Published in final edited form as:
Cogn Behav Neurol. 2008 December ; 21(4): 197–201. doi:10.1097/WNN.0b013e318190cf75.

Clinical Trial Design Issues in Mild to Moderate Alzheimer Disease

David S. Knopman, MD
Department of Neurology, Mayo Clinic, Rochester, MN.

Abstract

The field of clinical trials and therapeutics in Alzheimer Disease (AD) is little more than 20 years old. Considerable progress has been made in crafting appropriate designs for clinical trials of promising therapeutic agents for AD. This article reviews basic issues in diagnostic criteria, choice of outcome measures, duration of trials and analytic strategies. Through trial and error, a general set of strategies has evolved for the assessment of putative therapies for mild to moderate AD. The experience of the past two decades has set the stage for discovering the next generation of anti-AD drugs and introducing those therapies at milder stages of the disease.

Keywords

Alzheimer's disease; clinical trials; cholinesterase inhibitors

Introduction

Therapeutics in Alzheimer disease (AD) has entered only its third decade. Prior to 1986, the methodology for conducting clinical trials in AD was virtually non-existent. As a consequence of investigations with the cholinesterase inhibitor tacrine, a conceptual framework for AD clinical trials quickly gelled. Guidelines were established by the US FDA in 1990 that defined the general form of outcome measures[1]. The guidelines stipulated that a successful anti-AD drug would have to show benefits on both a cognitive test that reflected the core deficits in AD and a clinician's global impression. What was remarkable about the guidelines was that they did not specify an effect size, a minimum or maximum trial duration nor any kind of risk:benefit algorithm. Over the 1990's, an informal consensus appeared that a trial of 6 months duration with about 100-120 subjects per treatment arm was sufficient. Currently, there are 4 drugs that are approved and marketed in the USA for the treatment of mild to moderate AD. (A fifth drug, tacrine, was also approved but is not marketed due to its hepato- and gastro-intestinal toxicity.) Those 5 drugs were approved based on the 1990 FDA guidelines. Several other compounds also underwent phase III trials but failed to show efficacy.

While there is a growing recognition in the field that treatment at an earlier stage might be more effective, it is patients with mild to moderate disease that, by far, form the majority who present to physicians for treatment of dementia symptoms. Giving them the best possible therapy must still be a major goal for the field. The purpose of this review is to highlight some of the key neurocognitive issues in the design of clinical trials in mild to moderate AD.

Address correspondence and reprint requests to D.S. Knopman, MD, Department of Neurology, Mayo Clinic, 200 First Street SW, Rochester, Minnesota 55905 (telephone: 507-538-1038; fax: 507-538-6012; E-mail: knopman@mayo.edu)..

Disclosures: DSK has served on a Data Safety Monitoring Board for Sanofi-Aventis Pharmaceuticals, and is an investigator in a clinical trial sponsored by Elan Pharmaceuticals.

Diagnostic criteria

At the time that multicenter clinical trials were first initiated in AD, the NINCDS-ADRDA criteria[2] — had been accepted by the AD researchers world-wide. The AD criteria are usually used in tandem with the modified Hachinski Ischemic Index to reduce the probability of including patients with significant cerebrovascular disease[3,4]. Mild to moderate severity of dementia was defined by a range of scores on the Mini-Mental State Examination (MMSE) [5]. While this approach is not beyond criticism, experience suggests very broad acceptance in the context of clinical trials.

Trial design and duration

The randomized, double-blind, placebo-controlled, parallel-group design is the approach of choice for anti-AD drugs for mild to moderate AD. Alternative approaches such as cross-over designs are inadequate. The inevitable decline that occurs in AD violates a key assumption of the cross-over design, namely that patients are identical at the beginning of both phases. Studies that utilize a time-to-reach-endpoint (survival) design are feasible, and have been carried out, eg with vitamin E[6] and donepezil[7,8], but the majority of trials in mild to moderate AD have used endof-trial evaluations as the primary outcome measures. The instruments for assessing cognition, function and behavior are the focus of this review.

The initial multicenter trial of a cholinesterase inhibitor was only 6 weeks long[9]. In retrospect, it seems quite naïve to think that a drug could change cognition in AD in such a short time frame. Because it became clear that in-study placebo effects could take 6 weeks or longer to dissipate[10,11], longer trials were needed. For no empiric reason, the pivotal trials for all of the cholinesterase inhibitors were 5-6 months in duration[10-18]. Subsequently, 12 to 18 month duration trials have become the preferred approach for mild to moderate AD, for example [19-21].

Cognitive outcome measures

In keeping with the FDA requirement that one of the primary outcome measures be able to "detect changes in the core cognitive manifestations of dementia,"[1], the cognitive instrument that has been used in all of the cholinesterase inhibitor trials and many subsequent trials was the Alzheimer's disease assessment scale-cognitive (ADAScog)[22]. Developed specifically for clinical trials in AD, the ADAScog yields a single score that reflects the arbitrary weighting of performance in several cognitive domains. In keeping with the focus on AD, the ADAScog gives much greater weight to performance on learning and memory than to other domains. It also covers the domains of language and spatial cognition, but does not address executive function. Because of the lack of coverage of executive functions, additional items have been added to the ADAScog such as number cancellation and mazes[23].

The standard approach to analyzing ADAScog data in current clinical trials of mild to moderate AD is to use some form of linear regression such as analysis of variance or analysis of covariance in an intention-to-treat approach. Patients are included in analyses if they complete the trial or have at least one post-randomization ADAScog assessment. Covariates such as age, sex or other variables prospectively identified as relevant such as baseline MMSE scores may be included.

The properties of the ADAScog are well understood. There is a curvilinear relationship between disease severity and rate of change on the ADAScog. In mild patients (eg patients with MMSE scores in the 20-26 range at baseline), the rate of change is less than in the moderate range [24]. Patients with MMSE scores in the range of 14-21 generally show the greatest amount of change. With advancing severity of dementia such as in patients with MMSE scores less than

14, the rate of change in the ADAScog then tends to decrease with because performance approaches “floor” levels on several items.

The variability in the ADAScog’s change score has generally been equivalent to, or somewhat larger than, the amount of annual decline. For example, in an Alzheimer’s Disease Cooperative Study trial of 2 nonsteroidal antiinflammatory drugs in mild to moderate AD (baseline MMSE scores between 13 and 26)[20], the placebo group declined 5.7 ± 8.2 points over 1 year. In an industry trial of rofecoxib, the annual decline of a nearly identical group of mild to moderate AD patients (baseline MMSE 14-26)[25] was 5.44 (standard deviation not given in publication). Slightly smaller changes were seen in the placebo group of a B vitamin trial [26]. In contrast, in the AN-1792 immunization study in which only mild AD patients (MMSE 21-26) participated, the mean 1 year **decline** in the placebo group on the ADAScog was 0.6 ± 4.3 points. The very small annual change on the ADAScog in mild patients raises questions about the utility of the ADAScog at the mild end of the AD spectrum. Indeed, more sensitive cognitive test batteries have been proposed[27], and one was used in the clinical trial of AN-1792[28]. It consists of 9 widely used neuropsychological tests including subtests of the Wechsler Memory Scale, the Rey Auditory Verbal Learning Test, the Controlled Oral Word Association test, and Category fluency[27]. The neuropsychological battery was evaluated by creating a z-score composite of all of the elements of the battery. That z-score composite declined 0.21 ± 0.42 points which represented a much more favorable change to variance ratio than the ADAScog[27]. Because statistical power to detect treatment effects is a function of the size of the effect and the inverse of the between-subjects variability, fewer subjects would be needed to insure adequate power to detect a drug effect on this neuropsychological battery than the ADAScog.

The ratio of annual rate of change on the ADAScog and its variance are among the key elements that go into the possible sample size X study duration configurations of a clinical trial. The initial severity level of patients also influences the amount of annual change to be expected on the ADAScog in the control group. Longer duration trials will result in greater decline in the placebo group, with about the same level of variability between subjects. Hence sample sizes could be smaller in longer duration trials. However, attrition will be greater with longer studies, particularly if there are more moderately affected patients are enrolled. Thus, while effect sizes in longer trials imply the need for fewer subjects, greater attrition will require enrolling more subjects.

The ADAScog has proved to be a valuable tool because the 5 approved drugs exhibited treatment effects with it. The effect size in the cholinesterase inhibitor trials has been about 3-4 points on the ADAScog over 6 months[10-13] compared to the yearly decline in the ADAScog of about 6 points, eg[20,25] thus making the effect size equivalent to a delay in decline of about 6-8 months. Until there is more experience with alternatives, the ADAScog will probably remain as the cognitive instrument of choice for AD clinical trials for the immediate future. One persistent question about the ADAScog has been in defining what amount of therapeutic difference is clinically meaningful.

The ADAScog as a clinically relevant outcome in AD trials is not without its detractors; some critics of currently approved AD drugs would claim that the ADAScog is too sensitive to change in that the “statistically significant” changes observed in the trials of the approved drugs have not been clinically relevant. Because the cholinesterase inhibitors and memantine are palliative therapies whose impact is clearly modest in the majority of patients, the critics are really directing their skepticism at the drugs, not the ADAScog.

Clinician's Global Impression Ratings

The other primary outcome measure specifically mentioned in the FDA guidelines[1] was a clinician's global assessment. The form of a global assessment was not specifically defined, but rather the guideline required a tool that was meant to identify clinically meaningful change ("clinical utility")[1] over the course of the clinical trial, from the perspective of a skilled clinician that was separate from cognitive tests. Clinical meaningfulness was not defined by the FDA draft guideline, but the assumption of the regulatory agency was that clinicians "would know it when they see it." The original FDA guideline envisioned the use of a 7 point scale for the clinician's global rating, with a rating of "1" indicating marked improvement, "4" indicating no change and "7" indicating severe decline. The judgment as to whether the patient was worse, unchanged or better would be made by a clinician who was blinded to the scores from the cognitive tests of the primary outcome measure, but the judgment would, of necessity, include information from the clinician's own mental status examinations. Furthermore, the clinician involved in the global assessment was to be kept unaware, to the extent possible, of any adverse events that might unblind the treatment assignment. Initially, the FDA wished that only a patient examination be used for the determination, but ongoing discussions and dialogue made it clear that experts in the field felt that input from a primary caregiver was essential. A clinician's global impression then became a complex judgment involving evaluating evidence from two sources — the patient and the informant — and then assigning a rating to the overall impression of change[29]. Global change ratings have been routinely evaluated by intention to treat procedures and non-parametric statistics such as Cochran Mantel Haenzel chi-square test.

Because the FDA did not otherwise specify the structure and procedure for carrying out a valid assessment of change, industry and academic trialists experimented at first with several different versions. Although there may not be universal agreement on one version, the procedures and interview structure laid out by the Alzheimer's Disease Cooperative Study Clinical Global Impression of Change (CGIC) represent a validated and straightforward structured approach to the clinician's global that preserves the original intent of the FDA guidelines[30]. It utilizes information from both the patient and the caregiver.

Another approach to global ratings has been to use the Clinical Dementia Rating (CDR) Scale [31]. Its administration is somewhat different than a typical clinical global impression because the CDR is semi-structured and includes 6 individual domains of cognition and function. A severity score of 0 (normal) to 3 (severe impairment) is given for each domain. In addition, the CDR is scored both at baseline and follow-up, whereas the clinician's global impression is completed only at follow-up. Usually, the scores of the 6 domains are added together to yield a "sum of boxes" overall score. However, both the CDR and the clinician's global impression can be used to fulfill the spirit of the FDA requirement for the opinion of a skilled clinician who has weighed all relevant information from both informants and patients themselves. There is some empiric work on the properties of clinician's global impressions[29,30]. Generally, these instruments are reliable, but they have greater test-retest variability than cognitive tests. In contrast to the ADAScog, global impressions are insensitive to small amounts of change.

Clinician's global impressions have also been successful in trials of cholinesterase inhibitors and memantine. Thus, despite the potential drawbacks and inefficiencies of clinician's global impressions, some anti-AD drugs have been capable of generating a drug-placebo difference on the measure.

One of the criticisms of the clinician's global assessment is that it is too insensitive. At the individual patient level, while the ADAScog score can vary from 0-70, and the CDR sum of boxes can vary from 0 to 18 in 0.5 increments, the CGIC can move only 0-3 points better or

0-3 points worse. At the level of analysis by group, however, insensitivity is reflected in greater variability, which reduces power. In the placebo groups of the pivotal 24 week donepezil trial [11], the clinician's global impression declined 0.51 ± 1.0 rating points and the CDR sum of boxes declined 0.58 ± 1.73. Similar results in the placebo treated patients were seen at 6 months in a B vitamin trial, and at one year the CDR sum of boxes declined 1.60 ± 2.12 points[26]. The insensitivity may arise both from its moderate reliability as well as the coarseness of the rating scales. However, the feature of insensitivity is precisely what the FDA guideline intended for the instrument in order to detect only treatment effects that were visible to clinicians.

Functional assessments

Although the FDA guideline did not require assessments of daily functioning, most clinical trials have included some instrument that assessed caregiver opinions about the patients' daily functioning. Several instruments[32-34], as well as others, have been devised for this purpose, but no one instrument has emerged as the dominant one. There is of course much overlap between various instruments, but the lack of uniformity across trials has made it difficult to draw any conclusions about the magnitude of change demonstrable by functional assessments. In addition, the amount of change on a functional instrument that constitutes a clinically meaningful minimum is not established. In contrast to the ADAScog that often shows improvements in the placebo group in the first 6 weeks of study participation, functional assessments generally do not show improvement. An example of this phenomenon was illustrated in a trial of rivastigmine[13]. It appears that patients are not "allowed" to resume activities that they had previously given up.

Among the many concerns about measuring daily functioning is the wide variation in what prospective clinical trial participants ordinarily do. There are vast gender differences, differences depending upon whether a spouse is present or not, and vast differences in pre-morbid interests and abilities. Moreover, the informant who actually provides the information about daily functioning is another source of heterogeneity. Not only might perceptions of daily functioning on the part of the patient be colored by the informant's mood, abilities and expectations, large differences might be expected between how adult children and spouses perceive a patient's functioning.

Whether daily functioning is measured with a specific functional assessment tool or whether it is evaluated in the context of a global impression, it is highly desirable for a drug that shows benefits in cognitive assessments to show drug benefits on function. While it has not been possible to define a magnitude of benefit on functional measures that is clinically meaningful, a statistically significant result on a functional measure in a trial powered for an ADAScog drug effect would be encouraging.

One approach to the problem of measuring such a diverse construct like daily functioning has been to attempt to define an amount of clinically meaningful change prospectively for each patient[35], and then use that definition to define an end-point in a "survival" analysis. Such a strategy was successfully used with donepezil[7] and galantamine[35]. While this is a particularly clever and efficient approach, others have felt that the methodology for defining individual treatment failure is not sufficiently reproducible to allow it to be used as a primary outcome measure in clinical trials.

Neuropsychiatric Assessments

Neuropsychiatric symptoms are common in mild to moderate AD, but in the clinical trial context, behavioral outcomes have generally served as secondary outcomes, with some exceptions. The Neuro-Psychiatric Inventory (NPI)[36] has become the standard instrument used in AD clinical trials. There are two key issues that distinguish neuropsychiatric symptoms

from cognitive deficits in mild to moderate AD. First, in the mild to moderate AD patients who enroll in clinical trials, the burden of prevalent neuropsychiatric symptoms (ie, symptoms recorded at the time of enrollment) is modest. Second, the rate of treatment-emergent neuropsychiatric symptoms is also modest. To take a few examples, a clinical trial of nonsteroidal antiinflammatory agents in mild to moderate AD (MMSE 20.8 ± 3.6) noted that the baseline NPI score was 8.7 ± 10.6 and a one year decline of 3.4 ± 11.9 points[20], while a trial of memantine in mild to moderate AD (MMSE 17.2 ± 3.4) found that the placebo group had a baseline score of 12.2 ± 13.0 and a 24 week decline of 0.9 ± 16.33 points[37]. The B vitamin study group had a baseline NPI score of 6 with an interquartile range of 0 to 12[26]. In contrast, in a clinical trial specifically designed to test the antipsychotic agents in AD (MMSE score of 14.7 ± 5.8), the baseline NPI score in the placebo group was 39.1 ± 17.8 points[38]. Thus, at least in the clinical trials where AD patients were recruited for suitability to examine cognitive outcomes, the level of neuropsychiatric symptoms was far less than seen in patients recruited specifically because they had neuropsychiatric problems. These examples illustrate two additional points about neuropsychiatric symptoms in AD patients. There is a very substantial variability in the burden of neuropsychiatric symptoms. In addition, NPI scores rise with increasing baseline dementia severity.

Disease Modification versus symptomatic therapy

All stakeholders — patients, family members, physicians, insurors — in the field of AD therapeutics wish to have more potent medications, and it seems logical that treatments that get at the basic mechanisms of the disease would be more effective than those treatments that ameliorate a neurotransmitter deficit, for example, but don't alter the basic biological tempo of the disease. Disease modifying therapies are therefore the goal for AD therapeutics, where disease modification refers to slowing or aborting processes that are in the direct pathogenic pathway of the disease.

Proving that a therapy is disease-modifying turns out to be very difficult. A number of articles in the journal, *Alzheimer's & Dementia* in July 2006 extensively discussed the issues. It is difficult to summarize briefly, but there are analytic strategies using data from cognitive assessment that can, in principle, be used to claim that a treatment is disease-modifying. The goal when using cognitive test data to support the idea of disease modification is to show a divergence in the rate of decline in the slope of the cognitive test score curves between study drug and control group. Unfortunately slope analysis is analytically challenging. For example, as previously mentioned, many trials have observed that both placebo- and drug-treated show improvements on the ADAScog after 6 and sometimes 12 weeks of therapy. Thereafter, scores decline. Such a nonlinear pattern of performance is not well suited for analyses such as slope that assume linearity. Certain designs have also been proposed to test for disease modification. The randomized start and a randomized withdrawal approaches were formulated to demonstrate the divergence of cognition and global ratings[39] over the course of the trial, but there are many vexing conceptual issues with both of those designs, that discourage their use [40]. Just like cross-over designs, both of these approaches are compromised by the fact that the deficits in AD are not static, but instead worsen over time.

The idea that biomarkers can also support a disease-modifying claim is viewed as having great potential. Support for disease modification by neuroimaging evidence has been sought for several years, but so far the trials in which imaging was used[41,42] failed to find drug benefits. Although there are several candidate biomarkers such as structural imaging with MR, functional imaging with positron emission tomography and cerebrospinal fluid markers, none have had the good fortune to be included in a clinical trial that showed clear clinical improvements with treatment. Thus, it remains unknown whether biomarkers are capable of showing changes in the direction of improvement that paralleled clinical improvement.

Summary

Therapeutic trial methodology for mild to moderate AD has become much more sophisticated in the past 20 years, but many challenges remain for improvement in the efficiency of getting compounds from an interesting phase II profile to demonstration of efficacy of a drug in a phase III trial. Trial design issues should not impede progress and should not act as a barrier to test new drugs. As the field moves towards treatment of yet milder patients including those with mild cognitive impairment, the lessons learned with mild to moderate AD will be invaluable.

Acknowledgments

Supported by grants U01 AG 06786 (Mayo Alzheimer's Disease Patient Registry) and P50 AG 16574 (Mayo Alzheimer's Disease Research Center) from the National Institute on Aging

References

1. Leber, P. Guidelines for the Clinical Evaluation of Antidementia Drugs, First Draft. 1990.
2. 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;34:939–44. [PubMed: 6610841]
3. Hachinski VC, Lassen NA, Marshall J. Multi-infarct dementia. A cause of mental deterioration in the elderly. Lancet 1974;2:207–10. [PubMed: 4135618]
4. Rosen WG, Terry RD, Fuld PA, Katzman R, Peck A. Pathological verification of ischemic score in differentiation of dementias. Ann Neurol 1980;7:486–8. [PubMed: 7396427]
5. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98. [PubMed: 1202204]
6. Sano M, Ernesto C, Thomas RG, Klauber MR, Schafer K, Grundman M, et al. A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. The Alzheimer's Disease Cooperative Study. N Engl J Med 1997;336:1216–22. [PubMed: 9110909]
7. Mohs RC, Doody RS, Morris JC, Ieni JR, Rogers SL, Perdomo CA, et al. A 1-year, placebo-controlled preservation of function survival study of donepezil in AD patients. Neurology 2001;57:481–8. [PubMed: 11502917]
8. Courtney C, Farrell D, Gray R, Hills R, Lynch L, Sellwood E, et al. Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double-blind trial. Lancet 2004;363:2105–15. [PubMed: 15220031]
9. Davis KL, Thal LJ, Gamzu ER, Davis CS, Woolson RF, Gracon SI, et al. A double-blind, placebo-controlled multicenter study of tacrine for Alzheimer's disease. The Tacrine Collaborative Study Group. N Engl J Med 1992;327:1253–9. [PubMed: 1406817]
10. Corey-Bloom J, Anand R, Veach J, the ENA 713 B352 Study Group. 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 Geriatric Psychopharmacol 1998;1:55–65.
11. Rogers SL, Farlow MR, Doody RS, Mohs R, Friedhoff LT. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Neurology 1998;50:136–45. [PubMed: 9443470]
12. Knapp MJ, Knopman DS, Solomon PR, Pendlebury WW, Davis CS, Gracon SI. A 30-week randomized controlled trial of high-dose tacrine in patients with Alzheimer's disease. The Tacrine Study Group. JAMA 1994;271:985–91. [PubMed: 8139083]
13. Rosler M, Anand R, Cicin-Sain A, Gauthier S, Agid Y, Dal-Bianco P, et al. Efficacy and safety of rivastigmine in patients with Alzheimer's disease: international randomised controlled trial. BMJ 1999;318:633–640. [PubMed: 10066203]

14. Burns A, Rossor M, Hecker J, Gauthier S, Petit H, H M, et al. The Effects of Donepezil in Alzheimer's Disease - Results from a Multinational Trial. Dement Geriatr Cogn Disord 1999;10:237–244. [PubMed: 10325453]
15. Seltzer B, Zolnouni P, Nunez M, Goldman R, Kumar D, Ieni J, et al. Efficacy of donepezil in early-stage Alzheimer disease: a randomized placebo-controlled trial. Arch Neurol 2004;61:1852–6. [PubMed: 15596605]
16. Wilcock GK, Lilienfeld S, Gaens E. Efficacy and safety of galantamine in patients with mild to moderate Alzheimer's disease: multicentre randomised controlled trial. BMJ 2000;321:1445–9. [PubMed: 11110737]
17. Raskind MA, Peskind ER, Wessel T, Yuan W. Galantamine in AD: A 6-month randomized, placebo-controlled trial with a 6-month extension. Neurology 2000;54:2261–2268. [PubMed: 10881250]
18. Tariot PN, Solomon PR, Morris JC, Kershaw P, Lilienfeld S, Ding C. A 5-month, randomized, placebo-controlled trial of galantamine in AD. Neurology 2000;54:2269–2276. [PubMed: 10881251]
19. Aisen PS, Davis KL, Berg JD, Schafer K, Campbell K, Thomas RG, et al. A randomized controlled trial of prednisone in Alzheimer's disease. Alzheimer's Disease Cooperative Study. Neurology 2000;54:588–93. [PubMed: 10680787]
20. Aisen PS, Schafer KA, Grundman M, Pfeiffer E, Sano M, Davis KL, et al. Effects of rofecoxib or naproxen vs placebo on Alzheimer disease progression: a randomized controlled trial. JAMA 2003;289:2819–26. [PubMed: 12783912]
21. Mulnard RA, Cotman CW, Kawas C, van Dyck CH, Sano M, Doody R, et al. Estrogen replacement therapy for treatment of mild to moderate Alzheimer disease: a randomized controlled trial. Alzheimer's Disease Cooperative Study. JAMA 2000;283:1007–15. [PubMed: 10697060]
22. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. Am J Psychiatry 1984;141:1356–64. [PubMed: 6496779]
23. Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11(Suppl 2):S13–21. [PubMed: 9236948]
24. Doraiswamy PM, Kaiser L, Bieber F, Garman RL. The Alzheimer's Disease Assessment Scale: evaluation of psychometric properties and patterns of cognitive decline in multicenter clinical trials of mild to moderate Alzheimer's disease. Alzheimer Dis Assoc Disord 2001;15:174–83. [PubMed: 11723368]
25. Reines SA, Block GA, Morris JC, Liu G, Nessly ML, Lines CR, et al. Rofecoxib: no effect on Alzheimer's disease in a 1-year, randomized, blinded, controlled study. Neurology 2004;62:66–71. [PubMed: 14718699]
26. Aisen PS, Schneider LS, Sano M, Diaz-Arrastia R, van Dyck CH, Weiner MF, et al. Highdose B vitamin supplementation and cognitive decline in Alzheimer disease: a randomized controlled trial. JAMA 2008;300:1774–83. [PubMed: 18854539]
27. Harrison J, Minassian SL, Jenkins L, Black RS, Koller M, Grundman M. A neuropsychological test battery for use in Alzheimer disease clinical trials. Arch Neurol 2007;64:1323–9. [PubMed: 17846273]
28. Gilman S, Koller M, Black RS, Jenkins L, Griffith SG, Fox NC, et al. Clinical effects of A-betaimmunization (AN1792) in patients with AD in an interrupted trial. Neurology 2005;64:1553–62. [PubMed: 15883316]
29. Knopman DS, Knapp MJ, Gracon SI, Davis CS. The Clinician Interview-Based Impression (CIBI): a clinician's global change rating scale in Alzheimer's disease. Neurology 1994;44:2315–21. [PubMed: 7991118]
30. Schneider LS, Olin JT, Doody RS, Clark CM, Morris JC, Reisberg B, et al. Validity and reliability of the Alzheimer's Disease Cooperative Study- Clinical Global Impression of Change. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11(Suppl 2):S22–32. [PubMed: 9236949]
31. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993;43:2412–4. [PubMed: 8232972]

32. Gelinas I, Gauthier L, McIntyre M, Gauthier S. Development of a functional measure for persons with Alzheimer's disease: the disability assessment for dementia. Am J Occup Ther 1999;53:471–81. [PubMed: 10500855]
33. Galasko D, Bennett D, Sano M, Ernesto C, Thomas R, Grundman M, et al. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11(Suppl 2):S33–9. [PubMed: 9236950]
34. Pfeffer RI, Kurosaki TT, Harrah CH Jr. Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol 1982;37:323–9. [PubMed: 7069156]
35. Rockwood K, Fay S, Song X, MacKnight C, Gorman M. Attainment of treatment goals by people with Alzheimer's disease receiving galantamine: a randomized controlled trial. Cmaj 2006;174:1099–105. [PubMed: 16554498]
36. Cummings JL. The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. Neurology 1997;48:S10–6. [PubMed: 9153155]
37. Peskind ER, Potkin SG, Pomara N, Ott BR, Graham SM, Olin JT, et al. Memantine Treatment in Mild to Moderate Alzheimer Disease: A 24-Week Randomized, Controlled Trial. Am J Geriatr Psychiatry 2006;14:704–715. [PubMed: 16861375]
38. Schneider LS, Tariot PN, Dagerman KS, Davis SM, Hsiao JK, Ismail MS, et al. Effectiveness of atypical antipsychotic drugs in patients with Alzheimer's disease. N Engl J Med 2006;355:1525–38. [PubMed: 17035647]
39. Leber P. Slowing the progression of Alzheimer disease: methodologic issues. Alzheimer Dis Assoc Disord 1997;11(Suppl 5):S10–21. [PubMed: 9348423]discussion S37-9
40. Knopman D. Finding potent drugs for Alzheimer's disease is more important than proving the drugs are disease modifying. Alzheimer's & Dementia: Journal of the Alzheimer's Association 2006;2:147–9.
41. Fox NC, Black RS, Gilman S, Rossor MN, Griffith SG, Jenkins L, et al. Effects of A{beta} immunization (AN1792) on MRI measures of cerebral volume in Alzheimer disease. Neurology 2005;64:1563–72. [PubMed: 15883317]
42. Jack CR Jr. Slomkowski M, Gracon S, Hoover TM, Felmlee JP, Stewart K, et al. MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD. Neurology 2003;60:253–60. [PubMed: 12552040]