

Perspectives

Maximizing the Alzheimer's Disease Neuroimaging Initiative II

Maria C. Carrillo^{a,*}, Charles A. Sanders^b, Russell G. Katz^c

^aAlzheimer's Association, Chicago, IL, USA

^bFoundation for the National Institutes of Health, Bethesda, MD, USA

^cUnited States Food and Drug Administration, Washington, DC, USA

Abstract

The Alzheimer's Disease Neuroimaging Initiative is the largest public-private partnership on brain research underway at the National Institutes of Health. This 6-year study tracks cognitive and brain changes in normal subjects, those with mild cognitive impairment, and individuals with Alzheimer's disease. It was designed to provide better tools for performing effective clinical trials, and is slated to run until 2010. While data are being generated and analyzed, researchers involved in the study are developing an extension, i.e., the Alzheimer's Disease Neuroimaging Initiative II. The Foundation for the National Institutes of Health and the Alzheimer's Association convened a meeting to review the progress, evaluate future directions, and obtain the United States Food and Drug Administration's perspective on how the Alzheimer's Disease Neuroimaging Initiative could affect the drug-approval process.

© 2009 The Alzheimer's Association. All rights reserved.

Keywords: Alzheimer's disease; Biomarkers; Neuroimaging; Early detection; Longitudinal study; Natural history study; Mild cognitive impairment

1. Introduction

The largest public-private partnership on brain research underway at the National Institutes of Health (NIH), the Alzheimer's Disease Neuroimaging Initiative (ADNI), was launched by the National Institute on Aging in 2004 as an innovative, \$60 million collaboration among academia, industry partners, and the private sector (including support from the Alzheimer's Association). The principal investigator for ADNI is Michael Weiner, MD at the University of California, San Francisco.

This 6-year study tracks cognitive and brain changes in normal participants, in those with mild cognitive impairment (MCI), and in those with Alzheimer's disease (AD), to measure the progression of the condition. The public-private partnership that funds ADNI is coordinated by the Foundation for the National Institutes of Health, which has raised nearly \$25 million toward this effort.

The ADNI was designed to provide better tools for performing effective clinical trials. It grew partly from a need

to supplement clinical measures, which have diagnostic shortcomings, and to find valid surrogate markers for the early detection and monitoring of MCI and AD progression. The ADNI is slated to run until 2010, and while these data are being generated and analyzed, researchers involved in the study are developing a possible extension, the Alzheimer's Disease Neuroimaging Initiative II. This meeting was convened to review progress, evaluate future directions, and obtain the perspective of the United States Food and Drug Administration (FDA) on how ADNI could affect the drug-approval process.

2. ADNI: An update

The ADNI began in 2004 with plans to enroll 800 participants. Two hundred healthy normal participants, 400 participants with MCI, and 200 individuals with AD were slated to be followed every 6 months for 2–3 years with a variety of neuroimaging, fluid biomarker, and clinical/neuropsychological analyses [1]. Magnetic resonance imaging (MRI, at 1.5 Tesla for most, and 3 Tesla for some) was conducted on all participants; fluorodeoxyglucose positron emission tomography (FDG-PET) on half the participants; and Pittsburgh Compound B (PiB) PET amyloid imaging on 100 participants. The MRI structural imaging data are undergoing analysis by a variety

*Corresponding author. Tel.: 312-335-5722.

E-mail address: Maria.Carrillo@alz.org

The views stated here represent those of R.G.K., and do not represent an official statement of the United States Food and Drug Administration.

of methods. Clinical/neuropsychological, biomarker, and genetics data are being collected by separate core facilities, and data analysis is overseen by a statistical core.

Data are already being generated through ADNI. Imaging analysis demonstrates a statistically significant difference in the rate of brain volume atrophy between controls and MCI or AD participants, and this analysis indicates much less variability in the imaging measure than a standard clinical measurement such as the Mini Mental-State Examination [2]. In fact, analyses suggest that one of the major hopes of ADNI can be realized, i.e., that reliable surrogates can be found to allow for smaller and shorter clinical trials. The ADNI replicated several MRI studies that generated sample-size estimates, and demonstrated that smaller samples are feasible with MRI versus cognitive measures [3,4]. Using FDG-PET measures, ADNI participants with AD and MCI have characteristic reductions in regional FDG-PET measurements of cerebral glucose metabolism, which are correlated with clinical severity and progression [5,6]. After empirically defining a spatially distributed decline and reference regions-of-interest to characterize reliable metabolic declines in an ADNI training set, a preliminary analysis of an independent ADNI dataset suggests that a trial of only 56 AD participants per group would be required to demonstrate a 25% reduction in the regional rate of metabolic decline, with 80% power in a 12-month, multicenter, randomized clinical trial (K. Chen and E.M. Reiman, personal communication). Again, the estimated number of required participants is significantly smaller than that estimated from the same ADNI participants using the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) [7].

Cerebrospinal fluid (CSF) analysis, for example, largely confirms what others previously showed, i.e., that low CSF A β and high CSF τ and phospho- τ are found in participants with AD [8]. An analysis of longitudinal CSF data is eagerly awaited, particularly in regard to how these data are correlated with clinical and structural changes. Clearly there is a good correlation between CSF A β and several measures, including hippocampal atrophy, rate of change of FDG-PET signal, and conversion from MCI to AD: those with low CSF A β tend to have high conversion rates in contrast to those with high CSF A β . Pittsburgh Compound B imaging, which is yet to be completed, may shed some light on MCI converters, which are expected to be PiB-positive. The MCI participants with the highest τ and phospho- τ levels were also correlated with the greatest decline in cognition (according to the ADAS-Cog), again suggesting that biomarkers can help identify those at greatest risk.

Biomarkers may also be useful predictors for stratifying participants in prevention trials. Looking only at normal participants, ADNI data indicate that those with the lowest CSF A β show the highest rate of brain atrophy. Perhaps these individuals are already on the path to AD, but that will not be clear until further data are collected and analyzed. One of the most important contributions of ADNI would be the ability to detect changes early, which will be crucial when treating people in the earliest stages of disease.

The ADNI has had a tremendous impact. In addition to allowing free access to data that have been used in a large number of publications, there are now ADNI-like studies worldwide, including Australia, Japan, Europe, and China, creating a World Wide ADNI Working Group, supported by the Alzheimer's Association.

3. ADNI II: Renewal plans

Although ADNI began as a natural-history study to provide additional assessments of disease progression and to assist in drug approval, it is now moving toward measures that can be used for prediction, and that will ultimately lead to better treatments. As such, ADNI II plans to follow normal healthy participants and individuals with MCI, originally slated for a 3–4-year analysis, for an additional 5 years. The ADNI II also plans to add new cohorts to all three groups (controls, MCI, and AD) and to add a fourth group, made up of participants who do not fit the criteria of amnesic MCI or normal, i.e., people with very mild MCI.

The ADNI II will continue to collect structural MRI, FDG-PET, CSF, genetic, and other biomarkers. Plans may include PET scans for every new participant. The F-18 amyloid imaging agents have become more widespread, and by 2010, when renewal would begin, it is anticipated that the majority, if not all, of the ADNI sites would have PET scanners and access to those ligands.

The ultimate goal of ADNI is to identify the best biomarker methods to be used in phase 2 and 3 clinical trials: those with high rates of change, small standard deviations, high power, and correlations with clinical measures. Such measures will have to be validated in a treatment setting before eventual use as surrogates in prevention trials. For this reason, the FDA's perspective on ADNI and on the use of biomarkers in AD clinical trials in general is an important factor in deciding future ADNI directions.

4. ADNI: FDA perspective

The AD biomarkers that emerge from ADNI studies will be subject to regulatory-agency approval. How the approval process will function and how it will affect labeling is not entirely clear, because this is new territory for all concerned. But certain guiding principles, to which the FDA must adhere, are informative.

4.1. Biomarker validation/qualification

The FDA feels that biomarkers may eventually be useful for participant selection and stratification in clinical trials. Although further work is required to conclude reliably that individuals surrogates and their markers can predict different participant strata, it is very likely that ADNI and ADNI II can contribute a great deal in that regard. However, there will be ramifications. The use of biomarkers for participant selection or stratification will most likely be reflected in

labeling because, as a general rule, the principle in labeling is to describe the specifics of a trial, including the types of population studied. Ultimately, these markers might become approved diagnostics, in which case they might not be described in the label. However, it is likely that for the foreseeable future, any selection criteria that pertain to biomarker use will be included in the label.

Although biomarker-based stratification information would be included in labeling, the results of biomarker changes would not necessarily be included, because that would suggest that the biomarkers are appropriate measures of drug effect. The ADNI is not designed to test this. True validation of these markers only comes in the treatment setting, from adequately controlled trials that show how these markers function in relation to treatment, and how they correlate with clinical outcomes. In general, there must be good reason to believe that labeling statements have useful meaning. Although statements may be factual, they cannot misleadingly imply something that may not be true, such as a disease-modifying effect.

In addition to stratification, another potential use of biomarkers is for sample-size calculations based on biomarker changes. The FDA agrees that biomarker data could be useful in this way. In such a scenario, however, the implication arises that the biomarker would also have to be used as an outcome measure. In addition, the FDA would not currently approve a drug based solely on one of these biomarkers or imaging markers. Although the biomarker may allow a reduction in sample size of a trial arm, some clinical meaning must be attributed to a change in the biomarker alone: basically, the marker must be validated before it could stand alone as a measure.

4.2. *Qualification of biomarkers as surrogate endpoints*

The FDA would agree that any biomarker that changes consistently in response to treatment and is known to correlate with clinical benefit is, for all intents and purposes, a valid biomarker. When a particular biomarker (or biomarkers) responds in the same way to drugs that have multiple and different mechanisms of action, as shown in multiple, adequate trials, that would again be a scenario where a biomarker could be declared a validated surrogate. The FDA would posit that drug approval based on a change in such a validated biomarker is possible.

To consider a biomarker as a *suitable* surrogate endpoint for AD poses a more difficult challenge. A validated surrogate, as outlined above, would be acceptable, but there is no hard answer in regard to what a “suitable” surrogate might be. The law permits approval of a drug on the basis of its effect on a surrogate if it is “reasonably likely” to predict a clinical benefit. Thus it all depends on the definition of “reasonably likely.” The more knowledge we have about pathology and drug action, the better we can define “reasonably likely,” but it is impossible to know this in a vacuum. Considering specific examples may be helpful in addressing this question. If results from ongoing antibody and γ -secre-

tase trials demonstrate that lowering $A\beta$ correlates with clinical improvement and reduced rates of atrophy, for example, then using PiB-binding as a biomarker in a prevention trial, with reduced brain atrophy as an outcome, may be attractive.

The FDA is frequently asked whether a biomarker that was used as a selection criterion in a trial has to be approved as a diagnostic. The answer to that question is generally yes, and the diagnostic would have to be approved at the same time as the therapeutic agent.

The FDA is also often asked whether one diagnostic, e.g., an amyloid ligand such as PiB, is interchangeable with an equivalent. If the diagnostics are truly similar, then they may be interchangeable, but if not, then the label claims must be specific for each diagnostic.

4.3. *Prodromal/incipient AD diagnosis: role of biomarkers*

There is considerable interest in studying individuals with prodromal AD, but defining those individuals is not easy. Dubois et al [9] suggested that a diagnosis of incipient or prodromal AD could be based on a combination of deficit in cognitive function (e.g., >1.5 standard deviations on a delayed recall task) and at least one or more biomarkers linked mechanistically to AD pathology (e.g., CSF $A\beta_{42}$, total τ , amyloid load judged by single-photon emission computed tomography imaging, or regional/global brain atrophy according to MRI), but it is impossible for the FDA to render any judgments about the value of such criteria. Dubois et al [9], for example, described data from different sources that would have to be reviewed, and the quantitative standards of change for various biomarkers that would have to be related to diagnosis. In principle, the FDA is in favor of developing reliable criteria for the early identification of AD. It also remains to be seen if the criteria of Dubois et al [9] would make a suitable definition for MCI. In addition, in the case of defining MCI, it is important for a drug claim to show that a nonexpert can render a diagnosis of MCI.

The ADNI is the perfect data source for the type of data that could validate the criteria of Dubois et al [9]. The identification of early AD in several subgroups may be one of the most exciting results to come from ADNI. However, the FDA believes that autopsy data are at some point crucial to ensure a correct diagnosis.

4.4. *Validated biomarkers and other endpoints in prodromal/incipient AD trials*

Various measures would be acceptable to the FDA as endpoints in prodromal/prevention trials. Cognitive benefit over time, as judged by very sensitive neuropsychological testing, would be acceptable, for example, although the measures would have to be extraordinarily sensitive to be useful in individuals who have no complaint. The caveat here is that drugs for AD are currently required to exert an effect on a cognitive measure. Proving that a drug has a clinically meaningful effect is an issue that would need to be addressed in this context.

Time/delay to a clinical endpoint is also a perfectly acceptable measure. One problem with that approach, however, is subjectivity. Variability from investigator to investigator may introduce noise in the data. As long as that noise does not translate into a bias, time to a clinical endpoint is one way to demonstrate benefit in a nondemented population.

Biomarker measures (e.g., CSF A β / τ , or structural brain atrophy) would not be acceptable currently as a primary outcome measure, but they may have some validity in the future. This would be particularly helpful in the prodromal population, who may be 20 years away from becoming symptomatic, and in whom a surrogate (ideally a validated one) would be appropriate as an outcome measure. On the other hand, biomarker data could be used as a coprimary outcome, along with a sensitive neuropsychological measure. However, a clinically meaningful effect would still have to be achieved. The field is probably not at that point yet, but the FDA would certainly consider this approach, should it be applicable.

4.5. *Clinical trial design /statistics*

The FDA has nothing against the inclusion of different participant groups, such as individuals with prodromal AD and MCI, or those diagnosed with mild-to-moderate AD, in clinical trials. However, the FDA suggests that it may be better to separate studies when the participants are truly different, or when it is not known if they are not different. The key involves whether claims can be made for all groups, even though efficacy may not be shown within individual groups. In some cases, cross-group claims may be appropriate, but in other cases, depending on how different the groups are, the FDA will insist on seeing significant data for each group. It is likely that true statistical significance would have to be demonstrated in prodromal or nonsymptomatic participants, versus those with some measurable impairment. This could be accomplished, but the outcome measures would not be the same as for individuals with mild AD, so it may not make sense to combine those groups into one study.

This question of group stratification is also germane to biomarkers. The best way to show that a particular biomarker is acceptable as a primary outcome in presymptomatic participants is to show a correlation between the effect on that surrogate and clinical outcomes, i.e., symptomatic individuals or individuals close to becoming symptomatic. Time to diagnosis, for example, could be used in participants with imminent AD, and after that biomarker has proven itself in that population, it could be applied to a similar study in those who are not likely to show clinical symptoms for 20 years.

4.6. *Biomarkers as endpoints in AD studies*

On the use of biomarkers as endpoints, the FDA takes the position that if a trial demonstrates internal validation (e.g., if two 1-year randomized control trials in participants with mild-to-moderate AD revealed a decreased rate of change

in a prespecified brain region, plus a significant finding in a measure of cognition), then the FDA would seriously consider this a provisionally approvable application. That is not to say, however, that these data would warrant a disease-modification claim, because they do not represent a true validation. It is unclear that extending results to 2 years would remove the “provisional” modifier for such a claim, because it is unclear that 2-year data would help validate that surrogate any more fully than 1-year data. As for labeling, this would likely reflect the provisional nature of the claim, insofar as statements are added to labels not only because they are true and accurate, but because they provide some meaning to the clinician.

The FDA would also consider a combination of a cognitive measures and biomarker data as outcome measures, without the need for a global measurement. However, the question of clinical meaningfulness needs to be addressed. Whether or not a global measure can be discarded may depend on the type of claim desired. If a sponsor is interested in a disease-modifying claim but does not want to run a randomized-withdrawal or randomized-start design, a biomarker approach would be an option. If such a trial included three outcome measures (the usual cognitive and global, plus a biomarker), it would begin to address the validation question on a small scale. The FDA has been very clear that it would be willing to entertain such a trial as a possible way to obtain a disease-modifying claim. It is not necessarily the preferred way, but such an application would be reviewed.

On a slightly different variation of this theme, the FDA would have to evaluate the use of a global measure and a biomarker measure without a cognitive measure. In general, to obtain a claim for AD, a drug must have show a cognitive effect and a measure of clinical significance, i.e., a global measure. This is a reasonable approach, but the FDA would be willing to consider other criteria.

4.7. *Potential label claims from trials adding a novel therapy to approved symptomatic therapies*

According to the FDA's position on adding novel therapy to approved symptomatic therapies, an appropriately designed comparator study that includes biomarkers could certainly support a claim of superiority, although replication would probably be needed. As far as label claims are concerned, stratification by biomarker is perfectly acceptable, and would result in an adjunctive claim, which is common in various settings. The specific diagnostic criteria that defined the population would need to be described in the label, and would also be outlined in the indications.

4.8. *Biomarkers as baseline covariants*

Considerable data suggest that the rate of change of cognition and clinical state accelerates as AD progresses, and that an imaging or biomarker measurement of the extent of progression should have some predictive power for future

rate of change. The FDA does not object to the use of covariates in statistical models of primary outcomes in clinical trials, although having a statistician weigh in on the analysis could be beneficial.

5. Conclusions

The ADNI has proven itself a successful and impactful private-public partnership of academic laboratories, industry partners, and the private sector. Its original mission to search for clinical measures and valid surrogate markers for the early detection and monitoring of AD progression has evolved, and ADNI has shown great promise in the area of biomarker validation and in the potential validation of updated criteria for AD, such as those recommended by Dubois et al [9].

The ADNI has remained faithful to the precept that the data generated are freely accessible to the worldwide scientific community, a priceless commodity of unique proportions. The ADNI's outreach has attracted global collaboration and the creation of the World Wide ADNI Working Group. Support for ADNI II is widespread among the National Institute on Aging, industry partners, and the private sector. Many are looking forward to additional data and to answering many questions that can only be addressed by a natural-history study such as ADNI.

Acknowledgments

The Alzheimer's Association and the Foundation for the NIH thank the participants of this meeting, including Patricia Cole, MD, Eisai Global Clinical Development, ADNI Industry Scientific Advisory Board Chair; Michael Weiner, MD, University of California at San Diego, ADNI Principal Investigator; Neil Buckholtz, PhD, National Institute on Aging; and Tom Fagan, Science Writer. In addition to the National Institute on Aging, the ADNI public-private partnership includes federal support from the National Institute for Biomedical Imaging and Bioengineering (also part of the NIH), and the participation of the FDA. Private-sector funders include Abbott, AstraZeneca AB, Bayer Schering

Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Merck and Co., Inc., Novartis AG, Pfizer, Inc., F. Hoffmann-La Roche, Schering-Plough, Synarc, Inc., and Wyeth, as well as nonprofit partners, i.e., the Alzheimer's Association and the Institute for the Study of Aging. More information on the Foundation for the NIH is available at <http://www.fnih.org/>.

References

- [1] Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack CR, Jagust W, et al. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement* 2005;1:55–66.
- [2] 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.
- [3] 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 Alzheimer's. *Neurology* 2003;60:253–60.
- [4] Jack CR Jr, Shiung MM, Gunter JL, O'Brien PC, Weigand SG, Knopman DS, et al. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. *Neurology* 2004;62:591–600.
- [5] Langbaum JBS, Chen K, Le W, Reschke C, Bandy D, Fleisher AS, et al. Categorical and correlational analyses of baseline fluorodeoxyglucose positron emission tomography images from the Alzheimer's Neuroimaging Initiative (ADNI). *Neuroimage* 2009; in press.
- [6] Reiman EM, Chen K, Ayutyanont N, Lee W, Bandy D, Reschke C, et al. Twelve-month cerebral metabolic declines in probable Alzheimer's disease and amnesic mild cognitive impairment: preliminary findings from the Alzheimer's Disease Neuroimaging Initiative. *Alzheimers Dement* 2008;4:T58.
- [7] Mohs RC, Cohen L. Alzheimer's disease assessment scale. *Psychopharmacol Bull* 1988;24:627–8.
- [8] Shaw LM, Vanderstichele H, Clark CM, Aisen PS, Petersen RC, Blennow K, et al. Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging subjects. *Ann Neurol* 2009 March 18; [Epublication ahead of print].
- [9] Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* 2007;6:734–46.