A neurologist specializing in cognition was consulted to judge the appropriateness of the diagnostic criteria that were used in these trials. Trials not using NINCDS/ADRDA criteria were excluded from the review regardless of their quality scores.

Table 2.2 contains the final number of trials that were selected for review. Specific reasons for the rejection of studies from the original list of 65 are included in the drug review section (chapter 3) and appendix B.

The reviewers found the Jadad scale easy to use and were able to carry out each trial review in less than 10 minutes. Using the data gathered from the blinded reviews, interrater reliability was assessed using the Intraclass Correlation Coefficient (ICC) (15). The interrater reliability was found to be excellent with an ICC of 0.9.

While the blinded reviews were underway, the three other investigators (AP, LD, and MO) began a review of the psychometric properties of the primary and secondary outcomes used in the trials selected in the initial search. Given the time constraints of the project, the group necessarily reviewed some outcomes used in trials that were eventually excluded from further consideration.

Table 2.2: Number of RCTs Obtained from the Literature Search & Final Number of RCTs Selected for Review

Medication	# RCTs from Search	# RCTs for Review
Donepezil	4	4
Metrifonate	8	6
Rivastigmine	2	2
Galanthamine	0	0
Lecithin	12	1
Selegiline (incl. Vitamin E)	21	7
Ginkgo Biloba	7	3
Estrogen	4	0
Aspirin	0	0
Indomethacin	1	0
Prednisone	0	0
Linopirdine	2	1
Propentofylline	4	2
TOTAL	65	26

3 METHODOLOGICAL DESCRIPTION AND BEST EVIDENCE SYNTHESIS OF THE ANTIDEMENTIA TRIALS

3.1 Introduction

This chapter presents the results of the systematic reviews of selected randomized controlled trials for the treatment of Alzheimer's disease (AD). The treatments are grouped according to mode of action. The cholinesterase (ChE) inhibitors (donepezil, metrifonate, and rivastigmine) are reviewed first, followed by the anti-oxidants (selegiline and vitamin E). The review is rounded out by consideration of lecithin (a choline supplement), propentofylline (an agent that appears to have nerve growth factor enhancing properties), and linopirdine (a compound that enhances presynaptic endogenous acetylcholine release). Finally, clinical trials of an extract of ginkgo biloba, an over the counter substance that is believed to have anti-oxidant properties, are assessed.

The suggested use of AChE inhibitors in AD arises from the cholinergic hypothesis. The basis of this hypothesis is that it may be possible to correct a biochemical deficit in AD by stimulating the cholinergic system. In this way, it is hoped that impairment will be halted or slowed. One treatment strategy, related to AChE inhibition, consists of maximizing the available acetylcholine by inhibiting its breakdown. Two AChE inhibitors, tacrine and donepezil, have been approved in the United States for the treatment of AD. In Canada, donepezil has been approved but tacrine's adverse effects and modest clinical efficacy resulted in it not being apprived by Health Canada's Health Protection Branch. Metrifonate is under study but is on clinical hold to assess concerns about rare cases of proximal weakness (personal communications; Bayer Inc.). Rivastigmine is currently under study.

The development of medications with antioxidant properties was largely motivated by evidence for oxidative stress in patients with AD. Two drugs reviewed in the report are included in this category: selegiline and vitamine E. Vitamine E (Tocopherol) is a fat-soluble substance that blocks lipid peroxidation. Selegiline, also called *l*-deprenyl (Eldepryl® in the US), is a centrally and peripherally acting irreversible monoamine oxidase inhibitor (MAOI). Its pharmacological profile is unusual in that low doses (10 mg/day) selectively inhibit MAO-B, the predominant form of MAO in the brain in humans, whereas nonspecific inhibition of MAO-A and MAO-B occurs at higher doses. MAO-B elevation is known to occur naturally in the brain with age and is

significantly increased in AD. When too high, this enzyme elevation may represent a reversible functional neurochemical lesion.

Each drug review includes a description of the architecture of the selected trials (e.g., parallel vs. crossover, multi- vs. single- center, duration, size, and inclusion/exclusion criteria), treatment regimen and outcome measures. The results of the trials are presented in terms of cognitive outcomes, global outcomes and other, secondary, outcomes. Adverse events are reported in each drug review and tabulated in a standard format in section 3.10. The selection and use of outcome measures in clinical trials of AD is an area of concern. Thus, the review process included an in-depth examination of the outcome measures in the AD trials considered. The appropriateness of the outcome measures is included within each drug review but due to the length of the outcome measure review, it is presented to CCOHTA as a separate document (*Drug Treatments for Alzheimer's Disease II: A Review of Outcome Measures in Clinical Trials*).

3.2 Donepezil

Donepezil hydrochloride (ARICEPT^{TM}) is an acetylcholinesterase inhibitor specifically developed for the treatment of AD. In contrast to tacrine, donepezil has a high selectivity for central nervous system AChE and lacks peripheral activity as well as the undesirable side effects of peripheral AChE inhibition at least at the doses used in the trials reviewed. The known hepatotoxicity of tacrine is also absent. After administration of donepezil, AChE inhibition has rapid onset and a linear decline permitting once daily administration. Peak plasma concentration is achieved within 3 to 4 hours after administration and the half-life is 70 to 80 hours. Since the level of AChE inhibition in red blood cell (RBC) membranes mimics that of AChE inhibition in the cortex, RBC levels have been used as a marker of the clinical effectiveness of donepezil.

3.2.1 Methods

3.2.1.1 Design

One phase II (16) and three phase III (17-19) parallel, double blind donepezil trials are included in this review. All trials were supported by Eisai America Inc. and Eisai Co. Ltd., Tokyo and several of the authors (common to each trial) were employees of Eisai. The phase II trial, published in 1996, was primarily a dose ranging study designed to explore the potential efficacy and safety of donepezil for the treatment of AD. This trial included 3 treatment arms in addition to

placebo: 1 mg/day, 3mg/day and 5 mg/day. Subsequent trials incorporated two nonplacebo treatment arms: 5mg/day and 10 mg/day. Double blind treatment duration ranged from 12 weeks (16;18) to 24 weeks (17;19) with post treatment placebo washout periods of 2 weeks (16), 3 weeks (18) and 6 weeks (17;19). Multiple study sites in the US (16-18) and internationally (e.g., Canada, Australia, New Zealand, South Africa and Europe) (19) participated. The phase II trial included 161 patients (16) while in the phase III trials the number of patients randomized ranged from 473 (17) to 818 (19). Intention to treat analyses were reported for all trials, using the last observation carried forward approach. The amount of detail reported in each trial with respect to the primary and secondary outcomes was variable rendering direct comparisons difficult.

Inclusion criteria for the four trials were similar. Subjects with a diagnosis of probable AD of mild to moderate severity as assessed by an MMSE score 10-26 and a CDR of 1 or 2 were included. Three of the trials restricted enrollment to those 50 years of age and over (17-19) while the earliest (16) limited the age range to 55-85 years. There were a number of additional exclusion criteria primarily aimed at guarding against inclusion of subjects with other forms of dementia, and those taking concomitant medications with central nervous system effects. At baseline the average age of study subjects across the four trials ranged from 70.6 to 74.6 years of age. The majority of study subjects were women with the percentage of female participants ranging from 52% to 69%. Average MMSE scores as baseline ranged from 18 to 20 and the average ADAS-cog ranged from 25.3 to 29.2. The Burns et al. (19) trial did not provide baseline ADAS-cog scores.

3.2.1.2 Treatment

The phase II trial (16) was of 12-weeks duartion. Donepezil dosages of 1mg/day, 3mg/day and 5 mg/day, given once daily in the evening, were used. The double blind phase was followed by a 2-week placebo washout phase. Efficacy evaluations were carried out at baseline, 1,3,6,9,12 and 14 weeks. The plasma concentration of donepezil was also ascertained at each visit. The phase III trials (17-19) examined the efficacy and safety of 5 and 10 mg daily doses of donepezil versus placebo. The medication was taken once daily in the evening. Study subjects assigned to the 10 mg/day group were started on 5 mg/day for one week before increasing to 10 mg/day for the remainder of the double blind period. Efficacy evaluations were carried out at baseline and then either at 6 week intervals (17;19), or at 3 week intervals (18).

3.2.1.3 Outcome Measures

The primary outcome measure for cognitive performance in all trials was the ADAS-cog. The Clinical Global Impression of Change (CGIC) scale was the primary global outcome measure used in the phase II trial. The CGIC ratings were based on the impression of clinicians who were blind to the patient's performance on psychometric rating scales. The remaining trials used the CIBIC-plus, which is part of the CGIC. The CIBIC-plus includes caregiver-supplied information. All trials used the CDR-SB and the QoL as secondary outcomes measures and all but the Burns et al. (19) trial used the MMSE as a secondary cognitive performance outcome. The earliest trial (16) also included an ADL measure. The International trial group (19) used a modified Interview for Deterioration in Daily living activities in Dementia (IDDD), which is a caregiver completed measure of functional disability. In three of the trials (16-18) AChE activity in RBC membranes was measured to examine the relationship between AChE activity inhibition and donepezil plasma concentration.

3.2.2 Results

3.2.2.1 Drop-outs and Adverse Effects

The following table displays the percentage of completion in relation to trial duration and dosage.

Table 3.1 Completion percentages – Donepezil trials

Study Reference	(16)	(17)	(18)	(19)
Duration	12 weeks	24 weeks	12 weeks	24 weeks
#Patients randomized	161	473	468	818
% completed	Overall: 87.6% Placebo: 87.5% 1 mg: 80.9% 3 mg: 95.0% 5 mg: 87.2%	Overall: 77.8% Placebo: 80.0% 5mg: 85.0% 10 mg: 68.0%	Overall: 88.0% Placebo: 92.8% 5 mg: 89.9% 10 mg: 81.6%	Overall: 77.0% Placebo: 80.0 % 5mg: 74.0% 10 mg: 78.0%

Not surprisingly a larger proportion of randomized subjects completed the 12 week studies (87.6%, 88%) as compared to the 24 week studies (77.8%, 77%). There was, however, no clear relationship between dosage and percentage completion. The groups receiving the highest dosages were least likely to complete the study in two of the trials (17;18).

In general, donepezil was well tolerated and adverse events were reported to be mild to moderate in intensity. The prevalence of adverse events was high even in the placebo groups (65%, 69%, 76%), although this information was not directly retrievable from Rogers et al. (17). Adverse events were most common in the Burns et al. (19) trial, with 76% of placebo subjects reporting adverse events as compared to 79% in the 5 mg group and 86% in the 10 mg group. There were significantly more reports of nausea, diarrhea, vomiting and events involving the nervous system (ex. dizziness or confusion or insomnia) in the 10 mg treatment group (see section 3.10). All of the reported effects are predictable effects of cholinergic drugs and most resolved without a change in treatment. Rogers et al. (17) attributed the high prevalence of adverse effects in their 10 mg/day group to the forced rapid titration schedule (e.g., 5 mg/day for one week followed by 10 mg/day for the remainder of the double blind period). The investigators refer to an unpublished open trial in which dose titration took place over a period of 4-6 weeks and in which these problems were eliminated (ARICEPT^{TM}, Eisai Inc. [donepezil hydrochloride tablets], package insert, Teaneck, NJ). The types of that led to study adverse events drop-out were not described.

3.2.2.2 Efficacy

The primary cognitive outcome used in all of the donepezil trials was the cognitive subscale of the AD Assessment Scale (ADAS-cog). The mean difference between the baseline ADAS-cog and that measured at the endpoint (week 12 or week 24) was computed for each treatment group. The comparison of interest was between each treatment group difference and the placebo group difference. The table below summarizes the results. It should be noted that the method of presentation of results varied across trials. The observed differences in each treatment group needed for the following table were not available from the Burns et al. (19) publication but were provided by Pfizer (personal communication). In others the data were reported as means with ranges (16), or means with SEMs (17;18). For the comparisons with placebo, results were reported as p-values (16), mean differences with 95% confidence intervals (17), and mean differences with SEMs (18).

Table 3.2 Results – Donepezil trials

Study Reference	(16)	(17)	(18)	(19)
Placebo	0.7	1.82 (0.49)	0.4 (0.43)	1.63
Treatment	1 mg. -0.9 3 mg. -1.4 5 mg. -2.5	5 mg. –0.67 (0.51) 10 mg. –1.06 (0.51)	5 mg. –2.1 (0.43) 10 mg. –2.7 (0.43)	5 mg. 0.08* 10 mg. –1.4*
Treatment vs. placebo	1 mg. -1.6, NS[#] 3 mg. -2.1, p<0.05 5 mg. -3.2, p<0.01	5 mg. –2.49, p<0.0001 10 mg. –2.88, p<0.001	5 mg. –2.5 [-3.59 to -1.29] 10 mg. –3.1 [-4.22, -1.92]	5 mg. –1.5, p=0.002 10 mg. –2.9, p<0.0001
Duration	12 weeks	24 weeks	12 weeks	24 weeks
	[#]p>0.05	(SE)	(SE) [95% CI]	*values provided by Pfizer Canada

In two of the trials the subjects displayed, on average, an initial improvement in ADAS-cog scores up to 3 weeks (16;18), and in two trials the improvement was maintained to 6 weeks (17;19). In the phase II trial, at 3 weeks the treatment groups started to decline but remained above baseline scores throughout the double blind period. The placebo group, on the other hand, dropped below the baseline at 9 weeks and continued to decline until the end of the trial. A similar pattern was observed in Rogers et al. (17) and Rogers et al. (18) except that the decline started at 6 weeks. The International trial was the only study in which one of the treated groups (e.g., 5 mg/day) had dropped below the baseline scores at the study endpoint. Not surprisingly the deterioration in the placebo group was less in the 12 week trials than in the 24 week trials. Despite this, the 5mg to placebo differences were similar in the 12-week and 24 week trials i.e. -3.2, -2.49, -2.5 points, all statistically significant. The only exception was Burns et al. (19) where the 5 mg to placebo difference over 24 weeks was -1.5 (p=0.002). There appeared to be little additional benefit to the use of 10 mg/day as compared to 5 mg/day. In the trials with this high dose group, the 10 mg/day: placebo differences were 2.88, 3.1, and 2.9 points, all statistically significant. In an alternative analysis, Rogers et al. (17) reported that there was an increase in the percentage of study subjects who improved by 4 or more points on the ADAS-cog as a function of treatment group. Slightly more than a quarter (26.8%) of the placebo group achieved this outcome as compared to 37.8% of the 5mg/day group and 53.5% of the 10 mg/day group. Whether or not these differences in proportions reached conventional levels of statistical significance was not reported. Furthermore it

should be noted that it was not reported at what point in this study these 4 point improvements were achieved or, when achieved, whether they were maintained.

The impact of donepezil treatment on the overall clinical condition of study subjects was assessed in the phase II trial (16) using the CGIC and all remaining trials used the Clinician's Interview Based Impression of Change with caregiver input (CIBIC-plus). As a measure that directly incorporates change in its scoring, the quantity of interest is the mean score at the endpoint of the trial across treatment groups. The method of reporting the efficacy of donepezil with respect to the global measures differed across trials. In the phase II, 12-week trial (16), Rogers et al. reported that the percentage of patients in each treatment group who were rated as a treatment success (defined as a score <5 on the CGIC, e.g., no change, minimally improved, much improved, very much improved) was 80% in the placebo group, 83% in the 1 mg/day group, 87% in the 3 mg/day group and 89% in the highest dosage (5 mg/day) group. The difference between the 5 mg/day group and the placebo group was statistically significant (p=0.039). In interpreting this finding the authors state that since 12 weeks is a period of relative stability not much change would be expected to be detected on the CGIC, hence the low treatment failure percentage (20%) in the placebo group.

In two of the phase III trials (18;19), treatment efficacy was defined as an improvement on the CIBIC-plus scale (e.g., a score of 1, 2, or 3; very much improved, much improved, or minimally improved). Burns et al. (19) reported that 14% of the placebo group, 21% of the 5 mg group, and 25% of the 10 mg group had improved according to this definition. These figures are not as impressive as those reported in the Rogers et al. (18) trial in which 18% of the placebo group, 32% of the 5 mg/day group and 38% of the 10 mg/day group improved. In each case the difference between the treatment groups and the placebo group was statistically significant. The differences between treatment groups (e.g., 5 mg as compared to 10 mg), were not. Rogers et al. (17) defined treatment efficacy as no change OR improvement on the CIBIC-plus scale (e.g., a score of 1,2,3, or 4). In this trial, 55% of the placebo group, 67% of the 5 mg/day group and 75% of the 10 mg/day group either remained the same or improved on the CIBIC-plus. Differing definitions make comparisons with the other trials difficult but Burns et al. (19) did report the percentage of patients with a CIBIC-plus score of 4 or less and these are all lower than that reported by Rogers et al. (17) (e.g., 49% placebo, 57% 5 mg/day, 63% 10 mg/day).

Using the actual scores on the CIBIC-plus, Rogers et al. (17) found that all patient groups deteriorated on average over the double blind phase and this deterioration was largest in the placebo group. At 24 weeks, the 5 mg/day group and the 10 mg/day group experienced less deterioration than those in the placebo group remaining more or less stable from baseline. This is reflected in the endpoint scores of 4.51, 4.15 and 4.07 in the placebo, 5 mg/day, and 10 mg/day groups respectively. The treatment:placebo difference was statistically significant for both the 5 mg (0.36, p=0.0047) and the 10 mg/day group (0.44, p<0.0001). In the Rogers et al. (18) trial, on the other hand, both the 5 mg and 10 mg/day groups showed an improvement over baseline scores and this improvement was sustained until the end of study. Comparisons with the placebo group were statistically significant (5mg:placebo –0.3 (-0.5,-0.08), 10 mg:placebo –0.4 (-0.55, -0.13)) but there was no apparent benefit of the higher dose. Approximating from the figures in Burns et al. (19), a 5 mg:placebo difference of 0.3 (p=0.0326) and 10 mg:placebo difference of 0.4 (p=0.0009) were observed.

The evidence concerning the efficacy of donepezil using the secondary outcomes was variable. Rogers et al. (16) found no significant treatment effect on the ADL score used. Rogers et al. (17) and Rogers et al. (18) both reported a statistically significant improvement in MMSE scores for those taking 5 mg/day and 10 mg/day relative to the respective placebo groups. Specifically, in the 24 week trial (17) the end of trial differences in changes in MMSE scores were 1.21 points (p=0.0007) and 1.36 points (p=0.0002) for the 5 mg/day and 10 mg/day groups respectively. In the shorter trial (18) the differences were 0.6 points (p=0.004) and 0.9 points (p=001). A statistically significant difference in MMSE score changes was found between the 1 mg and 5 mg groups in the phase II trial (16) but no difference was found when compared to the placebo group. Two of the four trials found no apparent effect on the CDR-SB (16;18), while the two longer trials found evidence of a statistically significant improvement on CDR-SB scores. The lack of empirical evidence concerning the responsiveness to change of this instrument limits our interpretation of these findings. Only one of the trials found a statistically significant effect of donepezil on quality of life. In this trial however, although both the placebo and 5 mg/day groups showed improvement on quality of life, the 10 mg/day group exhibited a worsening accounting for the statistically significant difference between this high dose group and the placebo group.

3.3 Metrifonate

Metrifonate (Promem®) is an organophosphate that has been used in the treatment of schistosomiasis since the early 1960's. The recognition that metrifonate is a cholinesterase inhibitor has led to its study as a potential therapy for AD. Metrifonate taken orally is readily absorbed and rapidly distributed to the brain where rearrangement to 2,2,dichlorovinyl dimethyl phosphate results in irreversible enzyme inhibition and increased acetylcholine levels. It has been shown, as with donepezil, that there is a one-to-one relationship between red blood cell and brain AChE inhibition levels, and that once weekly dosing maintains a steady state of inhibition of red blood cell cholinesterase. A loading dose of metrifonate is needed to ensure that adequate RBC AChE inhibition levels are reached in a reasonable period of time. In the absence of a loading dose, adequate RBC AChE inhibition may not be reached for 6-8 weeks. The clinical benefit of a loading dose has been examined in one study (20). The findings of the study (discussed below) demonstrated that the use of a loading dose regimen did not expedite the onset of clinical benefits.

3.3.1 Methods

3.3.1.1 Design

Eight published studies were reviewed (21) (20;22-27). One study (24) was excluded as it did not meet the minimum requirements for methodological quality on the Jadad scale (8). In addition, the publication by Pettigrew et al. (27) was excluded since it presented only the pharmacokinetics and pharmacodynamics of metrifonate and was therefore outside the focus of the report. Six randomized, double-blind, parallel group clinical trials (21) (22) (20;23;25;26) were retained for in-depth review. These studies emanate from two groups of investigators. The first clinical investigation (an open trial) of metrifonate as a possible treatment for AD was published by Becker et al. (28). The same group, from the Department of Psychiatry at Southern Illinois University School of Medicine, subsequently published a single centre 3 month trial of 38 participants with mild to moderate AD in 1996 (21), and a single centre trial of 50 subjects of 6 months duration in 1998 (26). The principal author (Robert Becker) holds a patent on metrifonate. The two studies were funded by grants from the National Institute of Mental Health and the National Institute of Aging Alzheimer's Center. (25) The second group of investigators, The Metrifonate Study Group, published two trials in 1998 (23;25) and two in 1999 (20;22). All of the trials were multicentre, with recruitments from 264 (22) to 480 (23) subjects with mild to moderate

AD. These studies were funded by Bayer Corporation Pharmaceutical Division and several of the authors are Bayer employees. All of the studies included a two-week screening period but the duration of the double blind phase differed from trial to trial. Cummings et al. (23) was of 12 weeks duration, with a 2-week loading dose phase and a subsequent 10-week maintenance phase. Morris et al. (25) and Raskind et al. (22) were both of 26 weeks duration. The former trial (25) incorporated a 2-week loading phase followed by a 24-week maintenance dose, while the latter trial did not use a loading dose. The final trial from this group (20) was specifically designed to examine the benefits, if any, of a loading dose phase in clinical trials of metrifonate. This trial included 395 subjects and consisted of a 6-week double blind period. With the exception of the latter trial, all others included an 8-week post-treatment follow-up period.

For the most part, reasons for exclusion were based on the desire to omit study subjects with causes of dementia other than AD. For example, all studies required that study subjects score less than 4 on the modified Hachinski Ischemic Scale. This was done to detect those subjects who were suffering from dementia due to vascular causes. Three of the studies applied an age restriction; <90 years (21;26) and 45-90 years (20). All of the studies from the Metrifonate Study Group required that subjects score between 10 and 26 on the MMSE, while the Becker et al. studies required a score greater than 8 (21) or 10 (26). Potential participants with any other possible cause of dementia were systematically excluded from the Metrifonate Study Group trials as were those who did not have a caregiver. In all studies, those taking concomitant medication that may affect cognitive function were also excluded from participating.

Patients recruited to the trials ranged in age from 71.4 to 75 years of age and there were no apparent differences across trials or across treatment groups within trials. With the exception of Becker et al. (26) in which an equal number of males and females were recruited, all of the trials recruited a preponderance of women with the percentage of female participants ranging from 55% to 68%. At baseline, the study subjects were moderately impaired with mean MMSE scores ranging from 18.5 to 19.8. Baseline MMSE scores were not provided in the Cummings et al. (23) trial.

3.3.1.2 Treatment

Each trial included a pre-randomization baseline or screening period either on placebo (20-23;23;25;25) or metrifonate (26). All but one (22) of the trials incorporated a two-phase dosing strategy consisting of a loading dose phase followed by maintenance dosing to achieve a predetermined minimum level of AChE inhibition. The details of the treatment regimen for each of

the Becker et al. (21;26) trials are given in the first table below and the details for the Metrifonate Study Group are given in the subsequent table.

Table 3.3 Treatment regimen – Metrifonate trials

Study Reference	(21)	(26)
Screening period	3 weeks-placebo	4 weeks-metrifonate
Double blind duration	12 weeks	24 weeks
Loading dose	2 weeks (weekly) 5.0 mg/kg 1 week (weekly) 4.9 mg/kg	1 week (once daily) * 2.0 mg/kg for 5 days 0.95 mg/kg on day 6
Maintenance dose	9 weeks (weekly) 2.1 mg/kg	24 weeks (weekly) 2.9 mg/kg
Inhibition goal	40-60%	50-70% after one week
		*loading took place during the 4 week screening period

Table 3.3 – continued

Study reference	(23)	(25)	(22)	(20)
Screening period	2 weeks	2 weeks	2 weeks	2 weeks
Double Blind duration	12 weeks	26 weeks	26 weeks	6 weeks
Loading Dose	2 weeks (once daily) Low: 0.5 mg/kg Mid: 0.9 mg/kg High: 2 mg/kg	2 weeks (once daily) 2 mg/kg	None	2 weeks 100 mg or 150 mg daily based on weight
Maintenance dose	10 weeks (once daily) Low: 0.2 mg/kg Mid: 0.3 mg/kg High: 0.65 mg/kg	24 weeks (once daily) 0.65 mg/kg (e.g., 30 to 60 mg)	26 weeks (once daily) 50 mg	4 weeks (daily) 50 mg NLD* group received 50 mg per day for 6 weeks
Inhibition goal	Low: 30% Mid: 50% High: 70%	70%	Not stated	70%
				*Non loading dose

3.3.1.3 Outcome Measures

All of the metrifonate trials used the ADAS-cog as an outcome measure for assessing changes in cognition. In the trials published by Becker et al. (21;26) this was the primary measure of efficacy. The secondary measures in these two trials were the MMSE, the ADAS-noncog, the

GIS (a modification of the Clinical Global Improvement Scale), and the Activities of Daily Living Checklist (ADLC). It should be noted that no literature was found to describe the psychometric properties of the ADLC and it was only used in these two trials. The ADAS-cog and the CIBIC-plus were identified as primary outcome measures in the studies of Cummings et al. (23) and Morris et al. (25). Secondary outcome measures in these two trials were the MMSE, CIBIS-plus, the Geriatric Evaluation by Relatives Rating Instrument (GERRI), IADL, and the Physical Self-Maintenance Scale (PSMS) (23), the Neuropsychiatric Inventory (NPI), the Disability Assessment in Dementia (DAD), ADAS-noncog and the MMSE (25) respectively. There was no clear indication which of the eight outcome measures used in the trial published by Raskind et al. (22) were primary outcome measures. These measures were the ADAS-cog, MMSE, NPI, ADAS-noncog, DAD, CIBIC-plus, CIBIS-plus and the Global Deterioration Scale (GDS). Jann et al. (20) identified the ADAS-cog at the primary outcome measure and the secondary outcome measures were the MMSE, the ADAS-noncog, the CIBIC-plus and the CIBIS-plus.

3.3.2 Results

3.3.2.1 Drop-outs and Adverse Events

The percentage of patients who completed each of the trials was generally high. In the two Becker et al. studies (21;26), only 3/53 and 1/47 subjects, respectively, who entered the trial did not complete the protocol. Approximately 50% of participants reported adverse events in these two trials. All were reported to be mild and transient in nature and none resulted in any change in treatment. In the 1998 trial (26) adverse events were more frequent in the metrifonate group (14 subjects vs. 8 in the placebo group), whereas in the earlier trial (26) there were equal proportions of subjects with adverse events in each treatment group. For the Metrifonate Study Group trials findings were similar. There was some evidence of an effect of dosage and of duration of trial. The two shorter trials (20;23) had the highest proportion of subjects completing the trial, with between 89% (high dose metrifonate) and 96% (placebo) of subjects completing the 12 week trial (23). The two trials of 26 weeks duration (22;25) had the lowest completion rates with metrifonate:placebo rates of 79%:88% and 82%:84% respectively. The proportion of subjects discontinuing due to adverse events was low in all four trials ranging from 2 to 9% in the placebo groups and from 7 to 12% in the metrifonate treated groups. The adverse effects were primarily gastrointestinal in the placebo groups and of mixed type (e.g. cardiovascular, respiratory, etc.) in the treatment groups. In

the trial reported by Cummings et al. (23), a small percentage of the adverse effects were reported to be serious (e.g., 5% in the placebo group to 7% in the high dose group) requiring modification of dose. More than half of the placebo subjects reported at least one adverse event, and 86% of the high dose group experienced adverse events. There were similar findings in the Morris et al. (25) study. In the Jann et al. (20) trial, there was an indication that subjects assigned to the loading dose group experience more adverse effects than those assigned to the non loading dose. A dose related decrease in heart rate was noted in two trials (23;25).

3.3.2.2 Efficacy

The ADAS-cog was used to assess the efficacy of metrifonate in relation to cognitive performance in all of the trials. All of the trials found a statistically significant effect on the ADAS-cog scale at the end of the double blind phase. The magnitude of the change in ADAS-cog score, however, varied considerably across the trials, largely as a result of the different treatment modalities and the duration of the trials. In general, however, the findings indicated that the placebo group deteriorated over time, while the treated groups either did not change or improved. The average change from baseline to end of treatment is given below for each of the trials.

Table 3.4 Results – Metrifonate trials

Study reference	(21)	(26)
Placebo	1.1	1.67
Treatment group	-0.75	0.0
Treatment vs. placebo	1.86, p<0.01	1.67, p<0.03
Duration	12 weeks	24 weeks

Table 3.4 – continued.

Study reference	(23)	(25)	(22)	(20)
Placebo	Not available	2.5	Not available	0.5*
Treatment group	Not available	N/a	Not available	NLD: 0.0* LD: -1.0*
Treatment vs. placebo	Low: 1.5 [0.18,2.83] Mid: 1.3 [-0.02,2.62] High: 2.94 [1.61,4.27]	2.86 [1.37,4.34]	p=0.012	LD vs. placebo 1.5* p=0.01
Duration	26 weeks	26 weeks	26 weeks	6 weeks
	[95% CI]			*estimated from figures

Given the disparate methods of reporting results in the Metrifonate Study Group trials, it is difficult to make direct comparisons. The trial of Raskind et al. (22) does not report the magnitude of the change over time in the placebo group or the treatment group and provides only the value of the student-t test statistic for evidence of the effect of metrifonate on cognition as measured by the ADAS-cog. It is reported that there was a statistically significant benefit of metrifonate on cognitive performance using ADAS-cog ($p=0.012$). For the trials of similar duration (e.g., 24 or 26 weeks) the placebo:treatment differences in ADAS-cog scores ranged from as low as 1.5 (low dose metrifonate (23)) to as high as 2.94 in the same trial comparing high dose metrifonate to placebo. A very similar difference of 2.86 points (95% CI: 1.37 to 4.34) was observed in the Morris et al. (25) using a single metrifonate treated group with the same treatment regimen as the high dose group of Cummings et al. (23). In the only trial to assess the benefit of including a loading dose phase, there was an average improvement in ADAS-cog of 1.5 points in those subjects receiving the loading dose as compared to the placebo group ($p=0.01$) at 6 weeks (20). There was no difference at either 4 or 6 weeks in the changes in ADAS-cog score between the loading and non-loading dose groups

All of the trials from the Metrifonate study group found a statistically significant effect on the CIBIC-Plus. In the 12-week trial (23), however, there was no statistically significant effect in the low dose group, 0.04 (95% CI: -0.16 to 0.24). Both the placebo group and the low dose group exhibited deterioration in the CIBIC-plus from the baseline. This is in contrast to the mid and high dose groups, where a steady improvement was observed from baseline and the difference between the placebo and treatment groups was 0.29 (95% CI: 0.09 to 0.48) and 0.35 (95% CI:0.15 to 0.54) points on the CIBIC-plus respectively. In Morris et al. (25) a treatment:placebo difference of 0.28 (95% CI:0.06 to 0.50) was reported. In this study, the treated group started to improve early in the course of the trial and then began to deteriorate at approximately 18 weeks. In contrast to the preceding trials, the participants in the Raskind et al. (22) trial showed deterioration in CIBIC-plus scores from baseline. However, the metrifonate group did not decline as dramatically resulting in a 26-week treatment:placebo comparison of 0.2 ($p=0.039$). Jann et al. (20) reported an overall statistically significant positive effect ($p<0.05$) of the non-loading dosage regimen of metrifonate on CIBIC-plus scores at both 4 and 6 weeks after baseline. There was also a statistically significant improvement in the loading dose group relative to placebo at the 6-week point ($p<0.05$). However, the improvement in the loading dose group (0.2, estimated from the graph) was not as large as in the non-loading dose group (0.4, estimated from the graph).

The findings in relation to the secondary outcomes of the MMSE, ADAS-noncog, GIS, ADL, NPI, etc. were inconsistent and generally not statistically significant. A small effect on the MMSE score over 3 months was found in the Cummings et al. trial (23) with an improvement in the low dose group over the placebo group of 1.11 points and in the high dose over the placebo group of 1.37 points. In another study (21), the placebo group deteriorated on average 0.48 points on the Global Improvement Scale as compared to a much smaller deterioration in the treated group of 0.02 points (p=0.02).

3.4 Rivastigmine

Rivastigmine (Exelon®, ENA 713) is a brain selective cholinesterase inhibitor of the carbamate type, that prevents the degradation of acetylcholine in the synaptic cleft through the inhibition of acetyl- and butyryl- cholinesterase, thereby facilitating cholinergic transmission. The AChE inhibition shows brain selectivity particularly for the cortex and hippocampus. It has been shown that acetylcholinesterase is inhibited for as long as 10 hours after the drug has been eliminated from plasma (29). This inhibition is achieved at doses less likely to be associated with peripheral cholinergic effects or organ toxicity.

3.4.1 Methods

3.4.1.1 Design

Three studies were initially judged to be scientifically acceptable (29-31). During the consensus conference, however, the study by Sramek et al. (30) was identified as a safety/tolerability trial with no efficacy results presented. The study was thus excluded from in-depth review.

The objective of both retained trials (29;31) was to demonstrate the efficacy and safety of rivastigmine in the treatment of patients with AD. The ENA 713 B352 Study Group was coordinated by Dr. J. Corey-Bloom at the University of California (29). This was a multicentre study conducted at 22 sites across the United States. The second study (31), the B303 study, was also a multicentre study with 45 sites in both Europe and North America. The two studies were supported by Novartis and included co-authors who were employees of Novartis Pharmaceuticals Corporation in New Jersey (29), or Novartis Pharma in Switzerland (31).

These two trials were placebo controlled, double-blind parallel trials of 26 weeks duration. The outcome measures, chosen to fulfill the FDA's dual efficacy requirement for AD trials, improvement on a performance based cognitive instrument and demonstration of clinical meaningfulness, were the ADAS-cog, the CIBIC-plus and the Progressive Deterioration Scale (PDS). Evaluations for efficacy were performed at 12, 18, and 26 weeks or at early termination. Three types of statistical analyses were carried out: an intent to treat (ITT), a last observation carried forward analysis (LOCF) that included randomized patients with at least one evaluation while being treated and an observed cases (OC) analysis that included all patients with at least one evaluation while on study medication at designated assessment times.

Each of these trials recruited more than 600 patients, 699 in the US trial (29) and 725 in the European/North American trial (31). In contrast to most other trials of other AD therapies, these two studies did not exclude potential participants with concomitant diseases. Only those with severe and unstable medical illnesses were excluded. In addition, only patients taking anticholinergic drugs, acetylcholine precursor health food supplements, memory enhancers, insulin or psychotropic drugs were excluded (31) or were not permitted to continue using those medications (29). Unfortunately, reporting differences hinder direct comparison of patient characteristics across treatment groups. Corey-Bloom et al. (29) presented a large number of patient characteristics at baseline (screening) by treatment group while Rösler et al. (31) only provided baseline information by treatment group for two outcome measures, ADAS-cog and PDS.

The majority of trial participants were women. The European trial reported that 55% of participants overall were female. In the US trial, the proportion of female participants ranged from 57% in the low dose group, to as high as 68% in the high dose group. The participants were on average 72 years of age (range from 45-95) in the European trial (31), and slightly older, 74.5 years of age on average, in the US trial (31), although no range was given. In both trials, approximately 41% of study subjects had mild disease according to the DSM-IV criteria in Corey-Bloom et al. (29) and according to the GDS (32) in the Rösler et al. trial (31).

3.4.1.2 Treatment

In each trial, there were three treatment groups: placebo, a low dose group taking between 1 and 4 mg of rivastigmine per day and a high dose group taking between 6 and 12 mg per day. All study subjects received two capsules twice daily to be taken with food (morning and evening meal), although for the Rösler et al. (31) study this latter information was not stated in the trial report but

was subsequently obtained from a review article (33). There was an initial fixed dose titration phase of 7 weeks in the US trial within which gradually increasing doses were used. By week 7 it was expected that study subjects would be taking at least the minimum dose for their treatment group (e.g., 1 mg/day low dose or 6 mg/day high dose). The next phase (8 to 26 weeks) was a flexible dose maintenance phase in which the doses were further increased until the maximum dose (e.g., 4 mg/day low dose or 12 mg/day high dose) or a maximum tolerated dose was attained. A similar treatment regimen was used in the European trial, whereby fixed dosages increasing weekly in steps of 1.5 mg/day were given over the first 12 weeks of the study with a target dosage to be reached by week 7. During this period, no decreases in dosage were permitted. During the subsequent 13 to 26 weeks a flexible maintenance dosing schedule was applied wherein dosages could be increased or decreased with the goal of administering the maximum tolerated dose within the particular dosage range.

Since both trials used a dosing approach aimed at determining the maximum tolerated dose for each patient, it is prudent to consider the actual dosage received by subjects in each group. The trials were relatively successful in attaining maximal dosage in the low dose group. In the US trial (29), 83% of low dose subjects attained the maximum dosage of 4 mg/day and the average dosage in that group was 3.5 mg/day. Similarly, in the European trial, 90% of subjects attained the maximum dose in the low dose group and the average dosage for that group was 3.7 mg/day. In contrast only 55% (29) and 64% (31) of subjects assigned to the high dosage group reached the maximum dose of 12 mg/day. The average dosages in the high dose groups at study completion were 9.7 mg/day (29) and 10.4 mg/day (31) for the US and European trials respectively.

3.4.1.3 Outcome measures

Efficacy measures for both trials were the ADAS-cog, the CIBIC-plus and the PDS. Assessments were carried out at baseline, 12 weeks, 18 weeks, 26 weeks, or early termination. The first two of these outcome measures have been recommended by the FDA as essential to the assessment of the efficacy of anti-dementia drug products in controlled clinical trials. In both trials these instruments appear to have been used appropriately. The Progressive Deterioration Scale was developed in 1989 (34) to measure quality of life in AD patients. It is scored by the caregiver using a visual analogue scale. Scores range from 0 to 100 with a higher score indicating a better quality of life. The PDS was specifically designed to assess differences and changes in the quality of life of patients with AD, through the eyes of the caregiver. It has been shown to have good reliability and

validity, however, small changes may remain undetected because of large intra-subject variance (35). It would appear as if its use in these two trials was as a measure of functional ability rather than quality of life and the interpretation of improvements in PDS scores as being improvements in functional ability has been questioned (36).

Two instruments, the MMSE and the GDS, were used as staging measures in these trials and subjects were assessed at baseline and at 26 weeks, or early termination.

3.4.2 Results

3.4.2.1 Drop-outs and Adverse Events

The proportion of randomized patients who did not complete the trials was similar in the two studies and appeared to be related to dose. In the placebo group, 16% did not complete the US study as compared to 13% in the European study. The figures were similar in the low dose group with 15% and 14% not completing in the US and European trials respectively. In the high dose group, 35% of randomized subjects did not complete the trial in the US compared to 33% in the European trial. In the US study, 43% of withdrawals were attributed to adverse effects, while in the European trial, the figure was 52%. The figures were only slightly higher for the low dose groups in each trial with 56% of the US withdrawals and 53% of the European trial withdrawals ascribed to adverse effects. In the high dose group, the vast majority of withdrawals (80% in the US trial and 70% in the European trial) were due to adverse effects of the medication. This, combined with the dose related discontinuation, must be considered when interpreting the results of these trials. The authors of the European trial attribute the attrition in the high dose group to the forced dose escalation phase that would not be standard clinical practice. In fact, the goal in the forced dose escalation phase was not have patients reach their maximal tolerated dose.

3.4.2.2 Efficacy

Despite the similarities in methodology, and percentage completion, the results from the two trials differ. In the Corey-Bloom trial (29), both the low and high dose treatment groups showed statistically significant benefits on cognition and global functioning as compared to the placebo group. With respect the ADAS-cog (intent-to-treat analysis), the placebo group was seen to decline significantly over the 26 weeks with an average increase in score of 4.09 points (95% CI: 3.32 to 4.86). The two treatment groups, on the other hand, showed a significantly smaller decline

of 2.36 (95% CI:1.59 to 3.13) points on average in the low dose group and 0.31 (95% CI: -0.46 to 1.08, i.e., data consistent with no change) points on average in the high dose group.

The large decline in the placebo group in this trial may be due to the heterogeneous mix of study subjects recruited and the liberal entry criteria that permitted inclusion of AD subjects with concomitant diseases. Indeed, more than 90% of the recruited subjects at baseline had comorbid conditions and more than 90% were taking other medications. The differences in decline between the placebo group and the high dose group, whether using the intent to treat (difference of 3.78 points, 95% CI: 2.69 to 4.87) or the observed case analysis (difference of 4.95 points, 95% CI: 3.62 to 6.26) were highly statistically significant. The observed difference is more than the pre-specified level of an improvement of 3 points on the ADAS-cog set for efficacy. The results for the low dose group were less impressive with differences between the placebo and low dose group of 1.73 points (intent to treat) and 1.88 (observed case), only the latter being reported as statistically significant ($p<0.05$).

The results from the Rösler et al. (31) trial are more completely reported thereby allowing the reader to assess the differences between the three types of efficacy analyses as well as the differences between each of the treatment groups and placebo.

In contrast to the Corey-Bloom trial, the placebo participants in the Rösler trial deteriorated only 1.34 points on average (95% CI: 0.41 to 2.19) in the intent to treat analysis. Participants in the high dose group improved an average of 0.26 points (95% CI: -1.06 to 0.66, consistent with the null hypothesis of no change) according to the intent treat analysis. For the observed case analysis, while the magnitude of the effect (mean improvement of 1.17 points, 95% CI: -2.27 to 0.07) was different, the overall interpretation is the same with no evidence of change in score (i.e., no decline). Comparing the change in the placebo group to that in the high dose group yielded a statistically significant 1.6 point difference in favor of the high dose group ($p<0.1$, ITT), or 2.58 points ($p<0.001$, OC). For the low dose group, there is evidence of decline (mean decline of 1.37 points ITT, 95% CI: 0.53 to 2.27, mean decline of 1.24 points OC, 95% CI: 0.29 to 2.31). In comparison with the placebo group, neither the ITT analysis nor the OC analysis reveal statistically significant differences between the low dose group and the placebo group. These differences in the results from the two studies are somewhat puzzling. The large decline in the placebo group in the Corey-Bloom study may be attributed to the relatively “sicker” group of subjects recruited as participants, however, identical inclusion criteria were used for the European trial and the

magnitude of the effect is much smaller with no statistically significant effect of the low dose treatment. The heavy losses (primarily due to adverse effects of the medication) in both studies is also troubling and hinders conclusions regarding the efficacy of this medication in relation to cognitive stability or improvement. In the two studies the authors point out that the forced dose escalation phase in which doses were increased until dose-limiting side effects were observed maximized the incidence of adverse effects and does not reflect the procedure that would be followed in clinical practice.

The results relating to the CIBIC-plus were also presented differently in the two trials. In the first (29), the CIBIC-plus was scored as a numeric change from baseline, whereas in the second trial (31) the average score on the CIBIC-plus at the end point of the trial was presented. In the US trial, participants in the high dose group improved steadily from baseline to 18 weeks and then experienced quite a dramatic deterioration. At 26 weeks, the high dose group had an average CIBIC-plus score indicating deterioration. The average change from baseline in this group was 0.2 (95% CI: 0.04 to 0.36) indicating an overall worsening in global function. The low dose group and the placebo group experienced an ongoing global deterioration during the course of the trial, with the two groups being very similar until 18 weeks at which time a decline in the placebo group took place. The decline in the high dose group at 18 weeks resulted in the low dose and high dose groups looking very similar, in terms of average CIBIC-plus score, at 26 weeks. The change from baseline in the low dose group was 0.23 (95% CI: 0.07 to 0.39). The decline in the placebo group is reflected in an average change of 0.49 (95% CI: 0.33 to 0.65) points. Overall the high dose group was found to have statistically significant less deterioration in global function than the placebo group (average difference of -0.29 points with a 95% CI from -0.51 to –0.07). In an alternative analysis, the authors report that there was a significant different in the proportion of patients in each of groups who experienced an improvement in global function (e.g., CIBIC plus score of less than 4 at 26 weeks). This proportion was 24% in the high dose group, 25% in the low dose group and 16% in the placebo group.

There was no figure describing the CIBIC-plus scores over time in the report of the European trial and thus the pattern of changes over time is not available. However, the information given in the report indicates that the high dose group had improved by 26 weeks and that comparisons with the placebo group at each time point were statistically significant and in favor of the high dose treatment. From the ITT analyses we see that the high dose group improved on

average with a score of 3.91 (95% CI: 3.71 to 4.09), whereas both the low dose group and the placebo group deteriorated on average with CIBIC plus scores of 4.24 (95% CI: 4.02 to 4.38) and 4.38 (95% CI: 4.22 to 4.58) respectively. The results from the high dose group indicate that the data are consistent with no change (e.g., the confidence interval include the null value of 4.0). The same alternative analysis was performed as above and 37%, 30%, and 20% of the high dose group, low dose group, and placebo group respectively improved on the CIBIC-plus scale.

The PDS, a measure of quality of life, was used to assess activities of daily living and in both trials there were statistically significant differences between the high dose and the placebo group. Results from the ITT analysis of the US study data reveal that the high dose group declined an average of 1.52 points, the low dose group declined an average of 5.19 points which was similar to the placebo group which declined an average of 4.9 points. The results are also presented in terms of the proportion of subjects in each group who exhibited an improvement of at least 10% on the PDS. It is reported that 25% of the high dose group reached this level of improvement as compared to 15% of the placebo group. Data for the low dose group are not presented.

In the European study, based on the intent-to-treat analysis, the high dose group improved on average (although the confidence interval contained the null value), the low dose group declined an average of 3.37 points and the placebo group declined an average of 2.18 points. 29% of the high dose group was found to have improved 10% or more on the PDS, as compared to 20% of the low dose group and 19% of the placebo group. The effects of high dose rivastigmine were most evident using the last observation carried forward (LOCF) analysis. In this analysis the difference in the proportion of high dose subjects and placebo subjects who improved 10% or more and the PDS was 13% ($p<0.05$). However, given that 32% of the high dose patients did not have a 26 week follow-up as compared to only 13% of placebo patients, the use of LOCF analyses may have resulted in a overestimation of the improvement in the high dose group. The findings are difficult to interpret in light of the range of possible scores on the PDS. There are no empirical data on the clinical significance of a 10% change in PDS score. Furthermore, whether a 10% change in score has the same interpretation across the entire range of the scale is unknown. It was reported that the participants in the European trial had PDS scores of 53.8 to 55.2 on average at baseline. No baseline scores were provided for participants in the US trial.

3.5 Selegiline and Vitamin E

Selegiline: Also called *l*-deprenyl (Eldepryl® in the US), selegiline is a centrally and peripherally acting irreversible monoamine oxidase inhibitor (MAOI). Its pharmacological profile is unusual in that low doses (10mg/day) selectively inhibit MAO-B, the predominant form of MAO in the brain in humans, whereas nonspecific inhibition of MAO-A and MAO-B occurs at higher doses. MAO-B elevation is known to occur naturally in the brain with age and is significantly increased in AD. When too high, this enzyme elevation may represent a reversible functional neurochemical lesion. The administration of selegiline, which inhibits MAO-B, has been reported to reduce agitation, anxiety, and improve episodic learning and memory in some patients with AD. This promising property has been the impetus for several trials.

Vitamin E: It is believed that the pathology of AD may involve oxidative stress and the accumulation of free radicals leading to excessive lipid peroxidation and neuronal degeneration in the brain. Vitamin E (α- tocopherol) is a lipid soluble vitamin that interacts with cell membrane, traps free radicals and interrupts the chain reaction that damages brain cells. In animal models, vitamin E has been shown to reduce the degeneration of hippocampal cells after cerebral ischemia and to enhance motor function after spinal cord injury.

Of the 20 selegiline trials identified, ten scored sufficiently high on the blinded evaluation to be retained for review. However, upon examination of the full articles during the consensus conference, three trials were found to be single-blinded (37-39) and were thus excluded. Seven trials were included in the in-depth review. Only one vitamin E trial was identified (40). This trial, by Sano et al., was also included as a selegiline trial as it included 4 arms: placebo, selegiline, vitamin E, and selegiline+vitamin E. For the purpose of clarity, the selegiline arm of this trial will be discussed in the section on selegiline, and the vitamin E arm of the trial will be included in the section on vitamin E.

3.5.1 Methods

3.5.1.1 Design

Four trials (41) (40;42;43) used a parallel design and three others employed a crossover design (44-46). One of the parallel trials (42) included, after randomization, a 4-week single-blind placebo phase that preceded the double-blind phase. Most studies were conducted in the US with

the exception of one parallel trial, which was carried out in Canada (42) and another in Italy (43). Only two trials were multicentre. In one of the crossover trials (44), patients were recruited from among those who had already completed a multicentre tacrine trial. Patients from the parallel Sano et al. trial (40) were recruited from 23 centres participating in the Alzheimer's Disease Cooperative Study in the US. The other trials were single-centre: In the Canadian parallel trial, subjects were recruited from the Baycrest Centre for Geriatric Care in Toronto as well as from community practices (42). In the parallel study by Burke et al. (41), the clinic from which the patients were recruited was not specified. However, according to the affiliation of the authors, it is most likely have been from the University of Nebraska Medical Center. In the Italian parallel trial, patients were recruited from three neurology clinics, which were all part of the University of Milan (Italy) (43), it is difficult to establish, however, whether it was single- or multicentred. In the crossover study by Lawlor et al. (46), outpatients were recruited from the Mount Sinai Medical Center in New York, while those from the Tariot et al. trial (45) were identified from the Geriatric Neurology and Psychiatry Clinic of a community hospital in Rochester, New York. Three trials were publicly funded (42;44;45), and one was funded by Somerset Pharmaceuticals (41). The crossover study by Lawlor et al. (46) was funded jointly by the National Institutes of Health and Somerset Pharmaceuticals. The Sano et al. trial (40) was publicly funded but medications were supplied by Somerset Pharmaceuticals.

The duration of the trials varied widely. The parallel trials were of 3 months (24), 6 months (23), 15 months (22) and 2 years (40) respectively, while crossover trials tended to be of shorter duration: 8 weeks (25), 14 weeks (27), 20 weeks (26).

The study population was similar across all studies reviewed. The majority of trials excluded patients older than 80 years of age; resulting in a mean age of patients between 70 and 75 years. Such a restriction was not applied in the smallest crossover trial (44), however, where the mean age of patients was much higher (89 years). While all studies used the standard clinical criteria for the ascertainment of probable AD, some trials also considered the severity of dementia as an inclusion criterion, using either the CDR or the GDS scale. Studies involved patients with mild dementia (CDR=1) (41) or moderate dementia (CDR=2) (40). Another trial was also restricted to moderate dementia but CDR scores were not reported (45). These scores were probably selected to provide samples of AD patients that would maximize the chance of detecting any true treatment effects. Rubin et al. (47) pointed out that successful therapies for Alzheimer's

Disease are more likely to have an effect at milder stages of the disorder, so it would seem logical to exclude potential subjects with severe dementia (CDR 3). Also, the usefulness of enrolling subjects with a CDR of 0.5 has been called into question because these individuals sometimes have less predictable patterns of disease stage advancement and may not have dementia in the first place. Nevertheless, the Italian parallel trial by Mangoni et al. also included patients with severe dementia (GDS=3,4,5) (43). The problem with the GDS scale as a staging instrument is that it can misstate a patient's disease severity. It has been found that many of the descriptors of cognitive impairment for the seven GDS stages do not adequately reflect the actual clinical events that occur within those stages (32;48). The ability to enrol desired patients could be threatened if the GDS misidentifies the stage of dementia. If the trial includes patients who were not as demented as desirable, the variance of the results could be increased. As well, excluding patients who should be included has implications for generalizability. Because of these problems, the GDS should not be used to stage dementia in AD drug trials. The severity of disease was not reported in the three other studies.

Patients with a history of other conditions that might be responsible for cognitive impairment were excluded from all trials. Examples included head trauma, alcoholism, central nervous system disorders, epilepsy and a Hachinski score of greater than 2 (or 4 depending on the study). Most trials (44) (40;45;46) excluded patients who were using psychotropic medications. One parallel study (43) also excluded patients taking "selected medications" or those who had psychiatric disorders (which would also be associated with the use of psychotropics) (42). In light of the extensive use of psychotropic medications in the elderly population, such exclusion greatly limits the generalizability of the study results. Three of the trials required that patients have reliable caregivers (45) (40;46) and one (46) required that patients had exhibited behavioral disturbances in the form of agitation, anxiety, aggression, depression, psychosis, or sleep disturbance for at least 4 weeks. In addition, an important feature of the smallest trial (a crossover trial with 10 patients) (44) is that only patients who had already completed a tacrine trial were eligible.

Most of the studies were single-centered and this resulted in sample sizes that were relatively small: from 39 to 119 for parallel trials, and 10 to 50 for the crossover trials. The multi-centre trial by Sano et al. (40) included 341 patients but because of the four arms, the number of patients randomized to selegiline and placebo was smaller: a total of 171 patients in the two arms.

In the parallel trial by Burke et al. (41), a repeated measures analysis of variance was used to determine the drug effect using four assessments: baseline, 2, 8 and 15 months. The cutoff for statistical significance was not specified. In the Freedman et al. (42) trial, the main analysis was conducted on the BPRS scores using only the baseline and post-treatment assessments (average between visits 6 and 7, which corresponds to week 21 and 25, respectively). Paired t-test were used to assess the change in total BPRS score, and t-test for independent samples were used to assess the effect of selegiline on the change. Secondary and exploratory analyses were conducted using repeated measures analysis of variance on data obtained at each of the 7 visits. A cutoff of 0.05 was used on the primary analysis. In the Mangoni et al. trial (43), analyses were conducted using repeated measures analysis of variance on data obtained at baseline, first, second and third month of trial. In the Lawlor crossover trial (46), the primary analyses to evaluate the net drug effect on the BPRS, DMAS and ADAS-Cog were based on a single assessment at the end of each 6-week treatment period using a single sample t-test. The cut-off for statistical significance was 0.017. Secondary analyses involved repeated measures analyses of variance conducted on the behavior measures (BPRS and DMAS). In the small crossover trial, which involved patients who had completed a prior tacrine trial, a non-parametric signed rank test was used to compare each subject's performance while taking selegiline and while taking placebo. In addition, change scores were analyzed only for the comparison of the group that first received selegiline with the group that first received placebo). In the Tariot et al. crossover trial (45), patients were evaluated at baseline as well as monthly for 2 months during each treatment phase. Repeated measures analyses of variance were used to assess drug effect and the cut-off for statistical significance was 0.05. Surprisingly in this trial, dropout subjects were not analyzed. In study by Sano et al. (40) because the primary outcome was an event (dichotomous variable), a survival analysis was used to assess the risk of having an event on selegiline relative to the risk on the placebo (relative risk). This method also allowed the assessment of whether or not selegiline increased survival time as compared to patients receiving placebo.

The sex distribution was similar across comparison groups, with the exception of two studies: the Burke et al. parallel trial (41) had a larger proportion of men in the group receiving selegiline than in the placebo group. Also, in the Lawlor et al. crossover trial (46), the number of women was greater (14 vs. 6). When reported, the level of education varied little across studies, with a mean of 13 years of schooling (41;44;45). One study included patients with lower levels of

education (mean of 6 years) (43). Other baseline variables were similar across comparison groups. In the Sano et al. (40) trial, the mean MMSE score differed across study arms but this difference was taken into account during data analyses.

3.5.1.2 Treatment

The treatment dosage was always 10 mg/day and all trials used placebo comparisons. In the Canadian parallel trial, there was a single-blind placebo period of 4 weeks prior to the double-blind phase (42). In addition, during the initial one week of treatment, a dosage of 5 mg/day was administered before increasing the dosage to 10 mg/day. Following the double blind period (25 weeks), patients were placed on a single-blind placebo until 41 weeks.

3.5.1.3 Outcome measures

All studies used instruments to assess the effect of selegiline on one or more of the following domains: global assessment, cognition, functional status/quality of life, behavior and mood. Only four selegiline studies specified primary outcome measures (42) (40;44;46). In the Freedman et al. study (42), the BPRS, which belongs to the behavior and mood domain, was used while in the Lawlor et al. study (46) three primary outcomes were used: one cognitive scale (the ADAS-Cog), and two behavior scales (the BPRS and the DMAS). The only primary outcome measure used in the Schneider et al. trial (44) was a cognitive scale (ADAS-cog). The Sano et al. trial (40) differs from the other studies in that it focused on the functional status by using an event as primary outcome. The patient was considered to have the event if one of the following occurred: death, institutionalization, loss of the ability to perform the activities of daily living, or severe dementia (CDR=3). Scales were also used in the latter study but as secondary outcome measures. The scales used in each of the outcome domains are listed below:

Global assessment: Three selegiline trials did not include a global assessment scale as an outcome measure (43;44;46). The others used the CDR or CDR-SB scales (40;41;45) or, in one instance, the GDS scale (42).

Cognition: All trials used a cognitive scale and/or cognitive tests. The most frequently used was the ADAS-Cog and the MMSE. One trial did not include a cognition *scale* as an outcome measure (45). Instead a variety of cognition *tests* were used. Three trials combined cognition scales and cognition tests (41-43).

Functional status/Quality of life: Four trials did not include a functional status/quality of life scale (41;42;44;46). Others used the BDS-part I (43;45) or the modified BDS (40). Trials also included the IADL (45) or the DS scale (40).

Behavior and mood scale: With the exception of one trial (44), all included at least one behavior and mood scale, the most common being the BPRS (41;42;45;46).

3.5.2 Results

3.5.2.1 Drop-outs and Adverse events

Selegiline appeared to be well tolerated apart from intermittent mild nausea in one study (44) (this study involved patients who had completed a tacrine trial, and this may have influenced the occurrence of side effects either by a carry-over effect or simply greater awareness). In any event, no adverse effects led to withdrawals in any of the selegiline studies. The Sano et al. study (40), which also included a vitamin E arm, reported different results. Forty-nine categories of adverse events were defined a priori and there were significant differences among the groups in three categories: dental events (event that required dental treatment), falls and syncopal episodes. However, when adjustment for multiple comparisons was used, the differences were no longer statistically significant.

In the Sano et al. trial (40), 5% of the patients in the selegiline arm dropped out compared to 7% in the placebo but the number of withdrawals due to adverse events was not reported. There were also drop-outs in some of the other studies but these were not the result of adverse effects (41;42;46). Stated reasons in these cases were a perceived lack of efficacy (43;45) or inability on the patient's part to complete the trial (46). In the crossover trials, it was not stated during which phase the dropouts occurred (45;46). Additionally, some patients dropped out during the first few weeks of the trial and therefore did not have any measurements. These subjects were not included in subsequent analyses. The exclusion of these patients from the denominator affected the percentage completion only slightly given that the number of such patients was small.

3.5.2.2 Efficacy

In the trial by Sano et al. (40), selegiline did not appear to significantly reduce the risk of having one of the following events during the 2-year period: death, institutionalization, severe dementia, loss of ability to perform the activities of daily living (relative risk = 0.72, p = 0.087). However when analyses were adjusted for the differences in baseline MMSE scores, a significant

delay was found when all events were taken together: relative risk=0.57, p=0.012. It was estimated that selegiline increased the median survival (e.g., survival until the occurrence of a primary outcome) by 215 days compared to placebo. Analysis by endpoint showed that: 28% of selegiline patients and 31% of placebo had lost the ability to perform activities of daily living, 33% and 39%, respectively were institutionalized, 10% and 12%, respectively had died. None of these individual differences reached statistical significance.

The CDR was used primarily as an inclusion criterion in order to stage the severity of dementia (see above section on inclusion/exclusion criteria). However, it was also used in two trials as an outcome measure to assess the efficacy of selegiline. This scale measures cognitive performance in six areas: memory, orientation, judgment/problem solving, community affairs, home/hobbies and personal care. Two scoring methods were used: the original CDR score (varying from 0 to 3, with 3 being the most severe dementia), and the CDR-SB (the ratings in all six performance categories are summed to obtain a global dementia ranking that can go no higher than 18 points). The latter scoring method is preferred by some as it provides a more detailed and quantitative picture of a subject's severity of dementia (49). The Burke et al. parallel trial (41) used the CDR scale as well as the CDR-SB outcome measures. Over a period of 15 months patients on selegiline deteriorated from a mean CDR score of 1.0 at baseline (mild dementia) to a mean score of 1.3 (because it is still below a score of 2, the dementia is still considered mild). Although this trend was statistically significant, it did not differ from that observed in the placebo group, indicating that selegiline does not influence changes in the severity of dementia. Similarly, the CDR-SB scores increased significantly over the follow-up period (from a mean of 6.2 to a mean of 7.8 in the selegiline arm of the trial). However, this trend was not different from that observed in the placebo arm, which indicates that selegiline does not slow down the deterioration of dementia. The 25 interviews involving patients and caregivers were videotaped. Patients were scored by a geriatric psychiatrist at the time of videotaping. Another clinician, blinded to these first results, saw the tapes later and assigned a second set of ratings. Weighted kappas were used to measure the two clinicians' agreement. These kappas ranged from 0.75-0.94 (indicating excellent agreement). The crossover trial by Tariot et al. (45) used the CDR-SB only. Again the treatment had no effect on the scores over a period of 4 weeks. In the Sano et al. trial (40) patients had moderate dementia at baseline (CDR=2) and the outcome was an event, e.g., to have a severe dementia (CDR=3) at the end of the 2-year follow-up. In this study, 51% of patients included in the placebo group reached a

CDR of 3 compared to 43% in the selegiline group. This difference was not statistically significant.

The CDR is useful because: i) it provides physicians with a global rating that encompasses a broad range of patient characteristics; ii) it focuses on cognition, and not on items which may be related to other medical, emotional or social conditions (50). However, although it has been developed specifically in the setting of a longitudinal study on AD patients (49), it is a staging scale to assess the severity of dementia at one point in time, most often at baseline. Its sensitivity to change over time has not been documented. Therefore, it may not be an appropriate scale to assess changes in dementia severity over time. Furthermore, it tends to measure cognitive aspects of dementia, and not a patient's global health state with respect to AD.

The GDS scale was used as an outcome measure in the Canadian parallel trial by Freedman et al. (42). GDS scores did not change significantly from baseline to 25 weeks in the selegiline arm of the trial (mean of 4.3 and 4.4, respectively at baseline and post-treatment), nor in the placebo arm (mean of 3.9 and 4.0, respectively). Like the CDR, the GDS has been developed primarily as a staging instrument and its sensitivity to change has not been evaluated. It is intended to measure the progressive stages of cognitive impairment. However, it can misstate a patient's disease severity because many of the descriptors of cognitive impairment for the seven GDS stages do not adequately reflect the actual clinical events that occur within those stages (32;48).

The two crossover selegiline studies, which used the ADAS-cog scale, conducted statistical analyses on only 10 patients (46) (44). Yet, a statistically significant effect was found. In the Lawlor et al. trial, 10 out of the 20 subjects did not complete the ADAS-cog scale due to missed visits, language barrier or severity of cognitive/behavioral impairment. On this 11-item scale, which scores range from 0 to 70 (the higher the score the greater is the dysfunction), a statistically significant reduction of 10.8 points on the scale over a period of 6 weeks was found in the Lawlor et al. trial. In comparison, the magnitude of the reduction was 10-fold smaller in the Schneider et al. trial: -2.50 over a period of 4 weeks. Given that a 4-point change over a period of 6 months would correspond to what the FDA considers to be a clinically important change (51), one can assume that the results of the former trial are clinically significant. It is more difficult, however, to determine whether the results of the second trial are significant because these changes occurred over a period of 4 weeks and not 6 months. In addition, according to Stern (52), the severity of dementia at baseline should be taken into account. In the Lawlor et al. study, the severity of

dementia at inclusion was not reported, other than the patients exhibited behavioral symptoms that were distressing to the patient and the primary caregiver. In the Sano et al. parallel trial (40), the ADAS-cog was also used as a secondary outcome measure. Over a period of 2 years, patients had deteriorated regardless of treatment: the mean change in score was +8.3 for the selegiline patients and +6.7 for the placebo, a difference that was not statistically significant.

Four selegiline trials used the MMSE scale to assess cognitive performance but none employed it as a primary outcome measure (40-42;44). In the Schneider et al. crossover study (44), the baseline score was 15.7 and there was no significant improvement on the MMSE score over the period of 4 weeks (in fact there was a small numerical worsening). In the Burke et al. parallel study (41), the repeated measures analysis of variance indicated a significant decline in the MMSE score (which indicates worse performance) over a period of 15 months among patients who had a mean score of 18.8 at baseline. However, the rate of decline was not modified by selegiline. In the Sano et al. trial (40) patients declined by 5.1 points and 4.6 points, respectively in the selegiline and placebo arms from scores of 12.7 and 13.3, respectively at baseline. The difference was not statistically significant between the two groups. In the Canadian Freedman et al. study (45), baseline scores were 17.3 and 18.4, respectively in the selegiline and placebo groups (the difference was not statistically significant). There was no treatment effect because post-treatment values were almost identical in both groups. The usefulness of the MMSE as an outcome measure to assess the efficacy of selegiline is questionable for two reasons: i) Patients in the Sano et al. (40) and Burke et al. (44) studies had severe cognitive impairment at baseline: their mean score was below the cut-off of 17 which refers to severe cognitive impairment. The MMSE has been found to lose its responsiveness to change as the severity of dementia increases ("the floor effect") (53)*;* ii) the ability of the MMSE to detect changes over a short term has not been demonstrated because of its great variability across subjects (54). Although the MMSE shows good reliability and validity for its original purpose, that of screening for dementia, it was not designed to measure more subtle aspects of cognition. It appears that this could happen even in subjects who would otherwise be shown to have declined substantially if another scale had been used (54;55).

The Italian parallel trial used the BDS as its cognitive scale (although they also used a variety of individual cognitive tests) (43). There was a significant improvement on all items of the BDS scale over a period of 3 months (daily living, information/orientation, memory, concentration, and total score). However, the validity of this scale as evidence of drug efficacy is greatly

jeopardized by the fact that there is a low interrater reliability (56) and the results of validity testing have been inconclusive (57;58). Furthermore, its responsiveness to change has not been determined.

Four of the selegiline trials also included a variety of individual neuropsychological tests as secondary outcome measures (41-43) (45). In the crossover study by Tariot et al. (45), none of the 5 tests used showed significant improvement after 8 weeks (Selective Reminding Task, category retrieval, trails A and B, finger-tapping speed, choice reaction time). The Burke et al. trial (42) used a total of 18 tests to assess attention and concentration, motor, language, and memory domains. Although there was a significant change between baseline and 15 months in 6 of these tests (Trail-Making test Trails A, the Symbol Digit Modalities, Oral raw score, orientation component of the WMS-R, logical memory part I, paired associates part II), there was no treatment effect. The usefulness of these tests to monitor the progression of dementia and their ability to detect a treatment effect appears questionable. An extensive study (59) has shown that the tests of recent information recall were best for discriminating intact from demented subjects or detecting mild from more severe dementia. However, all tests have a poor ability to separate stages of AD. Only those that show a linear decline are best at following the course of disease. Consequently, the memory tests, such as the WMS, are not useful in longitudinal studies of cognition in AD because of floor effects early in the course of disease. However, tests of work fluency and visual recall should be included as outcome measures in trials.

In the Mangoni et al. trial (43), on the other hand, there was a significant treatment effect on all 5 tests used. After 90 days of treatment, the results of the Digit span, WMS, Toulouse-Piéron, Word fluency, drawing tests improved in the selegiline arm of the trial while they deteriorated in the placebo arm. Although, such results appear to confer a positive advantage to selegiline, these tests are not comprehensive enough to assess changes in the cognitive performance of patients. It was concluded that global tests are more appropriate (59).

Three trials included a functional status scale (40;43;45): none was used as a primary outcome and scales were heterogeneous across studies.

The Blessed Dementia Scale (BDS) part I was used as a secondary outcome measure in two trials (43;45) and a modified version, the mBDS was used in the Sano et al. trial (40). The scale was designed to measure the degree of intellectual and personality deterioration shown by the elderly. It includes two components: an evaluation of activities of daily living and changes in

personality, interests and drives, and an assessment of orientation, memory, and concentration. The higher the score the greater the impairment. In the Tariot et al. trial (45), selegiline was responsible for a 0.3 point change in BDS score over a period of 8 weeks. The baseline mean score was 4.2 and this effect was not statistically significant. In the Mangoni et al. parallel trial (43), patients in the selegiline group had a lesser level of dependence after 90 days (a mean score reduction of 1.9) while in the placebo the score actually increased by a mean of 0.45 points. This difference was statistically significant. In the Sano et al. trial (40), patients on selegiline had a higher score after 2 years (which indicates a greater level of dependence), but the deterioration was smaller than that observed in the placebo group (mean change of mean score of 4.2 and 5.4, respectively which was statistically significant). It is worth noting the BDS Part 1 specifically addresses a population of demented elderly, which is relevant for its use within AD clinical drug trials. However, there is a lack of evidence as to its sensitivity in detecting small changes that may have occurred following specific interventions.

The IADL, which was used in the Tariot et al. crossover study (45), was designed to measure elderly persons' competence in instrumental self-maintenance, that is, the complex activities required in everyday functioning. Scores range from 0 to 8 for women and from 0 to 5 for men. The authors stated that there was no difference in any of the measures that were used. IADL scores were not reported. The Dependence Scale (DS) is a 7-point scale stages the dependence level of patients, in terms of their need for help in performing occupational function. It was used in the Sano et al. trial (40). However, mean scores were not provided. Results were expressed as percentage of patients receiving higher score (the cut-off to define higher score was not specified). 80% of patients in the selegiline group were included in the higher score group (most dependent) compared to 86% in the placebo group. This difference was not significant.

All trials, except the Schneider al. (44) crossover trial included at least one behavior and mood scale. Only two (42;46) considered these as a primary outcome. The Brief Psychiatric Rating Scale (BPRS) was the most frequently used. In the Freedman et al. trial (42), where it was used as a primary outcome measure, there were no statistically significant differences between the selegiline and placebo groups on any behavioral measures after 2 months of treatment or more. The change in BPRS was also not significant during a comparable time period of 8 weeks in the Tariot et al. crossover trial (45). In the population of patients who had behavior disturbance (46), there was no significant net drug effect after 6 weeks in the primary analysis which involved a

single sample t-test, with and without intent-to-treat analysis. However, when the repeated measures analysis of variance was used on the three biweekly observations in the secondary analyses, there was a trend towards significant improvement ($p=0.06$). Nevertheless, these negative results do not support in themselves a lack of drug efficacy. In fact, it was determined that the BPRS does not cover the whole range of behavior and mood problems seen in AD because it has been developed for more general purposes (60;61). Direct patient interview, which is featured in the BPRS, may lead to inadequate behavior assessment since the patient is often unable to provide an accurate report; caregiver interview is preferable.

The Freedman et al. study (42) also used the ADAS-noncog to evaluate depression, concentration, cooperation, psychotic disturbances and motor activity among AD patients and no significant drug effect was found. The ADAS was developed to measure changes specifically in AD patients following pharmacologic treatment. This scale, which has been well evaluated in the literature, shows adequate reliability and validity. However, there is evidence that the scale is less reliable when used by nonclinicians, such as is the case in the Freedman et al. study (42). It is therefore difficult to determine whether absence of statistical significance is attributed to lack of drug efficacy or improper evaluator.

The Relative's Assessment of Global Symptomatology (RAGS), a scale intended to assess patients' behavior in the community, was used in the Tariot et al. (45) crossover trial as well as in the Freedman et al. trial (45). In the former trial, no significant change in mean value was observed following 8 weeks of treatment. Similar findings were found after 25 weeks of treatment in the Freedman et al. trial.

Two of the parallel trials included a scale to assess depression among patients (41;42). The Cornell Scale for Depression in Dementia (CSDD) is a 19-item scale, which is rated primarily on the basis of observation. The evaluation is conducted by a clinician and involves two phases: a caregiver's interview followed by a patient's interview. Scoring ranges from 0 to 57, a higher score indicating a more severe depression. The CSDD is unique in that it has been validated for patients with dementia (60). After about 25 weeks of treatment, there was no difference between baseline and post-treatment mean CSDD score in either the selegiline or placebo group (42). Analyses on repeated measures over 15 months did not show any trend in depression severity in neither the treatment nor the placebo group (41).

The Sano et al. trial (40) used the CERAD Behavior Rating Scale (C-BRSD) as a secondary outcome measure. This scale was developed by the Consortium to Establish a Registry for Alzheimer's Disease to assess behavior in AD patients through interviews with primary caregivers (62). This scale contains 48 items that can be grouped into 8 factors (depressive symptoms, psychotic symptoms, defective self-regulation, irritability/agitation, vegetative features, apathy, aggression and affective lability. Although there are no recommendations made on scoring the C-BRSD in the development article by Tariot et al. (63) one can either sum the frequency ratings across items or count the number of behavioral problems endorsed. The scoring method was not specified in the Sano et al. paper but in either case a higher score indicates more behavior problems. In both the selegiline and placebo groups, patients deteriorated but the mean increase in C-BRSD score was different in the two groups: 5.4 and 8.9, respectively for selegiline and placebo (the difference was not statistically significant).

To attempt to determine the effect of selegiline on mood and behavior is difficult because of the heterogeneity of scales that were used in the trials. By including widely varying items in their definition, the different scales are not easily comparable. The BPRS, for example, does not adequately cover an appropriate range of problems in AD. The ADAS-noncog (64-66) includes items outside the range of the domain, but does not include others important to AD. With the exception of the Lawlor et al. (46) trial, most trials excluded patients based on behavior and mood criteria and then included the domain as a secondary outcome. These RCTs are ill suited to assess an intervention's effect on behavior and mood symptoms.

Overall, selegiline did not have a significant effect on any of the four outcome domains: global assessment of dementia, cognition scales and tests, functional status/quality of life, behavior and mood. Four studies used the MMSE as a criterion for efficacy and none found a statistically significant difference between selegiline and placebo (40-42;44). In one study (41), disorientation was the only item that differed between the two groups, but there was no effect on the MMSE. In another study, there was a marginally significant ($p=0.08$) mean improvement of 2.5 points on the ADAS-cog (44). One other trial (46) found a significant difference of 10.8 (46). In the latter study, a difference of 5.6 points was found on the DMAS and of 2.2 on the Brief Psychiatric Rating Scale (BPRS). Neither of these results was statistically significant, but the study included only 20 patients and may have lacked the power to detect (as statistically significant) differences of this

magnitude. Overall it was concluded that there was no association between the ADAS-cog and the change in behavior in patients who were testable.

It is interesting to note that all three excluded studies (37-39) reported statistically significant differences on the majority of scales used. This stresses the importance of carrying out double-blind trials with rigorous methodology.

An issue of concern in the conduct of crossover trials is the possibility of carryover effects. One of the trials (44) had no washout period, one had a 2-week washout period (46) and a second had a 4-week washout period (45). In the former trial, a post hoc analysis was conducted to remove the carryover effects of the crossover design by comparing the two groups prior to the cross over. This was, then, analogous to a parallel design of 6 weeks (n=10). A statistically significant difference was found on the BPRS (e.g., 2.1 points improvement on selegiline versus 3.2 points decline on placebo). In the latter trial (45), no specific analysis was carried out to determine if there was a carryover effect.

3.5.3 Alpha-tocopherol (vitamin E)

The Sano et al. trial (40) described in the previous section was the only vitamin E trial. While its characteristics have already been described, the following discussion will focus on the vitamin E, placebo, and the vitamin E + selegiline arms.

3.5.3.1 Treatment

The α-tocopherol group received a racemic mixture of *dl*-α-tocopherol given in a dose of 1000 IU twice a day (morning and afternoon). Compliance with the medication was measured through the urine and a level of 2mg/deciliter or higher in 75% of samples from an individual patient was defined as positive for α-tocopherol.

3.5.3.2 Results

At baseline, the vitamin E group scored lower on the MMSE than the placebo group (11.3 vs. 13.3; p=0.07). Since the MMSE score was also found to be associated with the primary outcome, it was included in the models as a confounder. For the primary outcome “any of the events of interest”, there was a statistically significant delay in occurrence of the outcome in the treated group with a relative risk of 0.47 (p=0.001). Looking at each outcome separately, the finding was narrowed down to a statistically significant delay in the time to institutionalization in

those subjects assigned to the α-tocopherol group (RR=0.42, p=0.003). There was no apparent effect of treatment with α-tocopherol on the MMSE, EIS, BRS or UPD. However, the placebo group scored higher on the BDS (5.4 vs. 4.0 p=0.004) and required more care and supervision as measured by the DS (86% vs. 76% requiring assistance with moving, turning, eating or using the toilet, p=0.04).

The combination of vitamin E and selegiline did not confer any additional advantage. In fact, the combination appeared to reduce the beneficial effect of either vitamin E or selegiline alone in delaying the occurrence of the primary outcome : relative risk of 0.69, p=0.049. On the other hand, there was a higher frequency of falls (22% versus 5% in the placebo, p<.005), and syncope (16% and 4%, respectively). With respect to the secondary outcome measures, the effect of the combination was essentially the same as that with vitamin E alone : the mean change score on the Blessed Dementia Scale was the same, also significantly different from the placebo, and the percentage of patients which were most dependent was equally reduced. The only additional benefit was observed with the behavior rating scale for dementia (BRSD). While the score increased in all other treatment arms, including placebo, there was a significant reduction in mean BDS score of 1.1 (p=0.02).

3.6 Lecithin

Lecithins are phospholipids involved in cell membrane constitution and fat metabolism by the liver. Lecithin is used in AD mainly as a choline supplement, and then usually only in tandem with other active interventions, such as tacrine. Cholinesterase inhibitors are believed to deplete stores of choline, thus there may be a need for concurrent supplementation (67) (68). Early investigations of lecithin's effect were, for the most part, inconclusive and methodically weak (69). The present discussion focuses on results extracted from tacrine and lecithin combination clinical trials. The ideal comparison for studying the effect of lecithin would be trials of lecithin versus placebo or, as a secondary and less preferable comparison, trials of tacrine/lecithin combination versus tacrine alone.

3.6.1 Methods

3.6.1.1 Design

Four trials were selected following blinded review. On subsequent examination, three trials were excluded because either there appeared to be no control group (70) or the reported results were not sufficient to permit comment on lecithin's effect (71;72). The only lecithin study to be included for further review was a crossover trial of combination lecithin and tacrine (73). The study was an eight-week multicentre trial of 440 AD subjects. Participants were required to have no history of stroke, to have no concurrent use of centrally acting drugs, and were (descriptively) in good general health. Analysis was with ANOVA, ANCOVA, and Cochran-Mantel-Haenszel methods, but it is unclear whether an intent-to-treat approach was adopted.

Baseline characteristics were reported for all subjects, not solely for those who were randomized. The final sample to be randomized was a highly select 122 subjects -- the tacrine responders -- for whom no characteristics were reported other than (an indication of) similar sex and age distributions relative to the original sample of 440 subjects. These 122 subjects had higher ADAS and lower MMSE scores. In other words, the randomized subjects were those who responded on a cognition scale to tacrine during the titration phase, and were more cognitively impaired than the original sample. The original sample ranged in age from 50 to 89 years, with a mean age of 69.6 (SD 8.3) years. Nearly two thirds (62%) were male. At baseline screening, the average MMSE was 17.8 (SD 4.2 to 4.6).

3.6.1.2 Treatment

The design include of a 6-week dose titration phase (to determine a "best dose"), a 2-week placebo baseline period, and two 4-week treatment phases (comprising the crossover).
Tacrine was given in either 40 or 80 mg/day doses (in four capsules). The doses chosen for the double-blind phase were those associated with at least a four-point improvement on the total ADAS (e.g., the "best dose" as chosen in the titration phase). Lecithin was administered at 9.04 g/day (in four capsules).

3.6.1.3 Outcome Measures

Primary measures were the ADAS-cog and the CGIC; the ADAS-noncog was used as a secondary measure.

3.6.2 Results

3.6.2.1 Dropouts and Adverse Effects

Nearly 2/3 of participants did not achieve a best dose. In addition, prior to randomization, a further 26 (6%) subjects were lost during the placebo baseline phase. The remaining 122 subjects were assigned to tacrine plus lecithin, tacrine, or lecithin treatment arms. Adverse events ascribed to lecithin were not reported. Of the 122 subjects randomized, 4 (3.4%) withdrew during the first crossover phase, and 112 subjects completed both phases of the study. The distribution of dropouts across the treatment phases was not reported.

3.6.2.2 Efficacy

Results from the second phase of the crossover were not reported. The first phase of the study provided two possible comparisons for lecithin. Four weeks of lecithin use compared to baseline placebo resulted in small increases in the MMSE and the ADAS-cog, -noncog, as well as in the ADAS-total (although not statistically significant). The ADAS-cog baseline was 27.3 (SD 1.7) and the four-week difference was –1.3 (SD 0.8). However, on average, subjects showed minor improvement in all treatment groups. The combination of lecithin and tacrine compared (in parallel) to tacrine alone did not show any quantifiable improvement on outcome scores. In general, the addition of lecithin to tacrine did not appreciably alter the proportion of subjects who improved or worsened according to the CGIC. Treatment with lecithin alone resulted in less improvement and more worsening than the other two treatment arms (percent improved: lecithin 35%; tacrine 54%; tacrine plus lecithin 46% - percent worsened: lecithin 24%; tacrine 11%; tacrine plus lecithin 6%). There were three treatment arms in the first phase of this study, none of which was placebo alone. While the above discussion focused on lecithin at week four versus baseline, and lecithin/tacrine versus tacrine alone, there is no placebo (e.g., no active treatment) for contrast at week four.

Table 3.6 Results – Lecithin trial

Study Reference	(73)
Cognitive Outcomes	
ADAS-cog	-1.3 [-2.9, 0.3]
MMSE	0.9 [0.1, 1.7]
Global Outcomes	
CGIC	35% improved, 24% worsened
ADAS-noncog	-0.6 [-1.4, 0.2]
	[95% CI]

* The above results are difference-from-baseline scores and 95% confidence intervals for the lecithin-only group (at week-four).

The results in the table, along with those in the text above, are difficult to assess for clinical significance because of the short duration of the trial. The MMSE is unsuited for such short-term follow-ups. When compared to expected changes reported from validation studies, the above results are likely unimportant. Also, Dahlke et al. (74) have shown that nurses and physicians score the CGIC differently, and both were interchangeably used in the present study, possibly reducing the reliability of the CGIC results. No mention is made of whether training of the outcome raters at each of the centers occurred, nor whether different word lists were used for consecutive ADAS-cog administrations.

3.7 Propentofylline

Propentofylline (HWA 285) is an inhibitor of both adenosine-uptake and phosphodiesterase enzymes (75), with neuroprotective and anti-oxidant properties. It has been shown to ameliorate cognitive impairments in murine models (76), increase long-term potentiation in guinea pig hippocampus (77), and increase endogenous nerve growth factor (NGF) in mouse astroglial cells (78). It has been reported to prevent the normal decline of NGF immunoreactivity in aged rats (79) and this may promote cholinergic neuron survival.

3.7.1 Methods

3.7.1.1 Design

Two studies were selected following the blinded review, and both were parallel trials. One was a three-month, single-centre study (80) of 30 AD patients diagnosed using the NINCDS/ADRDA criteria. Subjects were excluded for any neurologic and psychiatric histories, vascular (including peripheral) diseases, carcinomas, and use of centrally acting drugs. The primary outcome measure was change to cerebral metabolism (via ^{18}F-2-fluro-2-deoxy-d-glucose positron emission tomography) during a continuous auditory recognition task. The other study (81) accrued 260 mild AD and vascular dementia outpatients (DSM-IIIR criteria) from multiple centres for a twelve-month trial. Eligibility was similar to the Mielke et al. study (80), except benzodiazepines were allowed if prescribed for sleep difficulties.

Subjects in the Mielke et al. study (80) were on average 64.8 years of age (range 52 to 78), 57% male, with mean baseline MMSE scores of 20.4 (SD 3.5). Marcusson et al. (81), did not report patient characteristics separately for those suffering from AD and vascular dementia. The mean age of the study subjects on propentofylline was 71.9 years (SD 7.3) and on placebo was 72.9 years (SD 7.1). The sex distribution was not reported.

3.7.1.2 Treatment

The treatment regimen for both trials was identical with a dose of 900 mg/day of propentofylline (in three capsules).

3.7.1.3 Outcome Measures

The main outcome measures were the Gottfries-Bråne-Steen (GBS) scale and the CGI item II for global functioning, along with the Syndrome Kurz test (SKT) and MMSE for cognition.

3.7.2 Results

3.7.2.1 Drop-outs and Adverse Effects

Two subjects (7%) were withdrawn from the Mielke et al. (80) study, while 73 (28%) were withdrawn from the Marcusson et al. (81) study.

Propentofylline increases cerebral blood flow, and can cause headaches, dizziness and vertigo. Since it is a xanthine derivative, gastro-intestinal (GI) difficulties can be expected. In one study (80), two (13%) propentofylline-treated subjects developed a rash, and one (7%)

propentofylline and three (20%) placebo-treated subjects developed GI pain. In the second study (81), 45 (35%) propentofylline and 23 (18%) placebo patients developed nausea, dizziness, GI pain, headache, or heartburn. These latter events were severe enough to cause 15 (12%) propentofylline and 11 (8%) placebo subjects to withdraw from the trial.

Table 3.7 Description– Propentofylline trials

Study Reference	(80)	(81)
Duration	Three months	Twelve months
Dose	900 mg/day	900 mg/day
Primary Outcome measures	Glucose metabolism	GBS, SKT
Subjects randomized	N=30	N=260
Percent Completed	93%	72%

3.7.2.2 Efficacy

Following three months of treatment, propentofylline-treated subjects showed statistically significant increases in the metabolism of specific cortical and sub-cortical regions compared to placebo-treated subjects (during the auditory word-recognition task) (80). It is difficult to translate such results into a clinically meaningful effect. Several neuropsychological tests were included, but no outcome measures differed appreciably between the treatment groups; however, the authors' reported worsening of dementia severity (as measured by the MMSE) in the placebo group is misleading. The propentofylline subjects actually declined more on average but exhibited greater variability than the subjects in the placebo group (respectively, -1.4 (SD 3.8) versus –1.0 (SD 1.6)). These results do not support a slowing of the progression of AD. On the neuropsychological scores, propentofylline administration improved DSST (three month difference 3.2 (6.2 SD) for propentofylline and –2.1 (7.5 SD) for placebo). This test, however, requires much of elderly subjects and is affected by visual impairment and manual dexterity. Little normative data for demented elderly are available to properly assess such changes.

In subjects diagnosed with AD, the Marcusson et al. (81) study indicated small yet consistent effects in favor of propentofylline administration, although only changes in SKT measures were statistically significant at 12 months. Baseline SKT scores were 16.2 (SD 6.4) for propentofylline and 15.2 (SD 6.5) for placebo, while twelve month differences were –1.3 (SD 4.0)

and –0.2 (SD 3.6), respectively. With the lack of information on what constitutes a clinically meaningful change on the SKT, the interpretation of the above differences is difficult; as such, it is suggested that the SKT is currently not appropriate as a primary outcome in clinical Alzheimer trials. Twelve-month differences from baseline on the GBS scale were 1.8 (SD 14.1) for propentofylline and 5.1 (SD 14.7) for placebo, where increased scores indicate worsening. The MMSE outcomes were not separated into dementia subgroups, so results cannot be reported for AD alone.

Table 3.8 Results – Propentofylline trials

Study Reference	(80)	(81)
Cognitive Outcomes		
SKT	NA	-1.1 [-2.3, 0.1]
DSST	5.3 [-0.1, 10.7]	NA
MMSE	-0.4 [-2.8, 2.0]	NA
Global Outcomes		
GBS	NA	-3.3 [-7.9, 1.3]
CGI Item I	NA	-0.1 [-0.3, 0.2]
CGI Item II	NA	-0.6 [-1.9, 0.8]

Difference scores and 95% confidence intervals [ ] between placebo and propentofylline for reference (80) are at three months, and reference (81) are at twelve months.

Marcusson et al. (81) presented their CGI results as mean changes from baseline, but the ordinal nature of the scoring would be better represented by percentages of improvement or worsening (something that could not be extracted for the above table). Further, while the interviewers were blind to the subjects' treatment, they were likely knowledgeable of scores on other scales and instruments, and so grading of the CGI is unlikely to have been objective.

3.8 Linopirdine

Linopirdine is a phenyllindolinone compound that enhances presynaptic endogenous acetylcholine release (82-84), thereby avoiding side effects commonly attributed to the AChE inhibitors. Other beneficial effects include the release of norepinephrin and serotonin, which are both implicated in AD (82).

3.8.1 Methods

3.8.1.1 Design

The one trial retained following blinded review was a multicentre parallel study, in which 375 AD subjects (diagnosed using DSM-IIIR and NINCDS/ADRDA criteria) were randomized and followed for six months (85). Eligibility criteria excluded those subjects with other psychiatric or neurologic histories, carotid or vertebral-basilar stenosis, alcohol or drug abuse, and concurrently using centrally active drug use. The trial was designed and sponsored by DuPont Pharma. An intent-to-treat analysis was used.

Linopirdine study subjects had a mean age of 71.5 (SD 7.9) and placebo subjects 71.7 years (SD 7.96). Overall, 56% were male and the average MMSE score at baseline was 19.5 (SD 3.1).

3.8.1.2 Treatment

Following a two-week placebo baseline phase, subjects were randomized to either 90mg/day linopirdine (in three capsules) or placebo.

3.8.1.3 Outcome Measures

The primary outcome measures were the CGI for global functioning and the ADAS-cog for cognitive performance

3.8.2 Results

3.8.2.1 Drop-outs and Adverse Effects

Overall, 20 (10.5%) placebo patients and 61 (31.6%) linopirdine patients withdrew from the study. Baseline characteristics were reported for all the screened subjects (N=382), but not for those randomized (N=375).

Treatment-related adverse events occurred in 68 (36%) of placebo- and 77 (41%) of linopirdine-treated subjects, all of who were more susceptible to elevated alanine transferase enzymes (ALT) in the liver than placebo-treated subjects. Elevations of ALT that were three times the upper limit of normal were seen in 37 (19.9%) linopirdine subjects and 2 (1.6%) placebo subjects.

Table 3.9 Description – Linopirdine trial

Study Reference	(85)
Duration	Six months
Dose	90 mg/day
Primary Outcome measures	ADAS-cog, CGI
Subjects randomized	N=375
Percent completed	78%

3.8.2.2 Efficacy

Not all subjects completed the full six months of treatment. For these individuals, the last assessment scores were carried forward and substituted as sixth-month measures. The following effect sizes are thus based on variable periods of treatment up to six months. According to the CGI results, few subjects deteriorated during the study, and the proportion of improved patients on linopirdine was similar to the proportion that improved in the placebo group (both approximately 10%). ADAS-cog results were presented graphically; linopirdine treated subjects showed small improvements (approximately one to two points) over placebo treated subjects over the course of the six months. The six-month ADAS-cog scores were 20.2 for linopirdine and 22.1 for placebo (P=0.01). Although the linopirdine group exhibited a statistically significant lower ADAS-total score than the placebo group (in the carry-forward analysis), the clinical relevance of this difference and previous ones are questionable. Also, no mention is made of whether different word lists were used for consecutive ADAS-cog administrations.

Although the trial reported assessing subjects with the Physical Self-Maintenance Scale, no results were given.

Table 3.10 Results – Linopirdine trial

Study Reference	(85)	
	Baseline	*End-point (six-months)*
Cognitive Outcomes		
ADAS-cog *	NA	-1.9 [-3.3, -0.5]
SKT	38.5 [26.2, 50.8] placebo 37.8 [24.7, 50.9] linopirdine	40.8 [36.1, 45.5] 39.5 [36.6, 42.4]
Global Outcomes		
ADAS-noncog	2.4 [0, 7.5] placebo 2.3 [0, 7.4] linopirdine	3.4 [1.6, 5.2] 3.0 [1.4, 4.6]

* Difference score between placebo and linopirdine and 95% confidence intervals [ ].

3.9 Ginkgo Biloba (EGb 761)

EGb 761 is a standardized extract from the leaves of the ginkgo biloba tree. There are a number of different chemicals in EGb but it is the flavenoids and terpenoids that are believed to have therapeutic potential. The mechanism of action is not well understood yet ginkgo biloba has recently been approved in Germany for the treatment of cerebral insufficiency. There are several studies, for the most part published in German, reported in a review by Kleijnen and Knipschild (86) that indicate that ginkgo biloba has an impact on cerebral circulation, neuronal cell metabolism, and elimination of free radicals. In the same review it is stated that the typical symptoms of cerebral insufficiency that are believed to be relieved by ginkgo treatment are: difficulties in concentration and memory, absent mindedness, confusion, lack of energy, tiredness, decreased physical performance, depressed mood, anxiety, dizziness, tinnitus and headache. EGb 761 is also known as Tebonin, Tanakan and Rökan. These are standardized on the amount of flavonoids (25%) and terpenoids (6%). A fourth preparation, Kaveri, is similarly standardized. One concern with these preparations of ginkgo biloba is that the preparations are different depending upon the manufacturing process used. One potential impact of these differences is variation in effect.

3.9.1 Methods

3.9.1.1 Design

In the initial evaluation, 7 trials of ginkgo biloba and its effects on cognition were identified (12-14;87-90) and met the minimum methodological standards. Four of these were subsequently excluded (12-14;90) either because the subjects were not demented (90) or it was not possible to determine with any degree of certainty whether the subjects would satisfy the NINCDS/ADRDA criteria for the diagnosis of probable Alzheimer's disease (12) (13;14;90). The remaining 3 trials (87-89) underwent in-depth review. Two of the trials were conducted in Germany (88;89) and the third in the US (87). The first (89) was a multicentre trial in which study subjects were recruited from 41 sites in Berlin in 1990-1991. The authors state that they had to provide on site training for the diagnosis of dementia as the site investigators were not adequately familiar with the diagnostic process. The source of funding for the study was not mentioned but one of the authors was an employee of Dr. Willmar Schwabe GmbH and Co, Karlsuhe, Germany, the manufacturer of EGb 761. The second German trial was a small single centre study with 20 patients with AD (88). The funding source was not stated. The US study conducted by the North American EGb Study Group (87) was the largest with 309 study subjects with AD (n=236) or multi-infarct dementia (n=73). The Dr. Willmar Schwabe Pharmaceutical Company funded the trial. In the two trials that included both AD and multi-infart dememtia (MI) study subjects (87;89) only partial details were reported for the AD group separately. For this reason it is difficult to make conclusions concerning the use of EGb as a treatment specifically for AD.

Although the validity of the diagnosis of AD and MID in the Kanowski et al. (89) trial raises some concerns, the authors do state that after the trial (but before data analysis) the differentiation of subjects into one of the two forms of dementia was assessed by the principal investigators. Le Bars et al. (87) used the DSM-IIIR and ICD-10 to establish a diagnosis of "uncomplicated dementia" either AD or MID, while Maurer et al. (88) used the DSM-IIIR and the NINCDS/ADRDA criteria for a diagnosis of probable AD. Kanowski restricted inclusion of subjects to those 55 years of age or older, with a MMSE score of 13-25 and with a Syndrom-Kurztest (SKT) score of 6-18 indicating very mild (SKT=5 to 8) to moderate (SKT=14 to18) impairment. Potential patients with disturbance of consciousness and patients suffering from other cerebral diseases or dementia of other origin were excluded. In addition, those taking medications that may have interfered with the assessment of efficacy were excluded. Maurer et al. (88) included

those with mild to moderate dementia as determined by a score of 3-5 on the BCRS and excluded those with a Hachinski score of >4 (e.g., those with vascular causes). Other reasons for exclusion were dementia of other etiology and any condition that would interfere with psychometric testing. Subjects between the ages of 50 and 80 were eligible for inclusion. Le Bars et al. (87) included those with uncomplicated dementia with an MMSE score between 9 and 26 and a GDS of 3 to 6 (mild to moderately severe dementia). No other significant medical conditions were permitted but medication use was allowed with the restriction that neither a change in regimen nor new prescriptions were permitted.

Kanowski et al. (89), reported the baseline characteristics of the combined group of subjects who completed the trial. Baseline information was not presented for those who were randomized. This is a serious limitation to the interpretation of this trial and precludes any comparison between patient characteristics across EGb trials. Furthermore, there were no data presented for the AD group separately. Those who completed the treatment were 70 years of age on average and age did not differ across treatment groups. The majority of study participants were women (66% in the treated group vs. 69% in the placebo group). The males assigned to the placebo group who completed the trial were 66 years of age on average as compared to 72 years of age for the females Based on measurement on the SKT, the majority of study subjects who completed the trial in both groups were mildly impaired (85% of those who completed the trial in the treated group scored $\leq$13 points as did 73% of the placebo group) at baseline. Baseline characteristics from the US trial (87) indicated that the majority of patients were female (58% overall), with 62% female in the placebo group and 54% in the treated group. The study subjects were 68 years of age on average (overall and in each intervention group) and average MMSE scores at baseline were 21.2 overall, 21.3 in the placebo group and 21.1 in the treated group indicating mild impairment.

In the Maurer et al. trial (88), the subjects in the placebo group were younger on average than those in the treatment group (60.6 years vs. 68.5 years). Forty four percent of the placebo group (4 of 9) was female as compared to 55% (5 of 9) in the treatment group. Although the data are not reported in the article, the authors state that the treatment group was “slightly more impaired” according to the BCRS than those in the placebo group. These baseline differences were not addressed in the analysis. The SKT scores in this trial (18.1 in the placebo group and 19.7 in the treated group, indicating moderate dementia) are quite different from those reported for trial

completers in the Kanowski et al. (89) trial (e.g., 12.2 for the AD subjects in the placebo group vs. 10.9 in the AD subjects in the treated group).

Although the sex distribution and age characteristics are similar in these trials, it is difficult to assess the comparability of level of impairment given the three different measures (BCRS, SKT and MMSE) that were used as well as the fact that one trial presents baseline characteristics only for those who completed the trial (89). Overall, however, it would appear as if the subjects recruited in the Maurer et al. (88) trial were more impaired than those in the other two trials.

3.9.1.2 Treatment

The duration of the trials were 12 weeks (88), 24 weeks (89) and 52 weeks (87). Each trial started with a single blind placebo run-in period of 7 days (88), 4 weeks (89), and 14 days (87). Each trial included a placebo group and a single treatment group. Two of the trials used dosages of 240 mg/day (88;89). In Kanowski et al. the regimen was 2 x 120 mg tablets, one taken in the morning and one in the evening before meals, whereas in Maurer et al. study subjects were given 2 x 40 mg tablets three times a day. In the US trial, the dosage was decreased by half and study subjects received 1 x 40 mg tablet three times a day before meals.

3.9.1.3 Outcome Measures

The primary outcome measures used to assess treatment efficacy in the Kanowski et al. (89) trial were the Clinical Global Impressions (CGI) scale, the Syndrom-Kurz Test (SKT) to assess cognition and the Nuremberg Geriatric Observation Scale (NAB) to assess activities of daily living. The primary outcome measures for the Le Bars et al. (87) trial were the ADAS-cog, the GERRI for daily living and social activities, and the CGIC for general psychopathology. Le Bars et al. state that the results related to secondary outcomes were presented elsewhere. Maurer et al. (88) chose only the SKT as a primary outcome measure but the Trailmaking Test (ZVT), the ADAS and the Multiple Choice Vocabulary Test (MWT-B) were selected as secondary outcome measures.

3.9.2 Results

3.9.2.1 Dropouts and Adverse Effects

No adverse events were recorded in the Maurer et al. (88) study and the only mention of dropouts is that 2 patients were excluded "one woman had been excluded because the P300 could not be recorded, and one other woman was excluded because of lack of compliance." The study results are based on the remaining 18 subjects. Kanowski et al. (89) report that, 205 subjects of the

original 216 who were randomized were available for the intention to treat analysis and 156 were available for the efficacy analysis. The majority of participants available for the efficacy analysis were suffering from AD (n=125). Adverse effects were reported only for the total group and 122 events were documented of which 7 were considered to be serious. The occurrence of any adverse event was equally divided between the study groups but 5 of the 7 serious adverse events occurred in the treated group. All but one of these 7 discontinued treatment. The authors state that with one exception the adverse events were not considered to be associated with the drug. Full details of the types of adverse events that occurred were not presented. However, the authors stated that skin diseases were more frequent in the treated group while gastro intestinal disorders were more frequent in the placebo group.

In the US trial (87), information on dropouts was presented for the AD and MID subjects combined. Beginning with 327 randomized subjects, only 78 (48%) and 59 (35%) completed the study in the treatment and placebo groups respectively. The majority of dropouts took place after 26 weeks of study (half way through the trial) in both groups. There were five serious adverse events (i.e., 2 deaths in the treated group; 1 death, 1 stroke and 1 subdural hematoma in the placebo group). One third of subjects in both groups reported at least one adverse event of any type. Of those classified as probably being related to treatment, there were 16% in the treated group and 12% in the placebo group. The majority of adverse events were considered to be mild with slightly more gastro intestinal symptoms reported by those in the treated group. The low percentage of completion in both the placebo and treatment groups in this trial severely limits the interpretation of the results.

3.9.2.2 Efficacy

To assess treatment efficacy Kanowski et al. (89) defined clinically relevant changes in each of the primary outcomes as follows:

- Changes in CGI to “much improved” or “very much improved”
- Decrease in the SKT total score of at least 4 points
- Decrease in the NAB total score of at least 2 points.

Using these definitions (and considering the total subject group), these authors report that 10% of the placebo group satisfied two or more of these criteria compared to 28% of the treatment group ($p=0.012$). Considering the individual outcomes, 32% of the treated group as compared to 17% of the placebo group experienced clinically relevant changes in CGI ($p<0.01$); 38% of the treated

group compared to 18% of the placebo group experienced clinically relevant changes in the SKT ($p<0.05$). There was no statistically significant difference between the two groups using the NAB results in terms of percentage of responders. Quantitative changes in these parameters were presented descriptively for the AD group separately. Treatment:placebo differences from baseline to 24 weeks for the CGI, SKT, and NAB were 0.7, 1.5 and 0.6 respectively. Due to the small numbers of subjects recruited from each study site, possible centre effects were not addressed. Unfortunately for two of these parameters, the CGI and the NAB, there is little or no English language literature examining responsiveness to change.

Despite the fact that the study subjects in the Maurer et al. (88) appeared to be more impaired than those in the Kanowski et al. (89) trial according to SKT scores, the treatment:placebo differences from baseline to endpoint were larger. The treated group in the smaller of the two trials improved an average of 2.88 SKT points in 12 weeks as compared to 2.4 SKT points on average over 12 weeks in the Kanowski et al. (89). In the latter trial the placebo group improved an average of 0.9 points but in the former trial there was deterioration in the placebo group of 0.8 points. There were no differences in ZVT in the Maurer et al. (88) trial. Using the same definition of CGI responder as in the Kanowski et al. (89) trial, a similar proportion of placebo subjects, 11% responded, compared to 33% in the treated group.

In the US trial (87), the placebo group exhibited a statistically significant decline on the ADAS-cog with an average increase in score of 1.5. The treated group, on the other hand, showed no significant change (average change score=-0.2). The difference between the two groups (1.7) was found to be statistically significant ($p<0.05$) and in favor of treatment over the 52 weeks of the trial. The study of Maurer et al. (88) also used the ADAS-cog as a secondary outcome and the results were similar in magnitude although the study was only of 12 weeks duration. The treated group improved with the ADAS-cog score decreasing by an average of 1.89 points, while the placebo deteriorated slightly with an average increase in score of 0.03 points. Probably due to the small sample size in this trial (n=18) the observed difference did not reach statistical significance.

Table 3.11a: Summary of Donepezil RCTs

Study	Rogers et al., 1996 (16) Jadad score=6-7 Funded by Eisai	Rogers et al., 1998 (17) Jadad score=6-7 Funded by Eisai	Rogers et al., 1998 (18) Jadad score=6 Funded by Eisai	Burns et al., 1999 (19) Jadad score=6.5 Funded by Eisai
Study Information	Multicentre, parallel; 12-wk, plus 2-wk washout; 40 patients planned per group; Completion: 20 patients did not complete the study; Baseline: mean age≅72.	Multicentre, parallel; 24-wk, plus 6-wk washout; placebo (n=162), 5mg donepezil (n=154), 10mg donepezil (n=157); Completion: not discussed; Baseline: mean age≅73, mean MMSE=19, CDR=0.5-2.	Multicentre, parallel; 12-wk, plus forced titration for 10mg dose, then 3-wk single-blind washout; 468 patients; Completion: placebo 93%, 5mg 90%, 10mg 82%; Drop-outs: 12%; Baseline: probable AD, CDR=0.5-2, mean age=74.	Multicentre, parallel; 24-wk double-blind, 6-wk single-blind washout; Placebo (n=274), 5mg (n=271), 10mg (n=273); Completion: placebo 80%, 5mg 78%, 10mg 74%; Baseline: mean age≅71 (50-93), mean MMSE=20 (9-26), CDR=0.5-2.
Treatment	Dose ranging study; Placebo, 1mg donepezil, 3mg donepezil, 5mg donepezil.	Placebo, 5mg donepezil, 10mg donepezil (10mg group had forced titration); Pills taken once daily in the evening.	Placebo, 5mg donepezil, 10mg donepezil (5mg taken for up to 1 week before receiving 10mg); Pills taken once daily at bedtime.	Placebo, 5mg donepezil, 10mg donepezil; Pills given orally once daily in the evening; 10mg group received 5mg for the first wk of the study.

Table 3.11b: Summary of Metrifonate RCTs

Study	Becker et al., 1996 (21) Metrifonate Jadad score=6-8 Funded by NIMH, NIA	Becker et al., 1998 (26) Jadad score=5 Funded by NIMH, NIA	Cummings et al., 1998 (23) Jadad score=8 Funded by Bayer	Morris et al., 1998 (25) Jadad score=7-8 Funded by Bayer
Study Information	Parallel; 3-wk placebo baseline, 3-month double-blind, 1 month off, open phase for maint. dose; Placebo (n=23), metrifonate (n=27); Drop-outs: 3 (not included in above totals); Baseline: mean age=71.4, mean education (yrs.) =12.6 (pl)/11.2 (m), mean MMSE=19.5 (pl)/19.3 (m), mean ADAS-cog=25.6 (pl), 26.4 (m).	Parallel; 4-wk patient-blind phase where all received metrifonate (1-wk dose finding, 3-wk safety), then 6-month double-blind (no washout between periods); Completion: 46 of 47 completed the trial; Baseline: mean age=71.7 (pl)/74.4 (m), mean education (yrs.) =13 (pl)/13.2 (m), mean MMSE=19.8 (pl)/20.6 (m), mean ADAS-cog=21.8 (pl)/20.6 (m).	Multicentre, parallel; 2-wk screening, 12-wk double-blind (2-wk loading,10-wk maintenance), then 40-wk open label extension; Completion: 463 of 480 enrolled; Drop-outs: 18; Compliance: low dose 34.5%, mid dose 52.2%, high dose 72.5%; Baseline: mean age≅74.	Multicentre, parallel; 2-wk screening, 26-wk double-blind (incl. 2-wk loading), then 8-wk follow-up; Placebo (n=135), metrifonate (n=273); Drop-outs: placebo 16, metrifonate 58; Noncompliance: placebo 3, metrifonate 3; Baseline: mean age≅73.6, mean MMSE=19.4 (pl)/18.8 (m), ADAS-cog=20.4 (pl)/22.9 (m).
Treatment	Placebo; Metrifonate dose was 5.0mg/kg wkly for 2 wks, then 4.9mg/kg wkly for 1 wk to reach 40-60% inhibition after 4 wks; Maint. dose = 2.1mg/kg wkly for 9 wks.	Placebo, 2mg/kg metrifonate for 5 days and .95 mg/kg on day 6, then 2.9mg/kg weekly for maintenance (6 months); Inhibition goal: 50-70% after 1 wk.	Pl, m → .5-.2mg/kg (ld), .9-.3mg/kg (md), 2-.65mg/kg (hd); Inhibition goals: 30% (ld), 50% (md), 70% (hd); Oral administration once daily at breakfast.	Placebo, 2mg/kg metrifonate loading dose, then .65mg/kg for 24 wks taken once daily before breakfast; Inhibition goal: 70%.

Notes: pl=placebo, m=metrifonate, ld=low dose, md=mid dose, hd=high dose

Table 3.11c: Summary of Metrifonate RCTs (con't) and Rivastigmine RCTs

Study	Raskind et al., 1999 (22) Jadad=7 Sponsored by Bayer	Jann et al., 1999 (20) Jadad=7 Sponsored by Bayer	Rösler et al., 1999 (31) Rivastigmine Jadad score=8 Funded by Novartis	Corey-Bloom et al., 1998 (29) Rivastigmine Jadad score=5-6 Funded by Novartis
Study Information	Multicentre, parallel, double-blind; 2-wk screening period, 26-wk double-blind treatment period, 8-wk post-treatment follow-up period; Placebo (n=87)- 84% completion, metrifonate (n=177)-82% completion; Drop-outs (adverse effects): pl=9%, m=11%; Baseline: probable AD (NINCDS/ADRDA), mean age=74.5 (pl)/74.6 (m), % male=32.2 (pl)/37.9 (m), mean MMSE=18.7 (pl)/18.7 (m)	Multicentre, parallel, double-blind; 2-wk screening period, 6-wk treatment period (ldg=2-wk loading dose, 4-wk maintenance/nldg=6-wk maintenance); # rand. and % compl. → pl (134/97%), ldg (133/89%), nldg (128/93%); Drop-outs (adverse effects): pl (2%), ldg (8%), nldg (5%); Baseline-means (pl/ldg/nldg): age=75.0/74.8/75.2, % male=45.9/40.9/40.6, MMSE=18.5/19.5/19.2, ADAS-c=22.4/20.4/21.5	Multicentre, parallel; 26-wks (12-wk fixed dose escalation, then 14-wk flexible dose maintenance); Placebo (n=239), low dose (n=243), high dose (n=243); Completion: 80% of randomized patients (581/725); Baseline: mean age=72, mean AD duration=39 months, severity (mild 41%, moderate 57%, severe 2%), mean MMSE=19.9, mean ADAS-cog≅23.5, mean PDS≅54.1.	Multicentre, parallel; 7-wk fixed dose titration, then 8-26-wk flexible dose maintenance; Placebo (n=235), low dose (n=233), high dose (n=231); Completion: placebo 84%, low dose 85%, high dose 65%; Baseline: mean age=74, 42% mild dementia, 57% moderate dementia.
Treatment	50mg taken once daily as maint. dose for 26 wks, identical placebo pill for control group; Inhibition goal: not stated.	Daily pill for pl or m; ldg → 2 wks of 6-wk double-blind at 100mg (<65 kg) or 150mg (>65kg), then 50mg for 4 wks; nldg → 50mg for 6 wks; Inhibition goal: 70%.	Placebo, low dose=1-4mg/day, high dose=6-12 mg/day; Doses increased weekly within assigned dose range in steps of up to 1.5mg/day during wks 1-12, then in weeks 13-26 doses could be increased or decreased within the assigned range with the aim to reach the highest well tolerated dose.	Placebo, low dose=1-4mg; high dose=6-12mg (2 capsules taken 2X/day w/food); Dosage titrated weekly within assigned dose range during weeks 1-7; then in weeks 8-26 doses could be increased or decreased within the assigned range with the aim to reach highest well tolerated dose.

Notes: pl=placebo, m=metrifonate, ldg=loading dose group, nldg=no-loading dose group

Table 3.11d: Summary of Selegiline RCTs

Study	Tariot et al., 1998 (45) Jadad score=5-6 Funded by NIMH, NIA	Freedman et al., 1998 (42) Jadad score=7-8 Funded by Ontario's Min. of Health and Mental Health Foundation	Sano et al., 1997 (40) (Selegiline and Vitamin E) Jadad score=5 Funded by NIH/Meds supplied by Hoffman Laroche Somerset	Lawler et al., 1997 (46) Jadad score=5 Funded by NIH and Somerset Pharm.
Study Information	Crossover; 8 wks selegiline, 8 wks placebo, 4-wk washout; 50 subjects w/2 drop-outs (depression, perceived lack of efficacy); Baseline: mild to moderate AD, outpatients, mean age=69.9, mean education=13.0 vs. 14.4.	Parallel; 4-wk single-blind placebo, 25-wk double-blind, then 12-wk single-blind placebo washout; 51 subjects; Compliance: >95%; Drop-outs: 1 placebo, 7 selegiline; Baseline: probable AD, mean age=70, GDS=3.9-4.2, MMSE= 17.8-18.5.	Multicentre, parallel; 2-year follow-up; 341 patients; Compliance: selegiline '+' → placebo 13%, selegiline 98%, vitamin E 11%, combined 93% - vitamin E '+' → placebo 12%, selegiline 9%, vitamin E 93%, combined 91%; Baseline: mean age≅75, mean education (yrs.) ≅12.5, mean MMSE=13.3 (pl)/11.3 (VE), mean BDS=6.1 (pl)/6.6 (VE).	Crossover; 6-wk selegiline, 2-wk washout, 6-wk placebo; 25 enrolled, but 5 dropped-out during first 2 wks because of inability to comply w/study protocol; No drop-outs b/c of adverse effects; Baseline: Prob AD, mean age=75, mean duration of AD=4.8, mean MMSE=11.9.
Treatment	Placebo, selegiline.	Placebo, selegiline 5mg for 1 wk, selegiline 10mg for 24 wks.	Placebo, 5mg selegiline 2x/day, 1000IU vitamin E 2x/day, combination selegiline and vitamin E.	Placebo, 10mg selegiline.

Note: VE=vitamin E

Table 3.11e: Summary of Selegiline RCTs, con't

Study	Schneider et al., 1993 (44) Jadad score=5 Funded by NIMH, State of California	Burke et al., 1993 (41) Jadad score=5-6 Funded by Somerset Pharm.	Mangoni et al., 1991 (43) Jadad score=7 Funding not specified
Study Information	Crossover; 4-wk selegiline, 4-wk placebo, no washout; 10 subjects w/no drop-outs; Baseline: outpatients, probable AD, mean age=88.8, mean MMSE=15.7, mean education=13.7.	Parallel; 15-month follow-up; 39 patients; Drop-outs: 1 placebo at the beginning, 2 more placebo and 3 selegiline during follow-up → not due to adverse effects; Compliance: 93%; Baseline: mild AD, CDR=1, mean age=74.8 vs. 71.4, education=13.3 vs. 11.4.	Parallel; 3-month follow-up; Placebo (n=51), selegiline (n=68); Drop-outs: placebo → 4 (poor efficacy) and 1 (adverse event), selegiline → 2 (poor efficacy), 3 (adverse events) and 1 (?); Baseline: probable AD, GDS=3, 4 or 5; mean age=68.7 vs. 69.0, mean education (yrs.) =6.9 vs. 6.4.
Treatment	Placebo, 5mg selegiline b.i.d.	Placebo, selegiline 5mg b.i.d.	Placebo, selegiline 10mg per day.

Table 3.11f: Summary of Propentofylline, Linopirdine and Lecithin RCTs

Study	Mielke et al., 1998 (80) Propentofylline Jadad score=5 Funding not specified	Marcusson et al., 1997 (81) Propentofylline Jadad score=5 Funding not specified	Rockwood et al., 1997 (85) Linopirdine Jadad score=6 Funded by DuPont Pharma.	Foster et al., 1996 (73) Lecithin Jadad score=5 Funded by Parke-Davis
Study Information	Parallel; 12 weeks; 15 randomized to placebo, 13 to propentofylline; Drop-outs: 1 refused randomization, 1 developed a rash; Baseline: 16 males (8 in each group), mean age=65.7 (pl) and 63.8 (prop), mean MMSE=21.3 (pl) and 19.5 (prop).	Multicentre, parallel; 1-month screening phase, 12-month double-blind phase; Placebo (n=131), propentofylline (n=129); Drop-outs: 73 (28%), incl. 11 (pl)/15 (prop) b/c of adverse events and 5 (pl)/11 (prop) at caregiver request; Baseline (pl/prop): mean age=72.9/71.9, mean MMSE=20.7/20.8.	Multicentre, parallel; 2-wk single-blind placebo lead-in, 6-month double-blind phase; Placebo (n=193), linopirdine (n=184) at start of lead-in period; Drop-outs: 20 placebo and 61 linopirdine between lead-in and study termination; Baseline (pl/lino): mean age=71.5/71.7, mean education (yrs.) =11.2/11.1, mean MMSE=19.6/19.4.	Multicentre, crossover; 6-wk double-blind dose titration, 2-wk placebo baseline, two 4-wk double-blind treatment phases; 440 entered the study, 116 completed the first double-blind period and 112 completed the second double-blind period; Baseline: high responders to tacrine.
Treatment	Placebo, 300mg propentofylline 3X/day	Placebo, 300mg propentofylline 3X/day.	Placebo, 30mg 3X/day linopirdine.	Tacrine 40 or 80mg, lecithin 9.04 g/day.

Notes: pl=placebo, prop=propentofylline, lino=linopirdine

Table 3.11g: Summary of Ginkgo Biloba RCTs

Study	Le Bars et al., 1997 (87) Jadad score=7 Funded by Schwabe	Maurer et al., 1997 (88) Jadad score=7 Funding not specified	Kanowski et al., 1996 (89) Jadad score=7 Funded by Schwabe
Study Information	Multicentre, parallel; 52-wk double-blind fixed dose phase, including 14-day single-blind run-in on placebo; 251 AD patients enrolled, 236 in ITT (pl=116; GB=120); Drop-outs: huge drop-out at 12 weeks → 38% of pl group completed trial vs. 50% of the GB group; Baseline (ITT subjects): females=58%, mean age=68 (SD=10), median education (yrs.) 14 (range 0-20), mean MMSE=21.2 (SD=5.7), mean ADAS-cog=20.0 (SD=15.8), mean disease duration=4.3 (SD=3.8).	Parallel; 3-month placebo-controlled double-blind trial, preceded by 7-day run-in phase; 20 patients (initial target was 40); Drop-outs: 2 patients were excluded b/c of lack of compliance and no P300 recorded); Baseline: mean age=68.5 (pl) and 60.6 (GB), ADAS-cog=36.1 (pl) and 31.2 (GB).	Multicentre, parallel; 4-wk single-blind run-in with placebo, then 24-wk double-blind phase; 222 enrolled, 156 completed (pl=77, GB=79); Drop-outs: 6 after run-in, 11 after 12 wks, 21 over course of trial, 28 protocol violations (failure to complete scheduled visits, poor compliance, use of prohibited medications); Baseline (available only for those who completed the study): Placebo group → n=24 males and 53 females, mean age=66.4 (males) and 72.0 (females), GB group → n=27 males and 52 females, mean age=69.7 (males) and 70.2 (females).
Treatment	Placebo; 40mg GB tablets 3X/day before meals, each tablet contains 24% flavoneglycosides, 6% terpenelactones (31% ginkgosides A, B and C, plus 2.9% bilobalide).	Placebo; 240mg GB (80mg 3X/day), each 40mg contains 9.6mg flavoneglycosides and 2.4mg terpenelactones.	Placebo; 120mg GB 2X/day before meals, each capsule contains 24% ginkgoflavoneglycosides and 6% terpenelactones (ginkgosides, bilobalide).

Notes: pl=placebo, GB=ginkgo biloba

Table 3.12: Summary of Most Frequently Used Outcome Measurement Scales

Scale	Subscale	Construct Measured	Rater	Scoring
Clinical Global Impression (CGI) *Note: The subscales can be treated as separate instruments or considered parts of a single instrument.	Clinical Global Improvement (CGI), Global Improvement Index, Clinician Global Impression of Change (CGIC or CGI-change)	Overall improvement in patient health status	Clinician	1 (very much improved)-7 (very much worse)
	Clinical Global Impression of Severity (CGI-severity)	Severity of illness	Clinician	1 (normal)-7 (among the most extremely ill)
	Efficacy index	Comparison of benefits versus side-effects of a medication	Clinician	1-16, calculated using a grid that contains descriptions of four possible drug effects and four possible levels of severity for drug complications
Clinician Global Impression of Change-plus (CIBIC-plus)	N/A	Overall improvement in patient health status	Clinician, with caregiver input	1 (very much improved)-7 (very much worse)
Clinical Dementia Rating (CDR)	N/A	Cognitive impairment in memory, orientation, judgment/problem-solving, community affairs, home/hobbies, and personal care	Clinician	0=none, 0.5=questionable, 1=mild, 2=moderate, 3=severe A score is awarded for each of the six categories; an algorithm is then used to obtain a single score (0-3) for overall impairment.

Table 3.12: Summary of Most Frequently Used Outcome Measurement Scales, con't

Clinical Dementia Rating-Sum of Boxes (CDR-Sum of Boxes)	N/A	Cognitive impairment in memory, orientation, judgment/problem-solving, community affairs, home/hobbies, and personal care	Clinician	0=none, 0.5=questionable, 1=mild, 2=moderate, 3=severe A score is awarded for each of the six categories; these scores are then added together to obtain a final score between 0-18.
Global Deterioration Scale (GDS)	N/A	Progressive stages of cognitive impairment	Clinician	1 (no cognitive decline)-7 (very severe cognitive decline)
Alzheimer's Disease Assessment Scale-cognitive (ADAS-cog)	N/A	Orientation, memory, language, and praxis	Clinician or other trained examiner (clinician must rate some parts of the scale – e.g., spoken language ability)	0-70, with higher scores indicating greater impairment
Mini-Mental State Examination (MMSE)	N/A	Orientation, memory, language, concentration, and praxis	Trained interviewer	0-30, with higher scores indicating less impairment ($\leq$23 = impaired; $\leq$17 = severely impaired)
Syndrome Kurtz Test (SKT)	N/A	Attention: naming objects/numerals, arranging/replacing blocks, counting symbols, reversal naming; Memory: immediate/delayed recall, recognition memory.	Trained interviewer	Original scores for German-speaking countries: range 0-27, with higher scores indicating more severe impairment (5 levels of severity)

3.10 Adverse Effects

Monitoring of adverse effects (AE) is mandatory in clinical trials. Globally, there are two main types of AE that can be observed: (i) effects that are expected on the basis of the pharmacological properties of the drug (type A adverse effects); and (ii) effects that are not expected (type B adverse effects). Type B adverse effects are thought to be dependent on patient characteristics because they do not occur in the majority of the treated population.

Type A adverse effects can be detected in clinical trials because they are relatively frequent and often occur soon after the initiation of treatment. On the other hand, because of limited sample size and insufficient trial duration, clinical trials are not appropriate to detect effects that are rare (i.e. type B adverse effects), nor those that occur some time after treatment initiation. Unfortunately, rare effects are often the most severe. Furthermore, even for the most common AE, incidence estimates are generally not precise due to small sample sizes. Thus, the safety profile of a drug cannot be fully determined using clinical trial data.

Nevertheless, all AE occurring during the course of a clinical trial are recorded, even if imputability cannot be assumed. Consequently, AE can be observed in both the treated patients and in those who are not receiving active treatment. Various classifications can be used to assess the severity of AE (91). One classification, proposed by Ibanez et al. (92), is based on survival: (i) fatal; (ii) severe - i.e. directly life threatening; and (iii) moderate - i.e. leading to hospital admission or to absence from work or school. Another classification (93) is based on treatment required for the AE: Grade 1 - Adverse Drug Reactions (ADR) did not lead to any change in medication; Grade 2 - ADR led to a change in medication in the form of dose reduction and/or additional treatment; and Grade 3 - ADR led to discontinuation of the medication suspected of causing the AE. Unfortunately, the AD trial reports reviewed here do not provide adequate data to classify adverse events using either classification, although it is usually noted whether or not the occurrence of the adverse event led to treatment interruption.

The grid below lists AE, their frequency of occurrence, and severity as reported in each of the reviewed trials. The severity column classifies AE incidence rates into three

categories: mild, moderate, or severe. Classification is based on the respective trial authors' stated impressions of severity, and incidence rates are based on the total number of subjects in a particular treatment arm. The severity column also shows the number (or percentage) of drop-outs due to AE.

Table 3.13a: Adverse events

Drug	Trial	Design	Group 1	Group 2	Group 3	Placebo	Severity
Lecithin	Foster et al. (73)	Crossover	Lecithin Not reported				
Propentofylline	Marcusson et al. (81)	Parallel	Propentofylline (N=129) Nausea 7% Dizziness 6% GI pain 5% Heartburn 5% Headache 5%			Placebo (N=131) Nausea 2% Dizziness 3% GI pain 1% Heartburn 1% Headache 1%	Propentofylline: 4 deaths, 4 serious adverse events, and 11 non-serious adverse events. *Placebo*: 2 deaths, 3 serious events, 8 non-serious adverse events. Above Deaths were not ascribed to study medications.
Propentofylline	Mielke et al. (80)	Parallel	Propentofylline (N=15) GI pain 7% Rash 13%			Placebo (N=15) Gi pain 20%	*Propentofylline*: 1 onset of urticarial rash.
Linopirdine	Rockwood et al. (85)	Parallel	Linopirdine (N=183) Any 41% (not specified)			Placebo (N=187) Any 36%	*Linopirdine*: 24 liver function test, 17 other. *Placebo:* 4 other.

Table 3.13b: Adverse events

Drug	**Trial**	Design	**Group 1**	**Group 2**	**Group 3**	**Placebo**	**Severity**
Donepezil	Rogers et al. (16)	Parallel (Phase II)	Donepezil (1 mg/day) (N=42) Nausea 7% Constipation 2% Dizzness 5% Nasal congestion 2% Common cold 10% Headache 10% Flushing 10% Agitation 7% UTI 2% Cough 2% Accident 2% Pain 7%	Donepezil (3 mg/day) (N=40) Diarrhoea 7% Gastric upset 5% Constipation 5% Dizziness 3% Nasal congestion 13% Common cold 5% Headache 5% Flushing 3% Agitation 5% UTI 8% Cough 10% Accident 3% Pain 3%	Donepezil (5 mg/day) (N=39) Nausea 10% Diarrhoea 10% Gastric upset 8% Constipation 8% Nasal congestion 5% Headache 3% Flushing 3% Agitation 3% UTI 3% Cough 3% Accident 10% Pain 5%	Placebo (N=40) Nausea 5% Diarrhoea 3% Gastric upset 5% Constipation 3% Dizziness 10% Nasal congestion 8% Common cold 8% Headache 8% Flushing 3% Agitation 5% UTI 5% Cough 5% Accident 3% Pain 3%	All adverse considered mild to moderate. No category specific incidence rates given for either the mild or moderate categories. Dropouts due to adverse events: Placebo: 2/5 dropouts 1mg/day: 5/8 dropouts 3 mg/day: 2/2 dropouts 5 mg/day: 3/5 dropouts Types of adverse events not stated.
Donepezil	Rogers et al. (17)	*Parallel*	Donepezil (5 mg/day) (N=154) Fatigue 5% Diarrhoea 9% Nausea 4% Vomiting 3% Anorexia 2% Muscle cramps 6% Dizziness 10% Rhinitis 1%	Donepezil (10 mg/day) (N=157) Fatigue 8% Diarrhoea 17% Nausea 17% Vomiting 10% Anorexia 7% Muscle cramps 8% Dizziness 8% Rhinitis 6%		Placebo (N=162) Fatigue 2% Diarrhoea 7% Nausea 4% Vomiting 2% Anorexia 2% Muscle cramps 1% Dizziness 4% Rhinitis 2%	Group 1 / Group 2 / Placebo Mild 33% / 54% / 18% Mod. 7% / 27% / 6% Severe 5% / 10% / 6% n=154 / n=157 / n=162 Dropouts due to adverse events: Placebo: 7% of patients 5 mg/day: 6% of patients 10 mg/day: 16% of patients Types of adverse events not stated

Table 3.13c: Adverse events

Drug	Trial	Design	Group 1	Group 2	Group 3	Placebo	Severity
Donepezil	Rogers et al. (18)	*Parallel*	Donepezil (5 mg/day) (N=157) Nausea 7% Insomnia 8% Diarrhoea 6% Pain 9% Headache 13% Dizziness 9% Muscle cramps 6% Fatigue 3% Accident 6% Agitation 4% Vomiting 3% Anorexia 4% Weight lose 2% Cold 5% Abdominal disturbance 6% UTI 6% Stomach upset 5% Rhinitis 5% URTI 5% Edema 1% Cough 1%	Donepezil (10 mg/day) (N=158) Nausea 22% Insomnia 18% Diarrhoea 13% Pain 13% Headache 12% Dizziness 9% Muscle cramps 8% Fatigue 8% Accident 6% Agitation 6% Vomiting 3% Anorexia 6% Weight lose 5% Cold 4% Abdominal disturbance 4% UTI 4% Stomach upset 3% Rhinitis 3% URTI 3% Edema 3% Cough 2%		Placebo (N=153) Nausea 8% Insomnia 5% Diarrhoea 3% Pain 7% Headache 8% Dizziness 7% Muscle cramps 4% Fatigue 5% Accident 7% Agitation 7% Vomiting 5% Anorexia 3% Weight lose 2% Cold 7% Abdominal disturbance 4% UTI 13% Stomach upset 1% Rhinitis 4% URTI 4% Edema 5% Cough 5%	Group 1 Group 2 Placebo Mild 21% 53% 16% Mod. none spec. none spec. none spec. Severe 4% 4% 5% n=157 n=158 n=153 Dropouts due to adverse events: Placebo: 2% 5 mg/day: 4% 10 mg/day: 10% Most common events leading to discontinuation were nausea and diarrhea. In the 10 mg/day group and 2% withdrew because of nausea and diarrhea respectively.

Table 3.13d: Adverse events

Drug	Trial	Design	Group 1	Group 2	Group 3	Placebo	Severity
Donepezil	Burnset et al. (19)	*Parallel*	Donepezil (5 mg/day) (N=271) Nausea 7% Diarrhoea 10% Vomiting 4% Anorexia 4% Dizzines 5% Confusion 7% Insomnia 7%	Donepezil (10 mg/day) (N=273) Nausea 24% Diarrhoea 16% Vomiting 16% Anorexia 8% Dizzines 9% Confusion 6% Insomnia 8%		Placebo (N=274) Nausea 7% Diarrhoea 4% Vomiting 4% Anorexia 1% Dizzines 5% Confusion 6% Insomnia 4%	Most adverse events considered mild and transient (no further breakdown provided). Nine percent of patients (n=818) had severe complications, although there was no specification of what these events may have been. Dropouts due to adverse events: Placebo: 10% 5 mg/day: 9% 10 mg/day: 18% Types of adverse events not stated.
Metrifonate	Becker et al. (21)	*Parallel*	Metrifonate (N=27) Upset stomach 7% Chest discomfort 15% Headache 7% Hiccups 4% Loss of hair 4% Sores 4%			Placebo (N=23) Nausea 13% A trial fibrillation 9% Headache 9% Sleep disturbance 4%	All adverse events regarded as probably or definitively associated with metrifonate were mild and transient. No other specifics provided. Not terminations of treatment due to adverse events.

Table 3.13e: Adverse events

Drug	Trial	Design	Group 1	Group 2	Group 3	Placebo	Severity
Metrifonate	Becker et al. (26)	*Parallel*	Metrifonate (N=23) 14 events reported. Type of events not reported.			Placebo (N=24) 8 events reported. Type of event not reported.	All adverse events regarded as probably or definitively associated with metrifonate were mild and transient. No other specifics provided. No alterations of treatment due to adverse events.
Metrifonate	Cummings et al. (23)	*Parallel*	Metrifonate Low dose (N=121) Abdominal pain 2% Diarrhea 9% Flatulence 1% Nausea 2% Leg cramps 1% Liver function abnormalities 0	Metrifonate Mid dose (N=120) Abdominal pain 4% Diarrhea 11% Flatulence 3% Nausea 8% Leg cramps 3% Liver function abnormalities 1%	Metrifonate High dose (N=119) Abdominal pain 12% Diarrhea 19% Flatulence 6% Nausea 16% Leg cramps 8% Liver function abnormalities 0	Placebo (N=120) Abdominal pain 3% Diarrhea 8% Flatulence 1% Nausea 9% Leg cramps 1% Liver function abnormalities 0	Group 1 / Group 2 / Group 3 / Placebo Mild 49% / 58% / 66% / 59% Mod. 23% / 35% / 34% / 26% Severe 6% / 7% / 8% / 5% n=121 / n=120 / n=119 / n=120 Placebo: 4% Low dose: 5% Mid dose: 7% High dose: 7% 3 patients discontinued because of asymptomatic bradycardia during the loading phase of the study.
Metrifonate	Morris et al. (25)	*Parallel*	Metrifonate (N=273) Nausea 12% Diarrhea 18% Leg cramps 9% Rhinitis 7% 325 adverse events reported.			Placebo (N=135) Nausea 10% Diarrhea 8% Leg cramps 4% Rhinitis 1% 148 adverse events reported.	Group 1 / Placebo Mild 71% / 67% Mod. 40% / 38% Severe 8% / 5% n=273 / n=135 Dropouts due to adverse events: Placebo: 4% Metrifonate: 12.4% Types of adverse events not reported

Table 3.13f: Adverse events

Drug	Trial	Design	Group 1	Group 2	Group 3	Placebo	Severity
Metrifonate	Raskind et al. (22)	*Parallel*	Metrifonate (N=177) Abdominal pain 10% Leg cramps 10% Agitation 8% Rhinitis 10% 216 Adverse events reported.			Placebo (N=87) Abdominal pain 2% Leg cramps 2% Agitation 2% Rhinitis 2% 84 Adverse events reported.	Group 1 / Placebo Mild 72% / 67% Mod. 43% / 23% Severe 7% / 7% n=177 / n=87 Dropouts due to adverse events: Placebo: 9% Metrifonate: 11% Types of adverse events not reported.
Metrifonate	Jann et al. (20)	*Parallel*	Metrifonate Loading dose (N=132) Diarrhea 14% Nausea 10% Leg cramps 4%	Metrifonate No loading dose (N=128) Diarrhea 7% Nausea 4% Leg cramps 0%		Placebo (N=133) Diarrhea 8% Nausea 3% Leg cramps 2%	Group 1 / Group 2 / Placebo Mild 55% / 49% / 56% Mod. 29% / 27% / 18% Severe 7% / 2% / 0% n=132 / n=128 / n=133 Dropouts due to adverse events: Placebo: 1.5% No loading: 4% Loading dose: 6% Types of adverse events not stated.

Table 3.13g: Adverse events

Drug	Trial	Design	Group 1	Group 2	Group 3	Placebo	Severity
Rivastigmine	Corey-Bloom et al. (29)	*Parallel*	Rivastigmine 1-4 mg (N=233) <u>Titration phase</u> Sweating 2% Fatigue 5% Asthenia 2% Weight decrease 1% Malaise 1% Allergy 2% Hypertension 4% Dizziness 15% Somnolence 7% Nausea 14% Vomiting 7% Anorexia 8% Flatulence 2% <u>Maintenance phase</u> Dizziness 8% Nausea 8% Vomiting 5% Dyspepsia 6% Sinusitis 1%	Rivastigmine 6-12 mg (N=231) Sweating 6% Fatigue 10% Asthenia 10% Weight decrease 4% Malaise 3% Allergy 0% Hypertension 3% Dizziness 24% Somnolence 9% Nausea 48% Vomiting 27% Anorexia 20% Flatulence 5% Dizziness 14% Nausea 20% Vomiting 16% Dyspepsia 5% Sinusitis 4%		Placebo (N=235) Sweating 2% Fatigue 4% Asthenia 2% Weight decrease 1% Malaise 1% Allergy 0% Hypertension 1% Dizziness 13% Somnolence 2% Nausea 11% Vomiting 3% Anorexia 3% Flatulence 1% Dizziness 4% Nausea 3% Vomiting 2% Dyspepsia 1% Sinusitis 1%	Only mention that the majority of reported nausea and vomiting was mild or moderate in severity. No other discussion of severity is provided. Dropouts due to adverse events: Placebo: 7% 1-4 mg: 8% 6-12 mg: 28% Types of adverse events not stated.

Table 3.13h: Adverse events

Drug	Trial	Design	Group 1	Group 2	Group 3	Placebo	Severity
Rivastigmine	Rösler et al. (31)	*Parallel*	Rivastigmine Low dose (N=243) Nausea 17% Vomiting 8% Dizziness 10% Headache 7% Diarrhoea 10% Anorexia 3% Abdominal pain 5% Fatigue 2% Malaise 1%	Rivastigmine High dose (N=243) Nausea 50% Vomiting 34% Dizziness 20% Headache 19% Diarrhoea 17% Anorexia 14% Abdominal pain 12% Fatigue 10% Malaise 10%		Placebo (N=243) Nausea 10% Vomiting 6% Dizziness 7% Headache 8% Diarrhoea 9% Anorexia 2% Abdominal pain 3% Fatigue 3% Malaise 2%	Only mention of severity is that most adverse events (none specified) were not sure. Dropouts due to adverse events: Placebo: 7% Low dose: 7% High dose: 23% Most dropouts due to adverse events occurred during dose escalation. Types of averse events not stated.
Ginkgo Biloba	Maurer et al. (94)	*Parallel*	Ginkgo Biloba (N=9) None stated.			Placebo (N=9) None stated.	None stated.
Ginkgo Biloba	Kanowski et al. (89)	*Parallel*	Ginkgo Biloba (N=79 completers) 63 adverse events recorded. Diseases of the skin: 6% GI disorders: 5%			Placebo (N=77 completers) 59 adverse events recorded. Diseases of the skin: 1% GI disorders: 11%	6 patients with serious adverse events + 1 patient with non-serious adverse events discontinued treatment prematurely. Type of adverse events not stated.
Ginkgo Biloba	Le Bars et al. (87)	*Parallel*	Ginkgo Biloba (N=155) 30% reported at least 1 adverse event. 97 adversse events. Types of adverse events not stated.			Placebo (N=154) 31% reported at least 1 adverse event 91 adverse events Types of adverse events not stated.	Dropouts due to adverse events: Placebo: 2.5% Ginkgo Biloba: 6% 1 death in the placebo group; 2 deaths in the treated groups. 1 stroke and 1 subdural hematoma in the placebo group.

Table 3.13i: Adverse events

Drug	**Trial**	**Design**	**Group 1**	**Group 2**	***Group 3***	**Placebo**	**Severity**
Selegiline	Sano et al. (40)	*Parallel*	Selegiline (10mg/day) Dental events 7%; Fall 9% Syncope 10%			Dental events 0%; Fall 5% Syncope 4%	Not stated
Selegiline	Mangoni et al. (43)	Parallel	*Selegiline (10mg/day)* Anxiety 1.5%; Allergy 1.5% Psychomotor agitation 4.4% GI heaviness 1.5 Heartburn 1.5% Dyspepsia 1.5%			Anxiety 2.0%; Agitation 3.9% Dizziness 2.0% GI 5.9%	*Selegiline:* Moderate (25%), Moderately severe (62.5%), Severe (12.5%) *Placebo:* Moderate (71.4%), Moderately severe (14.3%), Severe (14.3%)
Selegiline	Freedman et al. (42)	Parallel	*Selegiline (10mg/day)* Dizziness 30%; Irritability 23% Insomnia 15%; Restlessness 19% Mood changes 19% Hallucinations 8%;Confusion 50% Shakiness 4%; Headaches 12% Palpitations 4%;Bowel problems 15% Dry mouth 27%; Weight loss 8% Change in appetite 8% Rash 4%; Impotence 4%			Dizziness 20%; Irritability 16% Insomnia 16%; Restlessness 12% Mood changes 16% Hallucinations 20% Confusion 52% Shakiness 8% Headaches 20% Palpitations 4%; Bowel problems 8% Weight loss 12%; Rash 4% Impotence 4%; Other 12%	- No subject withdrawn because of AE - Severity class not reported
Selegiline	Burke et al. (41)	Parallel	Not reported			Not reported	No serious AE
Selegiline	Lawlor et al. (46)	Crossover	Not reported			Not reported	No « significant » AE No drop-outs because of AE
Selegiline	Tariot et al. (45)	Crossover	Not reported			Not reported	1 depression (x%) which required use of AD: not stated during which phase it occurred

Table 3.13j: Adverse events

Drug	Trial	Design	Group 1	Group 2	Group 3	Placebo	Severity
Selegiline	Schneider et al. (44)	Crossover	Nausea 10%				2 out of 3 patients on physostigmine experienced mild intermittent nausea throughout the trial
Ginkgo Biloba	Le Bars et al. (87)	Parallel	*Egb (120 mg/day)* Death 1.3% Severe AE 7.7% GI AE 11.6%			Death 0.6%; Stroke 0.6% ; subdural hematoma 0.6%; Severe AE 5.8% GI AE 7.1%	5 serious AE 3/21 severe AE led to withdrawal (2 Egb, 1 placebo)
Ginkgo biloba	Maurer et al. (88)	Parallel	*Egb (240 mg/day)* None observed			None observed	
Ginkgo biloba	Kanowski et al. (89)	Parallel	*Egb (240 mg/day)* Total 63 AE (not listed) GI events: 5.1% Possible AE: 5 patients including: Allergic skin reactions GI complaints Headache			Total: 59 AE (not listed) GI events: 11.7% Probable: 1 patient which was skin reaction	7 serious (5 in Egb group): 6 patients with serious and 1 with non-serious withdrew 3 serious AE during run-in period under placebo (imputability ruled out)

4 SUMMARY OF RESULTS OF THE BEST EVIDENCE SYNTHESIS

In this chapter we summarize the findings from the in-depth reviews of 9 treatments, or potential treatments, for AD based on published data. The blinded review and methodological scoring process ensured that only those clinical trials meeting rigorous standards of methodology were included. A total of 36 randomized clinical trials (either parallel group or crossover design) met the minimum criteria for inclusion based on blinded review. During the consensus conference that followed the independent blinded reviews, however, 12 of these trials (1 rivastigmine, 1 selegiline, 3 lecithin, and 7 gingko biloba) were excluded as it was determined that they did not meet the minimum criteria or there was serious concern about the diagnostic criteria used. This finding emphasizes the need for a consensus conference in which all trials submitted for blinded review are subsequently discussed by all reviewers prior to inclusion. Following a review of the first draft of this report, two additional trials of metrifonate were obtained and, as they both met the methodological criteria, were included. Thus 26, randomized clinical trials were included for the final report.

In a previous report (5) on the efficacy of tacrine for the treatment of AD, we concluded that the studies were so variable in dose, treatment duration, and outcome measures that to combine them to obtain single summary estimates of the efficacy of tacrine was not justifiable and not methodologically appropriate. It was our hope that since that report in mid 1997, the conduct of clinical trials of antidementia agents had become more standardized. Overall, we found that the methodological quality of the trials reviewed in this report was, in fact, superior to those reviewed to assess the efficacy tacrine. This is likely due, at least in part, to the fact that the majority of the trials reviewed in this report were published in 1995 or later. Selegiline was the one exception for which clinical trials were published between 1991 and 1998. The single vitamin E trial was published in 1992. In addition, several trials on the same agent (e.g. donepezil, rivastigmine, and metrifonate) were carried out by the same or overlapping groups of investigators, a feature that we anticipated would help to enhance comparability. We were frustrated, however, in our efforts to compare results from these studies due to

considerable differences in reporting strategies even from the same group of investigators.

The variety of outcome measures to assess cognitive performance, functional abilities and behavior has, and continues to be, impressive and clearly adds to the complexity of drawing overall conclusions about treatment efficacy. As the choice of outcome measure is crucial to the assessment of the efficacy of antidementia medications, this study includes an extensive report of the psychometric properties of the primary and secondary outcome measures used in the trials included in the review. See II: A Review of Outcome Measures in Clinical Trials.

Our original intent was to carry out quantitative meta-analysis of trials within treatments where warranted. However, after careful consideration of the differences in design and, particularly trial reporting, we determined that meta-analyses even across trials of the same pharmaceutical agent were not justifiable and in many cases, not possible due to lack of details in the trail reports. Although, it has become common practice to combine results from trials, there are concerns regarding the factors that may cause bias in a meta-analysis. In general, in conducting a meta-analysis one should pay particular concern to publication bias, quality assessment, and trial heterogeneity.

Publication bias: Our review was based entirely upon full published reports of trials and it is indeed likely that this criterion resulted in the exclusion of a number of trials. This exclusion argues against the conduct of a meta-analysis as exclusion of unpublished studies may lead to selection bias.

Quality assessment: Using the Jadad scale, we feel confident that from within the trials collected, we have selected those with the most rigorous methodology. However, the results of a meta-analysis will differ depending upon how the quality scores are used. We used the score as a cut-off and did not incorporate any information from trials that did not meet this cutoff. There was variation in the quality scores of the included trials and this was not taken into account in our review.

Trial heterogeneity: In some instances where the design of the trials was similar, reporting of results varied to the extent that in some trials baseline characteristics for randomized subjects were not reported in sufficient detail for adjustments to be made in

the analysis (e.g. (19). The exclusion of 1 out of 4 trials for this reason limits the utility of a meta-analysis.

In addition, following methodological review and the consensus conference, only a single trial remained for indepth review for Vitamin E (40), lecithin (73) and linopirdine (85) precluding the need for a meta-analysis.

Selegiline, although promising with its antioxidant activity, was not convincingly shown to be efficacious in any of the six trials that were reviewed. Some improvement was found on a few items of the cognitive scales used but these changes did not correlate with changes in patient behavior. Despite the lack of efficacy, it is notable that this drug exhibited a low incidence of adverse effects, none of which were serious enough to interrupt treatment.

Apart from an increase in the risk of falls and syncope, ∝-tocopherol (vitamin E) also appeared to be well tolerated. This treatment resulted in a statistically significant effect on the time from entry-to-trial to the occurrence of any of the events, which constituted the primary outcome. The choice of the primary outcome defined as time to: death, institutionalization; loss of ability to perform at least 2 out of the 3 of eating, grooming, or toileting as measured by the BDS; or severe dementia as measured by a CDR of 3 is problematic and renders interpretation of the results difficult. A sub-analysis by specific outcome revealed that for the ∝-tocopherol group there was a significant delay in the time to institutionalization. It is well known that the risk of institutionalization is related to a number of individual and family related variables and these were not accounted for (in fact, not measured) in this report. Furthermore, it is curious that with respect to this outcome, treatment with ∝-tocopherol alone appeared to have a greater effect than a combined treatment with ∝-tocopherol and selegiline.

With respect to lecithin, it is uncertain whether the selection procedure, based on good responses to tacrine, may have prejudiced the results for any comparisons with lecithin. A four-week duration of treatment may be too short to document beneficial effects, and crossover trials may be inappropriate for diseases that steadily progress, such as AD. Data from the second phase of the study were not presented and there is no clear placebo comparison. The study does not support a beneficial role for choline supplementation alone, which is in keeping with the results of earlier and methodically

weaker trials (95). Perhaps the cholinesterase hypothesis leading to such comparisons oversimplifies the true disease process of AD (69;95).

Based on only one trial, linopirdine does not appear to have beneficial effects for people with AD. This, along with the associated risk of liver compromise, casts doubt on the drug's clinical usefulness.

Propentofylline appears to have small beneficial global and cognitive effects with long-term administration, although the clinical relevance of these effects is uncertain. The drug appears to have a safe profile even at one year of continuous use.

The summary below focuses on the three AChE inhibitors (donepezil, metrifonate, and rivastigmine) and gingko biloba. We concluded that for selegiline, vitamin E, lecithin, linopirdine, and propentofylline the published data do not provide support for efficacy.

4.1 Donepezil

Based on the results of the phase II trial (16) using doses of 1 mg, 3 mg and 5 mg/day subsequent phase III trials (19) (18) (17) all used 5 mg/day and 10 mg/day dosage groups.

Overall, these trials are consistent in the finding that donepezil is a well tolerated drug at 5 and 10 mg/day and has moderate (statistically significant) effects both on cognition and the global clinical status of the patients over 12 and 24 weeks. The drug is administered once daily. The improvements in cognition as measured by the ADAS-cog were only slightly greater in the 10 mg treated patients than in the 5 mg treated patients; however, there were more withdrawals in the high dosage groups in two of the studies (17;18). The high prevalence of adverse events in the 10 mg group in Rogers et al. (17) is likely due to the forced rapid titration from 5 mg/day to 10 mg/day over 1 week. Treatment effects were virtually identical (2.88, 2.9, 3.1 points) using 10 mg/day over 12 or 24 weeks. The absence of specific data in the study of Burns et al. (19) makes direct comparisons with this trial impossible.

All of the trials used the last observation carried forward approach for dropouts and this may have resulted in an overestimation of a treatment effect. This is of

particular importance in Rogers et al. (17) in which only 68% of the 10 mg/day group completed the trial as compared to 80% of the placebo group. The concern arises that the time on treatment is not the same for the comparison groups and that ADAS-cog scores measured earlier in the trial for the treated group are higher than one would expect at the end of study as deterioration is underestimated. This discrepancy may result in a spurious increase in the observed difference between the treated and placebo group. The deterioration in the placebo groups in the 12 week trials was less than that in the 24 week trials but the treatment effect was greater resulting in similar findings across trials of differing length. It has been suggested that a change of 4 points on the ADAS-cog is expected to occur naturally in untreated AD patients (see report II), and that an antidementia medication should at least reverse this change. None of the donepezil trials reached this expectation.

With respect to global functioning, all of the donepezil trials reported a statistically significant difference between treatment and placebo group on the CGIC (16) or the CIBIC-plus (17-19). The absolute value of the differences are remarkably similar and do not vary by dosage group or trial duration. On average a 0.3 to 0.4 point difference between the placebo and treatment groups was observed. Findings presented in this way are difficult to interpret clinically. However, presented in another way, there was a statistically significant difference between the 5 mg/day, 10 mg/day groups and the placebo groups in all trials with respect to the percentage of participants who improved or did not change on the global scale. There was, however, no difference between the two treatment groups.

Overall, donepezil produced modest effects on both cognitive performance and global functioning. There is little evidence, however, of any improvement in efficacy with the 10 mg/day dose over the 5 mg/day dose. This should be investigated further. The high drop out rates, particularly in Rogers et al. (17) compromise the validity of the efficacy results of this particular study and the comparability with the other studies.

4.2 Metrifonate

The results of this review indicate that metrifonate given once daily is quite well tolerated with only a small proportion of subjects withdrawing from the trials due to adverse events. All trials reported a statistically significant effect of treatment on ADAS-cog scores at the end of the double blind phase. The magnitude of the difference, however, varied across the trials. In general, the placebo groups deteriorated over time while the treated groups either did not change or improved slightly. The magnitude of the effects on the ADAS-cog was maximal (at equivalent doses) at the highest dose whether the study was of 12 weeks or 26 weeks duration. Raskind et al. (22), however, did not report sufficient data to compare the magnitude of the effect in this trial to that of the other trials. In general, in comparison with placebo patients, who deteriorated over time, the metrifonate treated patients either did not change or improved slightly. In two trials of 26 weeks with differing dosages (both with loading dose phases) the average differences between the highest treatment dose and placebo were 2.94 points (23) and 2.86 points (25). These values are remarkably similar to those seen for donepezil but again, the treatment success criterion of reversal of the disease process by 6 months (e.g., 4 points on the ADAS-cog) was not reached.

All of the trials found a significant treatment effect on global function as measured by the CIBIC-plus. The treatment:placebo differences between baseline and end of the double blind phase were very similar. Morris et al.(25) report a treatment:placebo difference of 0.28 points (95% CI: 0.06 to 0.50), while Raskind et al. (22) report a difference of 0.2 points ($p<0.05$). Comparable differences were found in Cummings et al. (23) for the mid dose group (treatment:placebo difference of 0.29, 95% CI: 0.09 to 0.48) and the high dose group (treatment:placebo difference of 0.35, 95% CI: 0.15 to 0.54). The study by Jann et al. (20) demonstrated that a loading dose is not required.

Based on the review of the published literature the efficacy of metrifonate appears to be quite similar to that of donepezil. In a personal communication from Bayer we were informed that "metrifonate is currently on clinical hold to assess concerns about rare cases of proximal muscle weakness".

4.3 Rivastigmine

The two studies reviewed were designed in such a way as to enhance the comparability of results. However, the disparate methods of reporting of the two trials make it difficult to compare the magnitude of the results. Corey-Bloom et al. (29) report data on a variety of patient characteristics at baseline. Rösler et al. (31), however, present baseline data only on the ADAS-cog and the PDS. The two studies suffer from a high drop-out rate in the high dose group because of the fixed titration regimen (the very group for which differences are found) and the use of a variety of statistical analyses of efficacy. Our in-depth review reveals a moderate benefit on both cognition and global clinical status from high dose twice daily treatment with rivastigmine. Both studies included a heterogeneous mix of study subjects and more than 90% of subjects suffered from other medical conditions and/or used medications, theoretically enhancing the generalizability of the results. In the Corey-Bloom et al. (29) study, there was a high dose treatment: placebo difference of 3.78 points (95% CI: 2.69 to 4.87) over 26 weeks using an ITT analysis, larger than that found with either donepezil or metrifonate. In the Rösler et al. (31) the magnitude of the effect was much smaller for the high dose group with a high dose treatment:placebo difference of 1.6 points ($p<0.10$). The difference was not statistically significant for the low dose group.

A statistically significant effect of rivastigmine on global function as measured by the CIBIC-plus was observed in both trials. There was very little difference in the treatment:placebo difference from baseline in either of the two treatment groups in Corey-Bloom et al. (29) (e.g., 0.29 for the high dose group and 0.26 for the low dose group). Similar differences were reported in Rösler et al. (31) with a high dose to placebo difference of 0.47 and a low dose to placebo difference of 0.33, estimated from the available data.

Both studies report an improvement on the Progressive Deterioration Scale for those assigned to the high dose group as compared to placebo with no apparent benefit from low dose rivastigmine. In the Rösler et al. study these effects are most apparent using the LOCF analysis whereas the ITT analysis produced similar results in the U.S. Study.

The PDS was developed as a quality of life scale for AD patients (34) and its use in trials as a measure of activities of daily living has been criticized (36).

4.4 Ginkgo Biloba

Ginkgo biloba is widely available and recommended for use in Europe for cerebral insufficiency, which is suggested as an important feature in dementia. The trials reviewed here are of differing length, use two different dosages (of the two trials that used the same dosage, one administered the medication twice daily, the other three times daily), use a variety of primary and secondary outcomes, and vary greatly in sample size. In addition, only one of the trials (88) is restricted to subjects with AD and detailed information on drop-outs, baseline characteristics, and adverse effects is not adequately presented for AD patients in the two other trials. All of the above limitations make general conclusions on the efficacy of ginkgo biloba on treatment of AD difficult. Overall, it appears as if ginkgo biloba at 120 mg per day (87) has a modest effect on cognitive function as evidenced by a treatment: placebo baseline to endpoint differences of 1.7 points (95% CI: 0.2 to 3.2) on the ADAS-cog. This effect is markedly smaller than that seen for the AChE inhibitors.(as measured by the ADAS-cog or the SKT) in patients with AD. These results, however, must be interpreted in light of the unacceptably high dropout rates in both the treatment (50%) and placebo (62%) groups. Maurer et al. (88) reported a difference of 0.91 points on the ADAS-cog in their 12 week trial of only 20 AD subjects using 240 mg/day. Using the SKT as a measure of cognitive performance, Kanowski et al. (89) and a treatment regimen of 240 mg/day found a higher percentage of patients in the treatment group improved at least 4 points on the SKT (48%) than in the placebo group (18%), $p<0.05$. In addition a treatment:placebo difference of 1.5 points on the SKT was observed in this 24 week trial. In the shorter 12 week trial (88) using the same dosage, the difference was 3.67 ($p<0.013$) in favour of the treated group The dramatic difference in the efficacy evidence from these two trials with similar dosage must be interpreted in the light of the difference in disease severity of the study subjects in the two trials. Those in the Maurer et al. (88) trial appeared to be considerably more impaired cognitively than participants in the other two trials.

There was little or no impact of treatment with EGb on global impairment.

The use of primary outcome measures for which there is no English language literature (e.g., SKT) inhibits clear interpretation of the results of these three gingko biloba trials that met the minimum criteria for inclusion. Clinical trials using more widely accepted outcome measures, restricted to subjects with probable AD could greatly enhance this area of research.

4.5 Final Thoughts

In the past decade there have been numerous advances in the development of potential treatments for AD. In addition, the design, conduct, and reporting of clinical trials to assess these potential therapies has improved and has resulted in a greater number of trials that could be selected as "best evidence". Our review considered only full published reports of trials and it certainly is possible (and in fact, likely) that excellent ongoing or unpublished completed trials have been missed. Based on the evidence we reviewed, it is our conclusion that donepezil, metrifonate and rivastigmine all provide statistically significant modest benefit on cognitive performance and global functioning to the elderly with probable AD who are eligible for inclusion in clinical trials. In the absence of trials of longer duration, there is little evidence that these agents are more than symptomatic treatments. The efficacy of gingko biloba as a treatment for AD appears to be less than that of any of the above AChE inhibitors.

Although all of these medications appear to be well tolerated, in terms of the occurrence of adverse events, dropout rates are sometimes high and may have resulted in overestimation of apparent treatment effects. In a companion report, we review the psychometric properties of the primary and secondary outcome measures used in these trials and we remain concerned about the wide variety of scales used that do not have adequate psychometric assessment. Although a great number of the scales used have evidence of reliability and validity, responsiveness to change has not been assessed for many of the outcome measures. For this reason, the clinical significance of the treatment: placebo differences remains unclear. The ability to carry out a comparative analysis of therapies for AD depends not only on the comparability of design, duration, and outcome

measures used but also on the methods of reporting the results of the trials. There was no consistent method of reporting the results of the AD trials even when emanating from the same group of investigators. Although journals do have different guidelines for authors, we believe that the variability in reporting crucial details concerning the participants and the trial results, is unacceptable.

5 REFERENCE LIST

(1) Anonymous. Canadian study of health and aging: Methods and prevalence of dementia. CMAJ 1994; 150(6):899-913.

(2) Teng EL, Chui HC. The Modified Mini-Mental State (3MS) Examination. J Clin Psychiatry 1987; 48:314-318.

(3) American Psychiatric Association. Diagnostic and Statistical Manual for Mental Disorders. 3 ed. Washington, DC: American Psychiatric Association, 1987.

(4) McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984; 34:939-944.

(5) Wolfson C, Moride Y, Perrault A, Vida S. A Study of the Efficacy, Effectiveness and Economic Impact of Tacrine in Alzheimer's Disease. 1997. Ottawa, Canadian Coordinating Office for Health Technology Assessment.

(6) Slavin RE. Best evidence synthesis: An intelligent alternative to meta-analysis. J Clin Epidemiol 1995; 48(1):9-18.

(7) Nolan KA, Black RS, Sheu KF et al. A trial of thiamine in Alzheimer's disease. Arch Neurol 1991; 48(1):81-83.

(8) Jadad AR, Moore A, Carroll D et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Control Clin Trials 17, 1-12. 1996.

(9) Moher D, Jadad AR, Nichol G et al. Assessing the quality of randomized controlled trials: An annotated bibliograhpy of scales and checklists. Control Clin Trials 16, 62-73. 1995.

(10) DerSimonian R, Charette LJ, McPeek B. Reporting on methods in clinical trials. New Engl J Med 1982; 306:1332-1337.

(11) Clark HD, Wells GA, Huët C, McAlister FA, Salmi R, Fergusson D et al. Assessing the quality of randomized trials: Reliability of the Jadad scale. Control Clin Trials 1999; 20:448-452.

(12) Hofferberth B. The Efficacy of EGb 761 in patients with senile dementia of the Alzheimer type: A double-blind placebo-controlled study on different levels of investigation. Human Psychopharmacol 1994; 9:215-222.

(13) Wesnes K, Simmons D, Rock M, Simpson P. A double-blind placebo-controlled trial of tanakan in the treatment of idiopathic cognitive impairment in the elderly. Human Psychopharmacol 1987; 2:159-169.

(14) Rai GS, Shovlin C, Wesnes KA. A double-blind placebo-controlled study of ginkgo biloba extract ("Tanakan") in elderly outpatients with mild to moderate memory impairment. Can Med Res Opin 1991; 12:350-355.

(15) Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979; 86:420-428.

(16) Rogers SL, Freidhoff LT. 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 1996; 7(6):293-303.

(17) Rogers SL, Farlow MR, Doody RS et al. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer‘s disease. Donepezil Study Group. Neurology 1998; 50(1):136-145.

(18) Rogers SL, Doody RS, Mohs RC. Donepezil improves cognition and global function in Alzheimer disease, a 15-week, double-blind, placebo-controlled study. Donepezil Study Group. Arch Intern Med 1998; 158(9):1021-1031.

(19) Burns A, Rossor M, Hecker J et al. The effects of donepezil in Alzheimer’s disease - Results from a multinational trial. Dement Geriatr Cogn Disord 1999; 10:237-244.

(20) Jann MW, Cyrus PA, Eisner LS, Margolin DI, Griffin T, Gulanski B. Efficacy and safety of a loading-dose regimen versus a no-loading dose regimen of metrifonate in the symptomatic treatment of Alzheimer’s disease: a randomized, double-masked, placebo-controlled trial. Metrifonate Study Group. Clin Ther 1999; 21(1):88-102.

(21) Becker RE, Colliver JA, Markwell SJ et al. Double-blind, placebo-controlled study of metrifonate, an acetylcholinesterase inhibitor, for Alzheimer‘s disease. Alzheimer Dis Assoc Disord 1996; 10(3):124-131.

(22) Raskind MA, Cyrus PA, Ruzicka BB, Gulanski BI. The effects of metrifonate on the cognitive, behavioral, and functional performance of Alzheimer’s disease patients. Metrifonate Study Group. J Clin Psychiatry 1999; 60(5):318-325.

(23) Cummings JL, Cyrus PA, Bieber F et al. Metrifonate treatment of the cognitive deficits of Alzheimer‘s disease. Metrifonate Study Group. Neurology 1998; 50(5):1214-1221.

(24) Kaufer D. Beyond the cholinergic hypothesis: The effect of metrifonate and other cholinesterase inhibitors on neuropsychiatric symptoms in Alzheimer‘s disease. Dement & Geriatr Cog Disord 1998; 9(Suppl. 2):8-14.

(25) Morris JC, Cyrus PA, Orazem J et al. Metrifonate benefits cognitive, behavioral, and global function in patients with Alzheimer‘s disease. Neurology 1998; 50(5):1222-1230.

(26) Becker RE, Colliver JA, Markwell SJ et al. Effects of metrifonate on cognitive decline in Alzheimer‘s disease: A double-blind, placebo-controlled, 6-month study. Alzheimer Dis Assoc Disord 1998; 12(1):54-57.

(27) Pettigrew LC, Beiber F, Lettieri J et al. Pharmacokinetics, pharmacodynamics, and safety of metrifonate in patients with Alzheimer‘s disease. J Clin Pharmacol 1998; 38(3):236-245.

(28) Becker RE, Colliver J, Elble R et al. Effects of metrifonate, a long-acting cholinesterase inhibitor, in Alzheimer‘s disease. Drug Dev Res 1990; 19:425-434.

(29) Corey-Bloom J, Anand R, Veach J et al. 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 Psychiatry 1998; 1:55-65.

(30) Sramek JJ, Anand R, Wardle TS et al. Safety/tolerability trial of SDZ ENA-713 in patients with probably Alzheimer‘s disease. Life Sci 1996; 58(15):1201-1207.

(31) Rosler M, Anand R, Cicin-Sain A et al. Efficacy and safety of rivastigmine in patients with Alzheimer‘s disease: International randomized controlled trial. BMJ 1999; 318(7184):633-640.

(32) Reisberg B, Ferris SH, de Leon MJ, Crook T. Global Deterioration Scale (GDS). Psychopharmacol Bull 1988; 24(4):661-663.

(33) Schneider LS, Anand R, Farlow MR. Systematic review of the efficacy of rivastigmine for patients with Alzheimer's disease. Int J Geriatr Psychopharm 1998; 1 (Suppl. 1):S26-S34.

(34) DeJong R, Osterlund OW, Roy GW. Measurement of quality-of-life changes in patients with Alzeimer's disease. Clin Ther 1989; 11(4):545-554.

(35) Knopmann D, Gracon S. Observations on the short-term 'natural history' of probable Alzheimer's disease in a controlled clinical trial. Neurology 1994; 44(2):260-265.

(36) Sellwood E. Rivastigmine in Alzheimer's disease. eBMJ . 29-3-1999. www.bmj.com/cgi/eletters/318/7184/633.

(37) Monteverde A, Gnemmi P, Rossi F et al. Selegiline in the treatment of mild to moderate Alzheimer-type dementia. Clin Ther 1990; 12(4):315-322.

(38) Campi N, Todeschini GP, Scarzella L. Selegiline versus l-acetylcarnitine in the treatment of Alzheimer-type dementia. Clin Ther 1990; 12(4):306-314.

(39) Falsaperla A, Monici-Preti PA, Oliani C. Selegiline versus oxiracetam in patients with Alzheimer-type dementia. Clin Ther 1990; 12(4):376-384.

(40) Sano M, Ernesto C, Thomas RG et al. A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. N Engl J Med 1997; 336(17):1216-1222.

(41) Burke WJ, Roccaforte WH, Wengel SP et al. L-deprenyl in the treatment of mild dementia of the Alzheimer type: Results of a 15-month trial. J Am Geriatr Soc 1993; 41(11):1219-1225.

(42) Freedman M, Rewilak D, Cohen S et al. L-deprenyl in Alzheimer‘s disease: Cognitive and behavioral effects. Neurology 1998; 50(3):660-668.

(43) Mangoni A, Grassi MP, Frattola L et al. Effects of a MAO-B inhibitor in the treatment of Alzheimer disease. Eur Neurol 1991; 31(2):100-107.

(44) Schneider LS, Olin JT, Pawluczyk S. A double-blind crossover pilot study of l-deprenyl (selegiline) combined with cholinesterase inhibitor in Alzheimer‘s disease. Am J Psychiatry 1993; 150(2):321-323.

(45) Tariot PN, Goldstein B, Podgorski CA et al. Short-term administration of selegiline for mild-to-moderate dementia of the Alzheimer‘s type. Am J Geriatr Psychiatry 1998; 6(2):145-154.

(46) Lawlor BA, Aisen PS, Green C et al. Selegiline in the treatment of behavioral disturbance in Alzheimer‘s disease. Int J Geriatric Psychiatry 1997; 12:319-322.

(47) Rubin EH, Morris JC, Grant EA, Vendegna T. Very mild senile dementia of the Alzheimer type. Arch Neurol 1989; 46:379-382.

(48) Eisdorfer C, Cohen D, Paveza GJ et al. An empirical evaluation of the Global Deterioration Scale for staging Alzheimer’s disease. Am J Psychiatry 1992; 149(2):190-194.

(49) Berg L, Miller JP, Storandt M et al. Mild senile dementia of the Alzheimer type: 2. Longitudinal assessment. Ann Neurol 1988; 23(5):477-484.

(50) Hughes CP, Berg L, Danziger WL et al. A new clinical scale for the staging of dementia. Brit J Psychiat 1982; 140:566-572.

(51) Kramer-Ginsberg E, Mohs RC, Aryan M et al. Clinical predictors of course of Alzheimer‘s patients in a longitudinal study: A preliminary report. Psychopharmacol Bull 1988; 24:458-462.

(52) Stern RG, Mohs RC, Davidson M et al. A longitudinal study of Alzheimer‘s disease: Measurement, rate and predictors of cognitive deterioration. Am J Psychiatry 1994; 151:390-396.

(53) Small BJ, Viitanen M, Winbald B, Baeckman L. Cognitive changes in very old persons with dementia: The influence of demographic, psychometric and biological variables. J Clin Exp Neuropsychol 1997; 19:245-260.

(54) Clark CM, Sheppard L, Fillenbaum GG, Galasko D, Morris JC, Koss E et al. Variability in annual Mini-Mental State Examination score in patients with probable Alzheimer disease: A clinical perspective of data from the Consortium to Establish a Registry for Azlheimer’s Disease. Arch Neurol 1999; 56:857-862.

(55) Albert MS. Assessment of cognitive dysfunction. In: Albert MS, Moss MB, editors. Geriatric Neuropsychology. New York: The Guilford Press, 1988: 57-81.

(56) Cole MG. Interrater reliability of the Blessed Dementia Scale. Can J Psychiatry 1990; 35(4):328-330.

(57) Drachman DA, Swearer JM, Kane K et al. The Cognitive Assessment Screening Test (CAST) for dementia. J Geriatr Psychiatry Neurol 1996; 9(4):200-208.

(58) Heun R, Papassotiropoulos A, Jennssen F. The validity of psychometric instruments for detection of dementia in the elderly general population. Int J Geriatr Psychiatry 1998; 13(6):363-380.

(59) Losascio JJ, Growdon JH, Corkin S. Cognitive test performance in detecting, staging, and tracking Alzheimer's disease. Arch Neurol 1995; 52:1087-1099.

(60) Mack JL, Patterson MB. The evaluation of behavioral disturbances in Alzheimer's disease: The utility of three rating scales. J Geriatr Psychiatry Neurol 1994; 7:101-117.

(61) Ownby RL, Koss E, Smyth KA, Whitehouse PJ. The factor structure of the Brief Psychiatric Rating Scale in Alzheimer's disease. J Geriatr Psychiatry Neurol 1994; 7:245-250.

(62) Tariot PN, Mack JL, Patterson MB et al. The Behavior Rating Scale for Dementia of the Consortium to Establish a Registry for Alzheimer's Disease: The Behavioral Pathology Committee of the Consortium to Establish a Registry for Alzheimer's Disease. Am J Psychiatry 1995; 152:1349-1359.

(63) Tariot PN. CERAD Behavior Rating Scale for Dementia. Int Psychogeriatr 1996; 8 (Suppl. 3):514-518.

(64) Mohs RC, Rosen WG, Davis KL. The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. Psychopharmacol Bull 1983; 19(3):448-450.

(65) Rosen WG, Mohs RC, Davis K. A new rating scale for Alzheimer's disease. Am J Psychiatry 1984; 141:1356-1364.

(66) Mohs RC, Cohen L. Alzheimer's Disease Assessment Scale (ADAS). Psychopharmacol Bull 1988; 24(4):627-628.

(67) Ulus IH, Wurtman RJ. Prevention by choline of the depletion of membrane phosphatidylcholine by a cholinesterase inhibitor. N Engl J Med 1988; 318:191.

(68) Blusztajn JK, Holbrook PG, Lakher M et al. 'Autocannibalism' of membrane choline-phospholipids: Physiology and pathology. Psychopharmacol Bull 1986; 22:781-786.

(69) Bartus RT, Dean LR, Beer B, Lippa AS. The cholinergic hypothesis of geriatric memory dysfunction. Science 1982; 217:408-417.

(70) Gauthier S, Bouchard R, Lamontagne A et al. Tetrahydroaminoacridine-lecithin combination treatment in patients with intermediate stage Alzheimer‘s disease: Results of a Canadian double-blind, crossover, multicenter study. N Engl J Med 1990; 322(18):1272-1276.

(71) Fitten LJ, Perryman KM, Gross PL, Fine H et al. Treatment of Alzheimer‘s disease with short- and long-term oral THA and lecithin: A double-blind study. Am J Psychiatry 1990; 147(2):239-242.

(72) Weinstein HC, Teunisse S, van Gool WA. Tetrahydroaminoacridine and lecithin in the treatment of Alzheimer‘s disease. Effect on cognition, functioning in daily life, behavioral disturbances and burden experienced by the carers. J Neurol 1991; 238(1):34-38.

(73) Foster NL, Petersen RC, Gracon SI, Lewis K. An enriched-population, double-blind, placebo-controlled, crossover study of tacrine and lecithin in Alzheimer‘s disease. The tacrine 970-6 study group. Dementia 1996; 7(5):260-266.

(74) Dahlke F, Lohaus A, Gutzmann H. Reliability and clinical concepts underlying global judgements in dementia: Implications for clinical research. Psychopharmacol Bull 1992; 28(4):425-432.

(75) Parkinson FE, Rudolphi KA, Fredholm BB. HWA 285: A nucleoside transport inhibitor with neuroprotective effects in cerebral ischemia. Gen Pharmacol 1994; 6:1053-1058.

(76) Goto M, Demura N, Sakaguchi T. Effects of HWA 285 on disorder of learning and memory in rodents. Jpn J Pharmacol 1987; 45:373-378.

(77) Yamada Y, Nakamura H, Okada Y. HWA 285 enhances the formation of long-term portentiation in guinea pig hippocampal slices. Neuroscience Letters 1994; 176:189-192.

(78) Shinoda I, Furukawa Y, Furukawa S. Stimulation of nerve growth factor synthesis/secretion by HWA 285 in cultured mouse astroglial cells. Biochem Pharmacol 1990; 39:1813-1816.

(79) Nabeshima T, Nitta A, Hasegawa T. Impairment of learning and memory and the accessory symptoms in aged rats as senile dementia model: Oral administration of HWA 285 produces recovery of reduced NGF content in the brain of aged rats. Jpn J Psychopharmacol 1993; 13:89-95.

(80) Mielke R, Ghaemi M, Kessler J et al. Propentofylline enhanced cerebral metabolic response to auditory memory stimulation in Alzheimer‘s disease. J Neurol Sci 1998; 154(1):76-82.

(81) Marcusson J, Rother M, Kittner B et al. A 12-month, randomized, placebo-controlled trial of propentofylline (HWA 285) in patients with dementia according to DSM III-R. The European Propentofylline Study Group. Dement Geriatr Cogn Disord 1997; 8(5):320-328.

(82) Cook L, Nickolson VJ, Steinfels GF et al. Cognition enhancement by the acetylcholine releaser DUP 996. Drug Dev Res 1990; 19:301-314.

(83) Nickolson VJ, Tam SW, Myers MJ, Cook L. DUP 996 (3, 3-bix (4-pyrindinylmethyl) - 1-phenylindolin-2-one) enhances the stimulus-induced release of acetylcholine from rat brain in vitro and in vivo. Drug Dev Res 1990; 19:285-300.

(84) DeNoble VJ, DeNoble KF, Spencer KR et al. Comparison of DUP 996, with physostigmine, THA, and 3, 4-DAP on hypoxia-induced amnesia in rats. Pharmacology, Biochemistry & Behavior 1990; 36:957-961.

(85) Rockwood K, Beattie BL, Eastwood MR et al. A randomized, controlled trial of linopirdine in the treatment of Alzheimer‘s disease. Can J Neurol Sci 1997; 24(2):140-145.

(86) Kleijnen J, Knipschild P. Ginkgo biloba. Lancet 1992; 340:1136-1139.

(87) Le Bars PL, Katz MM, Berman N, Itil TM, Freedman AM, Schatzberg AF. A placebo-controlled, double-blind, randomized trial of an extract of ginkgo biloba for dementia. North american EGb study group. JAMA 1997; 278(16):1327-1332.

(88) Maurer K, Ihl R, Dierks T, Frolich L. Clinical efficacy of ginkgo biloba special extract EGb 761 in dementia of the Alzheimer type. J Psychiatr Res 1997; 31(6):645-655.

(89) Kanowski S, Herrmann WM, Stephan K et al. Proof of efficacy of the ginkgo biloba special extract EGb 761 in outpatients suffering from mild to moderate primary degenerative dementia of the Alzheimer type or multi-infarct dementia. Pharmacopsychiatry 1996; 29(2):47-56.

(90) Taillandier J, Ammar A, Rabourdin JP et al. Traitement des troubles du vieillissement cerebral par l'extrait de Ginkgo Biloba. Etude longitudinale multicentrique a double insu face au placebo. Presse Med 15[31], 1583-1587. 1986.

(91) Stephens MDB. Detection of New Adverse Drug Reactions. 3 ed. Reinhold, N.Y.: van Nostrand, 1993.

(92) Ibanez L, Laporte JR, Carne X. Adverse drug reactions leading to hospital admission. Drug Safety 1991; 6:450-459.

(93) Schmidt LG, Grohmann R, Helmchen H et al. Adverse drug reactions. Acta Psychiatr Scand 1984; 70:77-89.

(94) Bergamasco B, Scarzella L, La Commare P. Idebenone, A New Drug in the Treatment of Cognitive Impairment in Patients with Dementia of the Alzheimer Type. Functional Neurology 1994; 9(3):161-168.

(95) Little A. Evaluating treatments for dementia: Lessons from the maudsley lecithin trial. Int J Geriatr Psychiatry 1996; 11(4):293-304.

6 BIBLIOGRAPHY BY AUTHOR

Albert MS. Assessment of cognitive dysfunction. In: Albert MS, Moss MB, editors. Geriatric Neuropsychology. New York: The Guilford Press, 1988: 57-81.

American Psychiatric Association. Diagnostic and Statistical Manual for Mental Disorders. 3 ed. Washington, DC: American Psychiatric Association, 1987.

Anonymous. Canadian study of health and aging: Methods and prevalence of dementia. CMAJ 1994; 150(6):899-913.

Bartus RT, Dean LR, Beer B, Lippa AS. The cholinergic hypothesis of geriatric memory dysfunction. Science 1982; 217:408-417.

Becker RE, Colliver J, Elble R et al. Effects of metrifonate, a long-acting cholinesterase inhibitor, in Alzheimer‘s disease. Drug Dev Res 1990; 19:425-434.

Becker RE, Colliver JA, Markwell SJ et al. Double-blind, placebo-controlled study of metrifonate, an acetylcholinesterase inhibitor, for Alzheimer‘s disease. Alzheimer Dis Assoc Disord 1996; 10(3):124-131.

Becker RE, Colliver JA, Markwell SJ et al. Effects of metrifonate on cognitive decline in Alzheimer‘s disease: A double-blind, placebo-controlled, 6-month study. Alzheimer Dis Assoc Disord 1998; 12(1):54-57.

Berg L, Miller JP, Storandt M et al. Mild senile dementia of the Alzheimer type: 2. Longitudinal assessment. Ann Neurol 1988; 23(5):477-484.

Bergamasco B, Scarzella L, La Commare P. Idebenone, A New Drug in the Treatment of Cognitive Impairment in Patients with Dementia of the Alzheimer Type. Functional Neurology 1994; 9(3):161-168.

Blusztajn JK, Holbrook PG, Lakher M et al. 'Autocannibalism' of membrane choline-phospholipids: Physiology and pathology. Psychopharmacol Bull 1986; 22:781-786.

Burke WJ, Roccaforte WH, Wengel SP et al. L-deprenyl in the treatment of mild dementia of the Alzheimer type: Results of a 15-month trial. J Am Geriatr Soc 1993; 41(11):1219-1225.

Burns A, Rossor M, Hecker J et al. The effects of donepezil in Alzheimer's disease - Results from a multinational trial. Dement Geriatr Cogn Disord 1999; 10:237-244.

Campi N, Todeschini GP, Scarzella L. Selegiline versus l-acetylcarnitine in the treatment of Alzheimer-type dementia. Clin Ther 1990; 12(4):306-314.

Clark CM, Sheppard L, Fillenbaum GG, Galasko D, Morris JC, Koss E et al. Variability in annual Mini-Mental State Examination score in patients with probable Alzheimer disease: A clinical perspective of data from the Consortium to Establish a Registry for Azlheimer's Disease. Arch Neurol 1999; 56:857-862.

Clark HD, Wells GA, Huët C, McAlister FA, Salmi R, Fergusson D et al. Assessing the quality of randomized trials: Reliability of the Jadad scale. Control Clin Trials 1999; 20:448-452.

Cole MG. Interrater reliability of the Blessed Dementia Scale. Can J Psychiatry 1990; 35(4):328-330.

Cook L, Nickolson VJ, Steinfels GF et al. Cognition enhancement by the acetylcholine releaser DUP 996. Drug Dev Res 1990; 19:301-314.

Corey-Bloom J, Anand R, Veach J et al. 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 Psychiatry 1998; 1:55-65.

Cummings JL, Cyrus PA, Bieber F et al. Metrifonate treatment of the cognitive deficits of Alzheimer‘s disease. Metrifonate Study Group. Neurology 1998; 50(5):1214-1221.

Dahlke F, Lohaus A, Gutzmann H. Reliability and clinical concepts underlying global judgements in dementia: Implications for clinical research. Psychopharmacol Bull 1992; 28(4):425-432.

DeJong R, Osterlund OW, Roy GW. Measurement of quality-of-life changes in patients with Alzeimer’s disease. Clin Ther 1989; 11(4):545-554.

DeNoble VJ, DeNoble KF, Spencer KR et al. Comparison of DUP 996, with physostigmine, THA, and 3, 4-DAP on hypoxia-induced amnesia in rats. Pharmacology, Biochemistry & Behavior 1990; 36:957-961.

DerSimonian R, Charette LJ, McPeek B. Reporting on methods in clinical trials. New Engl J Med 1982; 306:1332-1337.

Drachman DA, Swearer JM, Kane K et al. The Cognitive Assessment Screening Test (CAST) for dementia. J Geriatr Psychiatry Neurol 1996; 9(4):200-208.

Eisdorfer C, Cohen D, Paveza GJ et al. An empirical evaluation of the Global Deterioration Scale for staging Alzheimer's disease. Am J Psychiatry 1992; 149(2):190-194.

Falsaperla A, Monici-Preti PA, Oliani C. Selegiline versus oxiracetam in patients with Alzheimer-type dementia. Clin Ther 1990; 12(4):376-384.

Fitten LJ, Perryman KM, Gross PL, Fine H et al. Treatment of Alzheimer's disease with short- and long-term oral THA and lecithin: A double-blind study. Am J Psychiatry 1990; 147(2):239-242.

Foster NL, Petersen RC, Gracon SI, Lewis K. An enriched-population, double-blind, placebo-controlled, crossover study of tacrine and lecithin in Alzheimer's disease. The tacrine 970-6 study group. Dementia 1996; 7(5):260-266.

Freedman M, Rewilak D, Cohen S et al. L-deprenyl in Alzheimer's disease: Cognitive and behavioral effects. Neurology 1998; 50(3):660-668.

Gauthier S, Bouchard R, Lamontagne A et al. Tetrahydroaminoacridine-lecithin combination treatment in patients with intermediate stage Alzheimer's disease: Results of a Canadian double-blind, crossover, multicenter study. N Engl J Med 1990; 322(18):1272-1276.

Goto M, Demura N, Sakaguchi T. Effects of HWA 285 on disorder of learning and memory in rodents. Jpn J Pharmacol 1987; 45:373-378.

Heun R, Papassotiropoulos A, Jennssen F. The validity of psychometric instruments for detection of dementia in the elderly general population. Int J Geriatr Psychiatry 1998; 13(6):363-380.

Hofferberth B. The Efficacy of EGb 761 in patients with senile dementia of the Alzheimer type: A double-blind placebo-controlled study on different levels of investigation. Human Psychopharmacol 1994; 9:215-222.

Hughes CP, Berg L, Danziger WL et al. A new clinical scale for the staging of dementia. Brit J Psychiat 1982; 140:566-572.

Ibanez L, Laporte JR, Carne X. Adverse drug reactions leading to hospital admission. Drug Safety 1991; 6:450-459.

Jadad AR, Moore A, Carroll D et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Control Clin Trials 17, 1-12. 1996.

Jann MW, Cyrus PA, Eisner LS, Margolin DI, Griffin T, Gulanski B. Efficacy and safety of a loading-dose regimen versus a no-loading dose regimen of metrifonate in the symptomatic treatment of Alzheimer's disease: a randomized, double-masked, placebo-controlled trial. Metrifonate Study Group. Clin Ther 1999; 21(1):88-102.

Kanowski S, Herrmann WM, Stephan K et al. Proof of efficacy of the ginkgo biloba special extract EGb 761 in outpatients suffering from mild to moderate primary degenerative dementia of the Alzheimer type or multi-infarct dementia. Pharmacopsychiatry 1996; 29(2):47-56.

Kaufer D. Beyond the cholinergic hypothesis: The effect of metrifonate and other cholinesterase inhibitors on neuropsychiatric symptoms in Alzheimer's disease. Dement & Geriatr Cog Disord 1998; 9(Suppl. 2):8-14.

Kleijnen J, Knipschild P. Ginkgo biloba. Lancet 1992; 340:1136-1139.

Knopmann D, Gracon S. Observations on the short-term 'natural history' of probable Alzheimer's disease in a controlled clinical trial. Neurology 1994; 44(2):260-265.

Kramer-Ginsberg E, Mohs RC, Aryan M et al. Clinical predictors of course of Alzheimer's patients in a longitudinal study: A preliminary report. Psychopharmacol Bull 1988; 24:458-462.

Lawlor BA, Aisen PS, Green C et al. Selegiline in the treatment of behavioral disturbance in Alzheimer's disease. Int J Geriatric Psychiatry 1997; 12:319-322.

Le Bars PL, Katz MM, Berman N, Itil TM, Freedman AM, Schatzberg AF. A placebo-controlled, double-blind, randomized trial of an extract of ginkgo biloba for dementia. North American EGb study group. JAMA 1997; 278(16):1327-1332.

Little A. Evaluating treatments for dementia: Lessons from the maudsley lecithin trial. Int J Geriatr Psychiatry 1996; 11(4):293-304.

Losascio JJ, Growdon JH, Corkin S. Cognitive test performance in detecting, staging, and tracking Alzheimer's disease. Arch Neurol 1995; 52:1087-1099.

Mack JL, Patterson MB. The evaluation of behavioral disturbances in Alzheimer's disease: The utility of three rating scales. J Geriatr Psychiatry Neurol 1994; 7:101-117.

Mangoni A, Grassi MP, Frattola L et al. Effects of a MAO-B inhibitor in the treatment of Alzheimer disease. Eur Neurol 1991; 31(2):100-107.

Marcusson J, Rother M, Kittner B et al. A 12-month, randomized, placebo-controlled trial of propentofylline (HWA 285) in patients with dementia according to DSM III-R. The European Propentofylline Study Group. Dement Geriatr Cogn Disord 1997; 8(5):320-328.

Maurer K, Ihl R, Dierks T, Frolich L. Clinical efficacy of ginkgo biloba special extract EGb 761 in dementia of the Alzheimer type. J Psychiatr Res 1997; 31(6):645-655.

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984; 34:939-944.

Mielke R, Ghaemi M, Kessler J et al. Propentofylline enhanced cerebral metabolic response to auditory memory stimulation in Alzheimer's disease. J Neurol Sci 1998; 154(1):76-82.

Moher D, Jadad AR, Nichol G et al. Assessing the quality of randomized controlled trials: An annotated bibliograhpy of scales and checklists. Control Clin Trials 16, 62-73. 1995.

Mohs RC, Cohen L. Alzheimer's Disease Assessment Scale (ADAS). Psychopharmacol Bull 1988; 24(4):627-628.

Mohs RC, Rosen WG, Davis KL. The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. Psychopharmacol Bull 1983; 19(3):448-450.

Monteverde A, Gnemmi P, Rossi F et al. Selegiline in the treatment of mild to moderate Alzheimer-type dementia. Clin Ther 1990; 12(4):315-322.

Morris JC, Cyrus PA, Orazem J et al. Metrifonate benefits cognitive, behavioral, and global function in patients with Alzheimer‘s disease. Neurology 1998; 50(5):1222-1230.

Nabeshima T, Nitta A, Hasegawa T. Impairment of learning and memory and the accessory symptoms in aged rats as senile dementia model: Oral administration of HWA 285 produces recovery of reduced NGF content in the brain of aged rats. Jpn J Psychopharmacol 1993; 13:89-95.

Nickolson VJ, Tam SW, Myers MJ, Cook L. DUP 996 (3, 3-bix (4-pyrindinylmethyl) - 1-phenylindolin-2-one) enhances the stimulus-induced release of acetylcholine from rat brain in vitro and in vivo. Drug Dev Res 1990; 19:285-300.

Nolan KA, Black RS, Sheu KF et al. A trial of thiamine in Alzheimer‘s disease. Arch Neurol 1991; 48(1):81-83.

Ownby RL, Koss E, Smyth KA, Whitehouse PJ. The factor structure of the Brief Psychiatric Rating Scale in Alzheimer’s disease. J Geriatr Psychiatry Neurol 1994; 7:245-250.

Parkinson FE, Rudolphi KA, Fredholm BB. HWA 285: A nucleoside transport inhibitor with neuroprotective effects in cerebral ischemia. Gen Pharmacol 1994; 6:1053-1058.

Pettigrew LC, Beiber F, Lettieri J et al. Pharmacokinetics, pharmacodynamics, and safety of metrifonate in patients with Alzheimer‘s disease. J Clin Pharmacol 1998; 38(3):236-245.

Rai GS, Shovlin C, Wesnes KA. A double-blind placebo-controlled study of ginkgo biloba extract ("Tanakan") in elderly outpatients with mild to moderate memory impairment. Can Med Res Opin 1991; 12:350-355.

Raskind MA, Cyrus PA, Ruzicka BB, Gulanski BI. The effects of metrifonate on the cognitive, behavioral, and functional performance of Alzheimer's disease patients. Metrifonate Study Group. J Clin Psychiatry 1999; 60(5):318-325.

Reisberg B, Ferris SH, de Leon MJ, Crook T. Global Deterioration Scale (GDS). Psychopharmacol Bull 1988; 24(4):661-663.

Rockwood K, Beattie BL, Eastwood MR et al. A randomized, controlled trial of linopirdine in the treatment of Alzheimer's disease. Can J Neurol Sci 1997; 24(2):140-145.

Rogers SL, Doody RS, Mohs RC. Donepezil improves cognition and global function in Alzheimer disease, a 15-week, double-blind, placebo-controlled study. Donepezil Study Group. Arch Intern Med 1998; 158(9):1021-1031.

Rogers SL, Farlow MR, Doody RS et al. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group. Neurology 1998; 50(1):136-145.

Rogers SL, Freidhoff LT. 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 1996; 7(6):293-303.

Rosen WG, Mohs RC, Davis K. A new rating scale for Alzheimer's disease. Am J Psychiatry 1984; 141:1356-1364.

Rosler M, Anand R, Cicin-Sain A et al. Efficacy and safety of rivastigmine in patients with Alzheimer‘s disease: International randomized controlled trial. BMJ 1999; 318(7184):633-640.

Rubin EH, Morris JC, Grant EA, Vendegna T. Very mild senile dementia of the Alzheimer type. Arch Neurol 1989; 46:379-382.

Sano M, Ernesto C, Thomas RG et al. A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer‘s disease. N Engl J Med 1997; 336(17):1216-1222.

Schmidt LG, Grohmann R, Helmchen H et al. Adverse drug reactions. Acta Psychiatr Scand 1984; 70:77-89.

Schneider LS, Anand R, Farlow MR. Systematic review of the efficacy of rivastigmine for patients with Alzheimer’s disease. Int J Geriatr Psychopharm 1998; 1 (Suppl. 1):S26-S34.

Schneider LS, Olin JT, Pawluczyk S. A double-blind crossover pilot study of l-deprenyl (selegiline) combined with cholinesterase inhibitor in Alzheimer‘s disease. Am J Psychiatry 1993; 150(2):321-323.

Sellwood E. Rivastigmine in Alzheimer’s disease. eBMJ . 29-3-1999. www.bmj.com/cgi/eletters/318/7184/633.

Shinoda I, Furukawa Y, Furukawa S. Stimulation of nerve growth factor synthesis/secretion by HWA 285 in cultured mouse astroglial cells. Biochem Pharmacol 1990; 39:1813-1816.

Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979; 86:420-428.

Slavin RE. Best evidence synthesis: An intelligent alternative to meta-analysis. J Clin Epidemiol 1995; 48(1):9-18.

Small BJ, Viitanen M, Winbald B, Baeckman L. Cognitive changes in very old persons with dementia: The influence of demographic, psychometric and biological variables. J Clin Exp Neuropsychol 1997; 19:245-260.

Sramek JJ, Anand R, Wardle TS et al. Safety/tolerability trial of SDZ ENA-713 in patients with probably Alzheimer‘s disease. Life Sci 1996; 58(15):1201-1207.

Stephens MDB. Detection of New Adverse Drug Reactions. 3 ed. Reinhold, N.Y.: van Nostrand, 1993.

Stern RG, Mohs RC, Davidson M et al. A longitudinal study of Alzheimer‘s disease: Measurement, rate and predictors of cognitive deterioration. Am J Psychiatry 1994; 151:390-396.

Taillandier J, Ammar A, Rabourdin JP et al. Traitement des troubles du vieillissement cerebral par l’extrait de Ginkgo Biloba. Etude longitudinale multicentrique a double insu face au placebo. Presse Med 15[31], 1583-1587. 1986.

Tariot PN, Goldstein B, Podgorski CA et al. Short-term administration of selegiline for mild-to-moderate dementia of the Alzheimer‘s type. Am J Geriatr Psychiatry 1998; 6(2):145-154.

Tariot PN, Mack JL, Patterson MB et al. The Behavior Rating Scale for Dementia of the Consortium to Establish a Registry for Alzheimer‘s Disease: The Behavioral Pathology Committee of the Consortium to Establish a Registry for Alzheimer‘s Disease. Am J Psychiatry 1995; 152:1349-1359.

Tariot PN. CERAD Behavior Rating Scale for Dementia. Int Psychogeriatr 1996; 8 (Suppl. 3):514-518.

Teng EL, Chui HC. The Modified Mini-Mental State (3MS) Examination. J Clin Psychiatry 1987; 48:314-318.

Ulus IH, Wurtman RJ. Prevention by choline of the depletion of membrane phosphatidylcholine by a cholinesterase inhibitor. N Engl J Med 1988; 318:191.

Weinstein HC, Teunisse S, van Gool WA. Tetrahydroaminoacridine and lecithin in the treatment of Alzheimer‘s disease. Effect on cognition, functioning in daily life, behavioral disturbances and burden experienced by the carers. J Neurol 1991; 238(1):34-38.

Wesnes K, Simmons D, Rock M, Simpson P. A double-blind placebo-controlled trial of tanakan in the treatment of idiopathic cognitive impairment in the elderly. Human Psychopharmacol 1987; 2:159-169.

Wolfson C, Moride Y, Perrault A, Vida S. A Study of the Efficacy, Effectiveness and Economic Impact of Tacrine in Alzheimer’s Disease. 1997. Ottawa, Canadian Coordinating Office for Health Technology Assessment.

Yamada Y, Nakamura H, Okada Y. HWA 285 enhances the formation of long-term portentiation in guinea pig hippocampal slices. Neuroscience Letters 1994; 176:189-192.

7 APPENDIX A

Table 1: The 3-Item Jadad Scale

Question	**Response Option**
1. Was the study described as randomized?	'Yes' or 'No'
2. Was the study described as double-blind?	'Yes' or 'No'
3. Was there a description of withdrawals and drop-outs?	'Yes' or 'No'
Scoring	
For each question, award one point for an affirmative response or zero points for a negative response. For question 1: ➢ award a bonus point if the method of randomization is appropriate (e.g., computer generated, etc.); ➢ deduct one point if the method of randomization is inappropriate. For question 2: ➢ award a bonus point if the method of double-blinding is appropriate (e.g., identical placebo, etc.); ➢ deduct one point if the method of double-blinding is inappropriate.	

Source: Jadad AR, Moore A, Carroll D et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Control Clin Trials 1996;17:1-12.

Table 2: The Additional Three Questions for the 6-Item Jadad Scale

Question	**Response Option**
Was there a clear description of the inclusion/exclusion criteria?	'Yes' or 'No'
Was the method used to assess adverse effects described?	'Yes' or 'No'
Were the methods of statistical analysis described?	'Yes' or 'No'
Scoring (for each question)	
One point is awarded for an affirmative response.	
A score of zero is awarded for a negative response.	

Source: Jadad AR, Moore A, Carroll D et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Control Clin Trials 1996;17:1-12.

APPENDIX B

RCTs Excluded from Review and Reasons for Exclusion

Medication	RCT	Reason for Exclusion
Estrogen	Schneider et al., 1997	Jadad < 5
	Phillips et al., 1992	Jadad < 5
	Ditkof et al., 1991	Jadad < 5
	Sherwin et al., 1988	Jadad < 5
Ginkgo Biloba	Hofferberth et al., 1994	NINCDS/ADRDA not satisfied
	Rai et al., 1991	NINCDS/ADRDA not satisfied
	Wesnes et al., 1987	NINCDS/ADRDA not satisfied
	Taillandier et al., 1986	Subjects not demented
Indomethacin	Rogers et al., 1993	Jadad < 5
Lecithin	Safford et al., 1994	Jadad < 5
	Crapper-McLachlan et al., 1993	Jadad < 5
	Minthon et al., 1993	Jadad < 5
	Uney et al., 1992	Jadad < 5
	Weinstein et al., 1991	*
	Chatellier et al., 1990	Jadad < 5
	Fitten et al., 1990	*
	Gauthier et al., 1990	*
	Vida et al., 1989	Jadad < 5
	Heyman et al., 1987	Jadad < 5
	Sorgatz et al., 1987	Jadad < 5
		*Results insufficient to comment on lecithin's effect
Linopirdine	van Dyck et al., 1997	Jadad < 5
Metrifonate	Kaufer et al., 1998	Jadad < 5

RCTs Excluded from Review and Reasons for Exclusion, con't.

Medication	RCT	Reason for Exclusion
Propentofylline	Mielke et al., 1996	Jadad < 5
	Moller et al., 1994	Jadad < 5
Selegiline	Filip et al., 1999	Jadad < 5
	Oakley et al., 1997	Jadad < 5
	Marin et al., 1995	Jadad < 5
	Agnoli et al., 1992	Jadad < 5
	Finali et al., 1992	Jadad < 5
	Finali et al., 1991	Jadad < 5
	Martignoni et al., 1991	Jadad < 5
	Agnoli et al., 1990	Jadad < 5
	Campi et al., 1990	Single-blind
	Falsapera et al., 1990	Single-blind
	Monteverde et al., 1990	Single-blind
	Piccinin et al., 1990	Jadad < 5
	Tariot et al., 1987	Jadad < 5
	Tariot et al., 1987	Jadad < 5

BIBLIOGRAPHY OF EXCLUDED RCTs

Agnoli A, Fabbrini G, Fioravanti M, Martucci N. CBF and cognitive evaluation of Alzheimer type patients before and after IMAO-B treatment: A pilot study. Eur Neuropsychopharmacol 1992; 2(1):31-35.

Agnoli A, Martucci N, Fabbrini G et al. Monoamine oxidase in dementia: Treatment with an inhibitor of MAO-B activity. Dementia 1990; 1:109-114.

Campi N, Todeschini GP, Scarzella L. Selegiline versus l-acetylcarnitine in the treatment of Alzheimer-type dementia. Clin Ther 1990; 12(4):306-314.

Chatellier G, Lacomblez L. Tacrine (tetrahydroaminoacridine; THA) and lecithin in senile dementia of the Alzheimer type: A multicentre trial. Groupe Francais d'Etude de la Tetrahydroaminoacridine. BMJ 1990; 300(6723):495-499.

Crapper McLachlan DR, Smith WL, Kruck TP. Desferrioxamine and Alzheimer's disease: Video home behavior assessment of clinical course and measures of brain aluminum. Ther Drug Monit 1993; 15(6):602-607.

Ditkoff EC, Crary WG, Cristo M, Lobo RA. Estrogen improves psychological function in asymptomatic postmenopausal women. Obst Gynecol 1991; 78:991-995.

Falsaperla A, Monici-Preti PA, Oliani C. Selegiline versus oxiracetam in patients with Alzheimer-type dementia. Clin Ther 1990; 12(4):376-384.

Filip V, Kolibas E. Selegiline in the treatment of Alzheimer's disease: A long-term randomized placedo-controlled trial. Czech and Slovak Senile Dementia of Alzheimer Type Study Group. J Psychiatry Neurosci 1999; 24(3):234-243.

Finali G, Piccirilli M, Oliani C, Piccinin GL. Alzheimer-type dementia and verbal memory performances: Influence of selegiline therapy. Ital J Neurol Sci 1992; 13(2):141-148.

Finali G, Piccirilli M, Oliani C, Piccinin GL. L-deprenyl therapy improves verbal memory in amnesic Alzheimer patients. Clin Neuropharmacol 1991; 14(6):523-536.

Fitten LJ, Perryman KM, Gross PL et al. Treatment of Alzheimer‘s disease with short- and long-term oral THA and lecithin: A double-blind study. Am J Psychiatry 1990; 147(2):239-242.

Gauthier S, Bouchard R, Lamontagne A et al. Tetrahydroaminoacridine-lecithin combination treatment in patients with intermediate stage Alzheimer‘s disease: Results of a Canadian double-blind, crossover, multicenter study. N Engl J Med 1990; 322(18):1272-1276.

Heyman A, Schmechel D, Wilkinson W et al. Failure of long term high-dose lecithin to retard progression of early-onset Alzheimer‘s disease. J Neural Transm 1987; Supplementum 24:279-286.

Hofferberth B. The Efficacy of EGb 761 in patients with senile dementia of the Alzheimer type: A double-blind placebo-controlled study on different levels of investigation. Human Psychopharmacol 1994; 9:215-222.

Kaufer D. Beyond the cholinergic hypothesis: The effect of metrifonate and other cholinesterase inhibitors on neuropsychiatric symptoms in Alzheimer‘s disease. Dement & Geriatr Cog Disord 1998; 9(Suppl. 2):8-14.

Marin DB, Bierer LM, Lawlor BA et al. L-deprenyl and physostigmine for the treatment of Alzheimer‘s disease. Psychiatry Res 1995; 58(3):181-189.

Martignoni E, Bono G, Blandini F et al. Monoamines and related metabolite levels in the cerebrospinal fluid of patients with dementia of Alzheimer type. Influence of Treatment with L-deprenyl. J Neural Transm Park Dis Dement Sect 1991; 3(1):15-25.

Mielke R, Kittner B, Ghaemi M. Propentofylline improves regional cerebral glucose metabolism and neuropsychologic performance in vascular dementia. J Neurol Sci 1996; 141(1-2):59-64.

Minthon L, Gustafson L, Dalfelt G et al. Oral Tetrahydroaminoacridine Treatment of Alzheimer‘s Disease Evaluated Clinically and by Regional Cerebral Blood Flow and EEG. Dementia 1993; 4(1):32-42.

Moller HJ, Maurer I, Saletu B. Placebo-controlled trial of the xanthine derivative propentofylline in dementia. Pharmacopsychiatry 1994; 27(4):159-165.

Monteverde A, Gnemmi P, Rossi F et al. Selegiline in the treatment of mild to moderate Alzheimer-type dementia. Clin Ther 1990; 12(4):315-322.

Oakley F, Sunderland T. Assessment of motor and process skills as a measure of IADL functioning in pharmacologic studies of people with Alzheimer‘s disease: A pilot study. Int Psychogeriatr 1997; 9(2):197-206.

Phillips SM, Sherwin BB. Effects of estrogen on memory function in surgically menopausal women. Psychoneuroendocrinology 1992; 17:485-495.

Piccinin GL, Finali G, Piccirilli M. Neuropsychological effects of L-deprenyl in Alzheimer‘s type dementia. Clin Neuropharmacol 1990; 13(2):147-163.

Rai GS, Shovlin C, Wesnes KA. A double-blind placebo-controlled study of ginkgo biloba extract ("Tanakan") in elderly outpatients with mild to moderate memory impairment. Can Med Res Opin 1991; 12:350-355.

Rogers J, Kirby LC, Hempelman SR et al. Clinical trial of indomethacin in Alzheimer‘s disease. Neurology 1993; 43(8):1609-1611.

Safford F, Baumel B. Testing the effects of dietary lecithin on memory in the elderly: An example of social work/medical research collaboration. Research on Social Work Practice 1994; 4(3):349-358.

Schneider LS, Farlow MR, Pogoda JM. Potential role for estrogen replacement in the treatment of Alzheimer‘s dementia. Am J Med 1997; 103(2S):46S-50S.

Sherwin BB. Estrogen and/or androgen replacement therapy and cognitive functioning in surgically menopausal women. Psychoneuroendocrinology 1988; 13:345-357.

Sorgatz H. Effects of lecithin on memory and learning. In: Hann I, Ansell GB, editors. Lecithin: Technological, Biological and Therapeutic Aspects. New York: Plenum Press, 1987: 147-153.

Taillandier J, Ammar A, Rabourdin JP et al. Traitement des troubles du vieillissement cerebral par l’extrait de Ginkgo Biloba. Etude longitudinale multicentrique a double insu face au placebo. Presse Med 1986; 15(31):1583-1587.

Tariot PN, Cohen RM, Sunderland T. L-deprenyl in Alzheimer‘s disease. Arch Gen Psychiatry 1987; 44:427-433.

Tariot PN, Sunderland T, Weingartner H et al. Cognitive effects of L-deprenyl in Alzheimer‘s disease. Psychopharmacology 1987; 91:489-495.

Uney JB, Jones GM, Rebeiro A, Levy R. The effect of long-term high dose lecithin on erythrocyte choline transport in Alzheimer patients. Biol Psychiatry 1992; 31(6):630-633.

van Dyck CH, Lin CH, Robinson R et al. The acetylcholine releaser linopirdine increases parietal regional cerebral blood flow in Alzheimer‘s disease. Psychopharmacology 1997; 132(3):217-226.

Vida S, Gauthier L, Gauthier S. Canadian Collaborative Study of Tetrahydroaminoacridine (THA) and Lecithin Treatment of Alzheimer‘s Disease: Effect on mood. Can J Psychiatry 1989; 34(3):165-170.

Weinstein HC, Teunisse S, van Gool WA. Tetrahydroaminoacridine and lecithin in the treatment of Alzheimer‘s disease. Effect on cognition, functioning in daily life, behavioral disturbances and burden experienced by the carers. J Neurol 1991; 238(1):34-38.

Wesnes K, Simmons D, Rock M, Simpson P. A double-blind placebo-controlled trial of tanakan in the treatment of idiopathic cognitive impairment in the elderly. Human Psychopharmacol 1987; 2:159-169.

APPENDIX C

Medications Excluded from Review (con't next page)

Acetyl-l-carnitine	LU 25-109
AF-102B	m-CPP
Amiridin	Melatonin
Aniracetam	Memantine
Arecoline	Milacemide
Besipirdine	Moclobemide
BMY 21,502	Naftidrofuryl
Carbamazepine	Nicergoline
Ceranapril	Nicotine
Cerebrolysin	Nimopidine
CI-979	Oral 5' – Meth.
Citalopram	Org 2766
Clonidine	Org 5667
Colchicine	Oxiracetam
Corticotropin 4-9	Partoid Salivary FRM
D-cycloserine	Phosphatidylserine
Denbufylline	Physostigmine
Desferrioxamine	Piracetam
DHEA	Posatirelin
Dihydroergotoxine	Pramiracetam
Dronabinol	Pyritinol
E2020	Risperidene
Eptastigmine	Sabeluzole
Essential fatty acids	Somatostatin
Ganglioside	SR-3
GM-1	Thiamine
Guanfacine	Thiothixene
Haloperidol	Thyrotropin
HP-128	Tranylcypromine

Huperzine A	Triazolam
Idebenone	Velnacrine
Inositol	Vincamine
Lazabemide	Vinpocetine
Lisuride	Xanomeline
Loxapine	Xantinolnicotinate
L-tryptophan	Yohimbine