Neurobiology of Aging publishes the results of studies in behavior, biochemistry, cell biology, endocrinology, molecular biology, morphol-
ogy, neurology, neuropathology, pharmacology, physiology and protein chemistry in which the primary emphasis involves mecha-
nisms of nervous system changes with age or diseases associated with age. Only original articles will be accepted. Reviews and primary research articles are included, occasionally accompanied by open peer
commentary. Letters to the Editor and brief communications are also acceptable. Brief reports of highly time-sensitive material are usu-
ally treated as rapid communications in which case editorial review is completed within six weeks and publication scheduled for
the next available issue. Negative results are published as one-page manuscripts (three typed double-spaced pages), with
supplementary material at the Neurobi-
ology of Aging web site.

The accepted abbreviation for Neurobiology of Aging for bibliographic citation is Neurobiol. Aging.

Instructions for submitting an article to Neurobiology of Aging may be found online http://ees.elsevier.com/nba.
NEUROBIOLOGY
OF
AGING

Editor-in-Chief
PAUL D. COLEMAN
Sun Health Research Institute, Sun City, AZ, USA
and
University of Rochester Medical Center, Rochester, NY, USA

Volume 32 (2011)
SUPPLEMENT

Emerging Consensus for Alzheimer's Biomarkers in Clinical Trials

EDITORIAL TEAM

Guest Editors:

John C. Morris, M.D.
Harvey A. and Dorismae Hacker Friedman Distinguished Professor of Neurology
Director, Knight Alzheimer Disease Research Center
Departments of Neurology, Pathology and Immunology
Washington University
St. Louis, MO 63110, USA

and

Dennis J. Selkoe, M.D.
Co-Director, Center for Neurologic Diseases
Department of Neurology, Brigham & Women’s Hospital
The Vincent and Stella Coates Professor of Neurologic Diseases
Harvard Medical School
Boston, MA 02114, USA
Emerging Consensus for Alzheimer's Biomarkers in Clinical Trials

Contributors:
Paul S. Aisen, M.D., University of California San Diego
Marilyn Albert, Ph.D., Johns Hopkins University School of Medicine
David M. Holtzman, M.D., Washington University, St. Louis
Clifford R. Jack, M.D., Mayo Clinic
William E. Klunk, M.D., Ph.D., University of Pittsburgh School of Medicine
Richard Mayeux, M.D., Columbia University
John C. Morris, M.D., Washington University, St. Louis
Eric M. Reiman, M.D., Banner Alzheimer's Institute
Dennis J. Selkoe, M.D., Brigham and Women’s Hospital, Harvard Medical School
Reisa Sperling, M.D., Brigham and Women’s Hospital, Harvard Medical School
John Q. Trojanowski, M.D., Ph.D., University of Pennsylvania

Organizers:
Cynthia Bens, Alliance for Aging Research
Daniel Perry, Alliance for Aging Research

Supporters:
Anonymous Donor
The Henry and Marilyn Taub Foundation
The Vradenburg Foundation
CONTENTS

SUPPLEMENT: Emerging Consensus for Alzheimer’s Biomarkers in Clinical Trials

INTRODUCTION

Recommendations for the incorporation of biomarkers into Alzheimer clinical trials: An overview
J.C. MORRIS, D.J. SELKOE ................................................................. S1

BIOFLUIDS

CSF biomarkers for Alzheimer’s disease: Current utility and potential future use
D.M. HOLTZMAN ................................................................. S4

Blood-based biomarkers for Alzheimer’s disease: Plasma Aβ40 and Aβ42, and genetic variants
R. MAYEUX, N. SCHUPF ..................................................... S10

IMAGING

Amyloid imaging as a biomarker for cerebral β-amyloidosis and risk prediction for Alzheimer dementia
W.E. KLUNK ................................................................. S20

The potential of functional MRI as a biomarker in early Alzheimer’s disease
R. SPERLING ................................................................. S37

Fluorodeoxyglucose positron emission tomography: Emerging roles in the evaluation of putative Alzheimer’s disease-modifying treatments
E.M. REIMAN, ON BEHALF OF THE ALZHEIMER’S DISEASE BIOMARKERS WORKING GROUP FOR THE ALLIANCE FOR AGING RESEARCH .................................................... S44

Alliance for Aging Research AD Biomarkers Work Group: Structural MRI
C.R. JACK, JR. ................................................................. S48

COGNITIVE

Changes in cognition
M.S. ALBERT ................................................................. S58

USE OF BIOMARKERS TO EXPEDITE CLINICAL TRIALS

Clinical trial methodologies for disease-modifying therapeutic approaches
P.S. AISEN ................................................................. S64
Introduction

Recommendations for the incorporation of biomarkers into Alzheimer clinical trials: an overview

John C. Morrisa,*, Dennis J. Selkoeb

a Departments of Neurology, Pathology, and Immunology, Knight Alzheimer Disease Research Center, Washington University, St. Louis, MO, USA
b Department of Neurology, Center for Neurologic Diseases, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA

The Alliance for Aging Research leads a coalition of not-for-profit organizations representing the interests of patients, caregivers, health advocates, and older Americans that aim to advance the development of transformational therapies for Alzheimer’s disease (AD). There are no currently approved therapies that can prevent, cure, or even substantially slow the progression of the dementia caused by AD. The aging of our society ensures that the already enormous annual costs of Alzheimer dementia will rise from $172 billion in 2010 (representing more than one-sixth of Medicare expenditures) to well over $1 trillion in 2050, even without inflationary adjustment (Alzheimer’s Association, 2011). In the absence of truly effective therapies for AD, the illness may well bankrupt our health care system as it robs generations of Americans of their most human qualities and creates a crushing caregiving burden for millions of families.

Unfortunately, to date all clinical trials of putative “disease-modifying” compounds in individuals with Alzheimer dementia have failed to demonstrate meaningful clinical benefit. There are several possible reasons for the failure of these trials, including: (1) inadequacies of the specific drugs tested; (2) inappropriate pathobiological targets; (3) inability to detect drug benefit, as might occur with insensitive outcome measures or if numerous trial participants were misdiagnosed and did not have AD; and (4) the key consideration that therapeutic intervention will be most successful prior to the onset of frank dementia, in the presymptomatic or very early clinical stage of AD before substantial neurodegeneration has occurred (Bateman et al., 2011). The development of biological markers for AD (i.e., objectively measured indicators of the disease), however, may help address many of these issues (Hampel et al., 2010; Perrin et al., 2009; Siemers et al., 2010).

Biomarkers for AD can be divided broadly into those that identify aspects of the molecular pathology of AD and those that reflect the “downstream” consequences of that pathology. In general, molecular biomarkers identify the disease process regardless of clinical status (i.e., they can identify AD in both its presymptomatic and symptomatic stages), whereas biomarkers reflecting downstream effects typically correlate more strongly with the symptomatic stage (i.e., Alzheimer dementia). Both these types of biomarkers have been evaluated in AD primarily for 3 classification purposes: (1) diagnosis, to assess whether symptomatic individuals have the condition; (2) screening, to assess whether asymptomatic individuals have the condition; and (3) prognosis, to assess whether individuals with the condition will achieve subsequent outcomes, such as appearance of symptoms or progression of symptom severity (Pepe et al., 2008). The process of biomarker development begins with the discovery of biomarker candidates and their initial validation in independent samples of human research participants. Then their classification accuracy is evaluated in clinical settings to determine sensitivity and specificity for diagnosis (e.g., in comparison with clinically and neuropathologically assessed expired brain donors), screening, or predicting outcomes. The final determination measures their usefulness in medical decision-making. Inherent in this process is the need to standardize protocols for measuring the biomarkers and to establish effective quality control procedures. While much progress has been made in these endeavors, still more work is needed before the appropriate roles of biomarkers are fully realized.

Although less extensively studied thus far, there also are important uses for biomarkers in AD clinical trials. For example, biomarkers may serve as endpoints in clinical trials, such that they could be used to assess an effect of a
therapeutic agent on the progression of AD, provide information needed by sponsors in deciding whether or not to invest in a larger study using clinical endpoints, or help to support a treatment’s disease-modifying effects. There is growing interest in the role they could play as surrogate endpoints in AD treatment trials, such that treatment effects on the biomarker could be used as a replacement for clinical endpoints in the accelerated approval of AD therapies by regulatory agencies. Because biomarkers likely have greater precision in measuring biological change than do clinical assessment tools in measuring dementia progression, incorporation of biomarkers into clinical trials may have the beneficial result of requiring fewer participants studied for shorter durations, thus notably reducing the costs of the trials. Further cost savings could be realized if biomarkers are used to enrich trials with individuals having a high likelihood of more rapid cognitive decline than other AD subjects; selective enrollment of “relatively rapid decliners” could mean fewer participants and less time needed to demonstrate a treatment effect (Snider et al., 2009). Biomarkers can also be used to enhance confidence that individuals enrolled in the trials are accurately diagnosed with AD, thus minimizing the masking of treatment effects that occurs when misdiagnosed persons are enrolled.

To advance the use of biomarkers in experimental studies of putative therapies for AD, and thus hopefully to accelerate effective drug development, the Alliance for Aging Research assembled a group of senior investigators engaged in AD biomarker research and commissioned them to provide an overview of the potential uses of specific biomarkers in AD clinical trials. This special issue is the culmination of the work of these experts. It addresses the current status of biomarkers for AD, their potential for enrichment of AD clinical trials. This special issue is the culmination of the work of these experts. It addresses the current status of biomarkers for AD, their potential for utilization in the design and conduct of clinical trials, and additional data that may be needed to fully establish their usefulness.

Although not all candidate biomarkers for AD are reviewed here, those that we selected are the most widely studied and thus most likely to be incorporated into clinical trials. Correspondingly, these biomarkers also are included in recently revised diagnostic guidelines for Alzheimer dementia intended to enhance diagnostic confidence for AD (McKhann et al., 2011), although the guidelines recognize that full validation of the diagnostic use of biomarkers still is needed. Molecular biomarkers for AD measure the major protein constituents of the neuropathologic hallmarks of AD, amyloid plaques and neurofibrillary tangles. In AD, the neurally secreted amyloid beta-protein (Aβ), particularly the species that is 42 amino acids long (Aβ), forms insoluble aggregates that are deposited as innumerable extracellular plaques. Hyperphosphorylated tau proteins accumulate intracellularly to form neurofibrillary tangles. Both Aβ and tau (total tau and phosphorylated tau) can be measured in biofluids.

David Holtzman (Washington University) reviews the status of cerebrospinal fluid (CSF) assays for Aβ and tau/phospho-tau as AD biomarkers (Holtzman, 2011). He indicates how these assays can be used in clinical trials to exclude individuals with a dementia disorder that is not associated with amyloid deposition (and therefore not AD).

Richard Mayeux (Columbia University) discusses investigations of plasma Aβ markers (Mayeux and Schupf, 2011). Mayeux also reviews the known genetic biomarkers for AD, namely dominantly inherited pathogenic mutations in the amyloid precursor protein (APP), presenilin 1 (PSEN1), and presenilin 2 (PSEN2) genes that appear to cause AD by increasing brain levels of Aβ. Mayeux also discusses the role of genetic variants such as apolipoprotein E (APOE) in assessing the risk of Alzheimer dementia.

The Aβ neuropathology of AD now can be assessed with amyloid imaging agents that detect fibrillar Aβ (amyloid) deposits in vivo. William Klunk (University of Pittsburgh) reviews the development of 11C-labeled and 18F-labeled amyloid tracers for positron emission tomography (Klunk, 2011). He also cites the initial use of amyloid imaging as an outcome measure in a clinical trial of passive immunotherapy directed against Aβ in patients with AD (Rinne et al., 2010).

Other imaging modalities appear to detect “downstream” events in the pathological cascade that is initiated by the molecular pathology of AD (i.e., Aβ and tau build-up). A potentially very early consequence of AD pathology may be the disruption of the functional integrity of neurons and their cerebral networks, as reviewed by Reisa Sperling (Harvard Medical School; Sperling, 2011) as regards functional magnetic resonance imaging and by Eric Reiman (Banner Alzheimer’s Institute; Reiman, 2011) as regards fluorodeoxyglucose positron emission tomography. The imaging of neuronal loss caused by AD with magnetic resonance imaging (MRI) techniques is the most commonly used and best validated AD biomarker (cerebral atrophy reflects neurodegeneration), and Clifford Jack (Mayo Clinic) notes that the topographic pattern of atrophy in imaging studies of AD correlates with the pattern of neurofibrillary pathology (Jack, 2011). Importantly, these neurofibrillary lesions are composed of the tau protein, which can be sensitively measured in CSF (above). Finally, Marilyn Albert (Johns Hopkins University) and Paul Aisen (University of California, San Diego) consider the relevance of AD biomarkers to the major clinical feature of AD, cognitive decline (Albert, 2011), and to improving clinical trial design for “disease-modifying” therapies (Aisen, 2011).

These authors also indicate what remains to be accomplished regarding the standardization of biomarker assays to minimize interlaboratory variability and to establish uniform protocols for sample collection. For example, because the CSF levels of Aβ may exhibit diurnal fluctuations (Butem et al., 2007), it may be important to develop guidelines for the standard timing of sample collection.
Another example is the need for rapid, automatic quantification of regional volume loss in the MRI assessment of cerebral atrophy and for applicable age-adjusted normative values. Once uniform, well-standardized protocols for measuring a particular biomarker are widely available, its utility should be revaluated in appropriately large cohorts. Attention should be given to evaluating combinations of the above biomarkers and to changes in biomarker levels over time, as it is unlikely that a single biomarker assayed at a single time point will yield adequate sensitivity and specificity for diagnosing AD. Efforts to discover and develop additional biomarkers for AD also should be strongly supported. Nonetheless, certain of the currently available AD biomarkers discussed herein—in particular, structural MRI, amyloid imaging, and CSF levels of Aβ1–42 and tau—have been sufficiently studied and validated in many populations that the authors of this special issue recommend their use to improve the selection of appropriate individuals for AD clinical trials and to serve as surrogate outcome measures in trials of potentially disease-modifying AD therapies. While further research lies ahead, these specific biomarkers are now appropriate for incorporation into AD clinical trial design, and their inclusion is likely to speed the identification of effective disease-modifying treatments and prevention for this currently incurable disorder.

Disclosure statement

The authors disclose no conflicts.

Acknowledgements

This special issue was sponsored through the nonprofit Alliance for Aging Research with charitable gifts from the Henry and Marilyn Taub Foundation, the Vradenburg Foundation, and by an anonymous donor. We acknowledge with thanks the contributions of Marc Walton, MD, PhD, Associate Director, Office of Translational Sciences, who served as liaison to the Food and Drug Administration (FDA), and of John Trojanowski, MD, of the University of Pennsylvania and Co-Director of the Biomarker Core of the Alzheimer’s Disease Neuroimaging Initiative (ADNI). Dr. Morris is supported by National Institute on Aging grants P50 AG05681, P01 AG03991, P01 AG026276, and U01 AG032438; Dr. Selkoe’s work is supported by National Institute of Health R37 grant AG06173.

References

Biofluids

CSF biomarkers for Alzheimer’s disease: current utility and potential future use

David M. Holtzman

Department of Neurology, Developmental Biology, Knight Alzheimer’s Disease Research Center, and the Hope Center for Neurological Disorders, Washington University, St. Louis, MO, USA

Abstract

Over the past 15 years, cerebrospinal fluid (CSF) biomarkers have been shown to be useful for both the diagnosis as well as the prognosis in Alzheimer’s disease. It has been shown the CSF levels of amyloid-β (Aβ)42 are a very good marker for the presence of amyloid deposition in the brain regardless of clinical status and that total tau and phosphorylated forms of tau are useful in detection of neurodegeneration. When combined together, these CSF markers are useful not only in differential diagnosis but also in predicting conversion and rate of progression from mild cognitive impairment/very mild dementia to more severe impairment. The markers are also useful in predicting conversion from cognitive normalcy to very mild dementia. This field is briefly reviewed and recommendations for future studies in this area are provided. © 2011 Elsevier Inc. All rights reserved.

Keywords: Alzheimer’s disease; CSF; Tau; Biomarkers

1. Introduction

With the emergence of disease-modifying strategies for the treatment of Alzheimer’s disease (AD), impetus has intensified to diagnose the condition in its early “preclinical” stages—before significant brain damage has occurred. Since their first description in 1907, amyloid plaques and neurofibrillary tangles (NFTs) have been the hallmark histopathological features of AD. Historically, they have also been associated with the dementia caused by the disease. It is clear, however, that these lesions begin to accrue in significant amounts in many “cognitively normal” elderly individuals (Crystal et al., 1988) and likely together with other synaptic and cellular damage, must reach a threshold to enable the clinical manifestations we recognize as dementia. A growing body of evidence now supports the idea that amyloid plaques and NFTs actually define but do not fully repre-
portantly, CSF biomarkers can now be useful for clinical trials and for clinicians to predict prognosis of individuals who have mild cognitive impairment/very mild dementia as well as for those who are cognitively normal.

The sampling of CSF represents the most direct and convenient means to study the biochemical changes occurring in the central nervous system. Therefore, CSF is an attractive resource for ongoing research into AD biomarkers. Implicated by biochemical and immunohistological studies of AD brain tissue, the major protein constituents of the hallmark pathological features of the disease (amyloid-β [Aβ]42, tau, and phosphorylated forms of tau) have emerged as the current leading diagnostic and prognostic CSF biomarkers (Table 1).

### 2. CSF amyloid-β

Amyloid-β is a secreted peptide of unknown physiological function that is cleaved from the amyloid precursor protein (APP) by the sequential activities of β-secretase and γ-secretase enzymes. The majority of Aβ is produced in the brain and secreted into the brain extracellular space. Some fraction of central nervous system (CNS)-produced Aβ diffuses into the CSF, appearing in modest concentrations (approximately 10–15 ng/mL). Aβ occurs in multiple forms, including those ranging from 37 to 43 amino acids in length. Among these, Aβ40 is the most abundant species, but Aβ42 seems to be essential for initiating amyloid-β aggregation and is considered central to the amyloid cascade hypothesis of AD (Hardy and Selkoe, 2002). Of these 2 species, Aβ42 has emerged as a useful biomarker for AD.

Although the finding is initially counterintuitive, the mean concentration of Aβ42 in the CSF is significantly reduced by about 50% in subjects with AD relative to age-matched controls (Motter et al., 1995; Sunderland et al., 2003); this phenomenon is thought to result from deposition of Aβ42 in amyloid plaques, preventing its transit from the brain into the CSF (plaques acting as a sink). In support of this hypothesis, when antemortem CSF Aβ42 concentrations are compared with results from amyloid imaging in the same individual or to postmortem measurements of brain Aβ load, virtually all individuals with fibrillar Aβ deposits show low concentrations of Aβ42 in the cerebrospinal fluid, independent of cognitive status (Fagan et al., 2006, 2007; Jagust et al., 2009; Tapiola et al., 2009). Thus, CSF Aβ42 can serve as a diagnostic and surrogate biomarker for Aβ deposition in the brain. Unlike CSF Aβ42, CSF Aβ40 levels are not different in individuals with AD compared with controls (Shoji et al., 1998). The decrease in CSF Aβ42 appears to precede amyloid retention as detected by amyloid imaging using compounds such as 11C-labeled Pittsburgh Compound B (11C-PIB), signifying what is perhaps the first evidence of AD pathology in cognitively normal individuals (Cairns et al., 2009; Fagan et al., 2006, 2009). While CSF Aβ40 does not differentiate individuals with AD from controls, CSF Aβ40 has recently been shown to be decreased in a subset of subjects with cerebral amyloid angiopathy (CAA) (Verbeek et al., 2009). Aβ42 alone is less useful in
differentiating AD from other dementias, because low levels have also been documented in patients with frontotemporal dementia (FTD), vascular dementia, and dementia with Lewy bodies (DLB), though it is possible it is low in many of these patients because of the concomitant presence of fibrillar Aβ deposits, such as occurs in the majority of individuals with DLB.

Despite its utility in the detection and differential diagnosis of dementia, CSF Aβ42 does not correlate well with disease duration or severity. This is consistent with results from 11C-labeled Pittsburgh Compound B studies showing that amyloid retention does not change appreciably during the symptomatic stages of AD (Rowe et al., 2007), and further supports results from pathological studies of AD; amyloid pathology occurs very early in the disease process and has relatively stabilized by the time the first clinical signs of dementia appear.

3. CSF tau

Tau is a cytosolic protein predominantly expressed in neurons, wherein its primary function seems to be regulation of microtubule stability within the axon. This function is regulated by several different posttranslational modifications, principally phosphorylation of numerous serine and threonine residues. In AD, hyperphosphorylated tau often fills the dystrophic neurites of neuritic plaques, and is the principle component of the paired helical filaments that constitute NFTs that are present in neuronal cell bodies. The precise forms of tau that appear in the CSF, and the mechanism or mechanisms by which they get there, are not entirely understood, but recent studies (Portelius et al., 2008) demonstrate that virtually all domains of the protein are represented, and it is widely assumed (but not proven) that the major sources of increases in tau and phosphorylated tau in the CSF in AD are either due to synaptic and/or neuronal injury, cell death, or possibly neurofibrillary tangles.

3.1. Total (T)-tau

Tau is the major protein component of intraneuronal NFT and is elevated in the CSF in most patients with AD. In addition to the presence of tau in neurofibrillary tangles, it has been shown that tau levels in CSF can increase rapidly as a result of neuronal injury, and therefore, may indicate the severity of the underlying neurodegeneration (Blennow, 2004). Over 50 studies have demonstrated an increase in the concentration of total tau (t-tau) by approximately 2–3-fold in AD compared with nondemented elderly subjects (Blennow et al., 2001). Elevation of CSF tau differentiates AD from nondemented, age-matched elderly with a sensitivity and specificity of approximately 90% (Sunderland et al., 2003). As mentioned previously, tau elevation seems to occur at the early symptomatic stages of disease (mild cognitive impairment [MCI] and/or very mild dementia) and in some cognitively normal individuals, where its levels correlate with the amount of amyloid deposition and together with Aβ42 predict cognitive decline (see below). Cognitively normal individuals with evidence of amyloid deposition and increased tau are likely to have preclinical AD (see below) (Fagan et al., 2009). However, it is important to consider that tau elevation can be seen in other neurodegenerative diseases, potentially limiting the utility of tau alone in the differential diagnosis of AD (Arai et al., 1997). Tau, as a marker of neuronal injury, can be transiently increased after any acute brain injury (such as stroke or trauma) (Hesse et al., 2001). Moreover, tau levels seem to remain relatively stable throughout the clinically symptomatic period of AD (Sunderland et al., 1999) and do not correlate well with dementia severity. Age might affect the CSF levels of tau; however, studies have been conflicting regarding the direction and significance of such an effect (de Leon et al., 2007).

3.2. P-tau

Abnormal tau phosphorylation is present in neurofibrillary tangles and has been investigated as a marker of AD pathology. As many as 30 different phosphorylation sites of phosphorylated (p)-tau have been identified (Bueé et al., 2000), and enzyme-linked immunosorbent assays (ELISAs) have been developed for at least 5 of them. Studies examining the utility of different forms of p-tau in the early diagnosis of AD, and in the differentiation from other causes of dementia, have consistently shown that p-tau 181 (Arai et al., 2000), p-tau 231–235, or p-tau 396–404 (Hu et al., 2002) offer at least equivalent diagnostic utility for AD as compared with total tau. Studies comparing the diagnostic performance of different phosphorylation sites (p-181, p-199, and p-231) suggest that all 3 assays are equally effective in differentiating AD from nondemented controls. P-tau 231 may provide diagnostic specificity for AD and may improve the differentiation between AD and frontotemporal dementia (Buerger et al., 2002), while there is some evidence that p-tau 181 improves the differentiation between AD and DLB (Hampel et al., 2004). P-tau 396–404, and the ratio of p-tau 396–404/t-tau, but not tau alone, has been shown in 1 study to differentiate AD from vascular dementia (Hu et al., 2002). In contrast to t-tau, p-tau does not appear to be increased secondary to acute brain injury, further adding to its diagnostic specificity.

4. Combination of Aβ42 and tau

4.1. Diagnosis

Based on current data, the use of CSF Aβ42 alone but especially together with t-tau or p-tau 181 is very useful in
both diagnosis and prognosis of individuals with MCI/very mild dementia and also in predicting progression from cognitive normalcy to MCI/very mild dementia. This is likely due to the fact that the levels of the markers together can identify 2 aspects of AD pathology, plaques ($\alpha B42$), and tauopathy/neurodegeneration (tau).

The combination of $\alpha B42$ and total tau or p-tau as a ratio or as an index provides the best discriminative value to date for individuals with AD compared with healthy controls of the same age, with a sensitivity of approximately 85%–90% and a specificity of approximately 85%–90% (Shaw et al., 2009; Welge et al., 2009) as verified by autopsy. Slightly lower sensitivities and specificities are seen with MCI/very mild dementia versus age-matched controls; however, this is likely due to the fact that a somewhat larger percentage of patients diagnosed with MCI have an underlying diagnosis that is not AD (Fagan et al., 2007; Shaw et al., 2009). Lower specificities are obtained when these ratios are used to differentiate AD from other dementia etiologies.

4.2. Prognosis

4.2.1. Progression from MCI/very mild dementia to AD.

Several studies have shown that either decreased CSF $\alpha B42$ or increased tau or phosphorylated forms of p-tau predicts progression from MCI/very mild dementia to AD. However, ratios of tau/$\alpha B42$ and p-tau/$\alpha B42$ may be more predictive than an individual marker. The relative risk of progression from MCI to AD was increased in patients who had high tau, p-tau, and low $\alpha B42$ at baseline with 90% sensitivity and 100% specificity in 1 study (Arai et al., 1997). An increased tau/$\alpha B42$ ratio was seen in 90% of individuals with MCI who later progressed compared with 10% of those who did not, in a large longitudinal study of MCI patients followed for 18 months (Riemenschneider et al., 2002). The combination of tau/$\alpha B42$ and the p-tau 181/$\alpha B42$ ratio, in a longitudinal study of almost 200 subjects with average follow-up of 4–6 years, strongly predicted progression of MCI to AD (Hansson et al., 2006). The utility of the AD CSF profile (defined by decreased $\alpha B42$ and increased tau) to detect progression from MCI to more advanced stages of AD was recently confirmed in a longitudinal study of 100 individuals with mild AD, 196 individuals with MCI, and 114 controls (Shaw et al., 2009) as well as in a very large multicenter study of over 1000 subjects in both Europe and the USA (Mattsson et al., 2009). In a smaller study, both markers predict not only conversion from MCI to AD, but also the rate of progression of cognitive decline as measured by the clinical dementia rating (CDR) sum of boxes and neuropsychological test scores (Snider et al., 2009). This is important as it is difficult to determine if an individual subject reaches a qualitative clinical cutoff defined as impaired enough to now have “AD” instead of “MCI.” However, being able to predict the rate of clinical progression should be quite useful in future clinical trials of disease-modifying drugs.

4.2.2. Progression from cognitive normality to MCI.

The increased ratio of tau/$\alpha B42$ and p-tau/$\alpha B42$ in normal individuals has been associated with an increased risk of conversion from normal to very mild dementia/ MCI in 3 recent studies. In a study by Fagan et al., approximately 70% of those with a high ratio, compared with only 10% of those with a normal ratio, converted from normal to MCI over a 3-year period (Fagan et al., 2007). Li et al. reported that over a follow-up of 42 months, all subjects who converted to MCI had elevated tau/$\alpha B42$ ratios, while no conversions occurred in the normal ratio group (Li et al., 2007). A larger study of 174 individuals that were enrolled initially as cognitively normal at the time of lumbar puncture has confirmed and extended these data (Craig-Schapiro et al., 2010). Although the current data come from relatively small sample sets, they suggest that over a 3- to 4-year period, the CSF tau/$\alpha B42$ ratio is a very good marker of predicting conversion from cognitively normal to MCI/very mild dementia. Low levels of CSF $\alpha B42$ alone in cognitively normal elderly is also predictive of conversion to MCI/very mild dementia; however, in the absence of an increase in tau or p-tau, the conversion time is significantly longer (3–8 years) (Gustafson et al., 2007; Skoog et al., 2003; Stromrud et al., 2007). This is likely due to the fact that amyloid deposition alone in the absence of significant degeneration as marked by increased tau indicates an earlier phase of the AD pathological process. It appears that the subgroup of normal elderly with a high ratio of tau/$\alpha B42$ have developed both $\alpha B$ deposition and neurodegeneration and represent individuals with preclinical AD.

5. Recommendation and future directions for use of CSF biomarkers in clinical trials

5.1. Clinical trials

For clinical trials in subjects with mild clinically defined cognitive abnormalities (MCI/very mild dementia), enrollment based on the use of CSF $\alpha B42$, tau, and p-tau would assist in several ways. First, by only enrolling individuals with a CSF $\alpha B42$ below a cutoff score for that laboratory that indicates a greater than 90% chance of having brain amyloid deposition, this will ensure that almost all of the subjects being enrolled have AD and will exclude subjects that lack amyloid deposition. Second, by using cutoff scores for the tau/$\alpha B42$ ratio or at minimum $\alpha B42$ as entry criteria, it will ensure that most individuals enrolled will decline at an appreciable rate such that smaller number of subjects will be required per arm to determine whether there are significant effects of the treatment. Third, assessment of these markers over time may allow one to visualize a therapeutic
effect (e.g., decreased tau) that might precede and ultimately correlate with a clinically beneficial effect such as reduced neurodegeneration. Although the data are less extensive in cognitively normal subjects, they support enrollment of normal subjects with a high CSF tau/Ab42 ratio into clinical trials with a goal of reducing conversion from normal to MCI/very mild dementia or in slowing cognitive decline as measured by neuropsychological test score performance. In untreated persons, such a biomarker result would predict conversion rates from CDR 0 to CDR > 0 of approximately 50% over 3 years.

5.2. Future directions

1. Larger numbers of cognitively normal people should be followed over time to validate the usefulness of CSF Ab42, tau, p-tau, and the ratios as predictors of cognitive decline.
2. Especially for prognosis in both normal and those with MCI/very mild dementia, it will be useful to validate the CSF markers in larger, epidemiologic-based sample sets. In addition, it will be informative to combine both fluid and imaging biomarkers to determine the most useful combinations of tests for diagnosis and prognosis.
3. From a technical standpoint, for a clinical trial to be conducted now using CSF biomarkers, it would be recommended to assess CSF markers in a single laboratory due to interlab variability. Assessment of these markers in multiple sites will require interlab standardization.
4. Other CSF biomarkers have been identified that may assist with diagnosis and prognosis. It will be important to determine if they provide added value to the current markers in appropriately assessed populations. Table 1.

Disclosure statement

David Holtzman is on the scientific advisory board of Satori, En Vivo, and C2N Diagnostics. He has consulted for Pfizer, Bristol-Myers Squibb, and Innogenetics. His laboratory receives research grants from Eli Lilly, Pfizer, Astra-Zeneca, and C2N Diagnostics.

Acknowledgements

This work was supported by NIH grants AG03991 and AG026276.

References


Blood-based biomarkers for Alzheimer’s disease: plasma Aβ40 and Aβ42, and genetic variants

Richard Mayeux*, Nicole Schupf

Taub Institute for Research on Alzheimer’s Disease and the Aging Brain, the Gertrude H. Sergievsky Center, and the Departments of Neurology, Psychiatry, and Epidemiology, Columbia University, New York, NY, USA

Abstract

Identifying a biomarker for Alzheimer’s disease that can be obtained from a blood sample has been a goal of researchers for many years. Over the past few years a number of investigators have studied several plasma biomarkers but most frequently plasma amyloid beta (Aβ)40 and Aβ42 while others have explored the use of genetic variants as biomarkers for diagnosis or risk. This review considers the cross-sectional and longitudinal data regarding plasma Aβ40 and Aβ42 as diagnostic biomarkers as well as risk biomarkers. Review of recent genome-wide association studies indicates as many as 10 genetic variants have been associated with susceptibility to Alzheimer’s disease (AD). Further analysis suggests that these factors have modest effects on risk and are thus not helpful, as yet, in the diagnosis of disease. Until the function of these genes is understood, their role in risk and diagnosis will remain uncertain. Thus, there are several types of peripheral biomarkers under investigation, but more work is required before they can be deemed clinically useful.

© 2011 Elsevier Inc. All rights reserved.

Keywords: Plasma amyloid β; Cross-sectional study; Prospective study; Genome-wide association studies

1. Introduction

Alzheimer’s disease (AD) is among the most frequently encountered diseases in aging societies with an estimated 5 million people in the United States and 17 million people worldwide suffering from the disease. It is expected that these numbers will quadruple by the year 2040, by which 1 out of 45 Americans will be affected, leading to a considerable public health burden.

To date, there are no definitive diagnostic tests or biological risk markers of the disease. The diagnosis of AD during life is based on clinical examination using the criteria of the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer’s Disease and Related Disorders Association (ADRDA) Work Group (McKhann et al., 1984). Although these criteria have been found to be reliable and valid any measure that would increase diagnostic sensitivity and specificity would be highly valuable for improving early detection and intervention.

2. Plasma Aβ40 and Aβ42

There have been approximately 26 investigations assessing plasma amyloid beta (Aβ)40 and Aβ42 as a diagnostic or as a biological risk factor (Tables 1–3). Studies of high risk populations (Table 1) have been consistent in showing elevated plasma Aβ42 levels in individuals from families multiply-affected by AD (early onset and late onset). Scheuener et al. (1996) first reported elevated levels among symptomatic carriers of presenilin mutations. Recently Ringman et al. (2008) found that Aβ42 and Aβ42/Aβ40 ratio levels were elevated in unaffected familial AD mutation carriers compared with unaffected individuals with familial AD without mutations. However, Aβ42 levels were lower in mutation carriers with incipient AD characterized as having a clinical dementia rating (CDR) = 0.5 (Hughes et al., 1982), supporting the hypothesis that Aβ42 decreases prior to overt disease. First degree relatives of patients with...
Table 1
Studies of plasma Aβ40 and Aβ42 in high risk populations: Down syndrome, familial Alzheimer’s disease, and first degree relatives of patients with Alzheimer’s disease

<table>
<thead>
<tr>
<th>Year</th>
<th>Study</th>
<th>Outcome</th>
<th>n (Ca/Co)</th>
<th>Result</th>
<th>Interval</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>Scheuner et al., 1996</td>
<td>FAD</td>
<td>12 APP/31 control; 9 PS1/4 PS2/14 control</td>
<td>Plasma Aβ42 higher in FAD</td>
<td>Cross-sectional</td>
<td>FAD with multiple affected members</td>
</tr>
<tr>
<td>2007</td>
<td>Schupf et al., 2007</td>
<td>AD in Down syndrome</td>
<td>44 inc/130 control</td>
<td>Elevated Aβ42, not Aβ40</td>
<td>3.9 y</td>
<td>Prospective</td>
</tr>
<tr>
<td>2007</td>
<td>Ertekin-Taner et al., 2008</td>
<td>LOAD first degree relatives</td>
<td>217 relatives/103 controls</td>
<td>High Aβ42</td>
<td>Cross-sectional</td>
<td>Families with multiple affected members, no known mutations</td>
</tr>
<tr>
<td>2008</td>
<td>Ringman et al., 2008</td>
<td>FAD MUC at risk in contrast to FAD NC</td>
<td>9 MUC/8 NC</td>
<td>Aβ42 and Aβ42/Aβ40 levels higher in MUC vs. NC: Aβ42 lower in MUC CDR = 0.5</td>
<td>Cross-sectional</td>
<td>Cross-sectional. Preclinical stage: Aβ42 levels decrease with onset of symptoms prior to overt disease</td>
</tr>
</tbody>
</table>

Key: Aβ, amyloid beta; AD, Alzheimer’s disease; APP, amyloid precursor protein; CDR, clinical dementia rating; FAD, familial Alzheimer’s disease; LOAD, late-onset Alzheimer’s disease; MUC, mutation carrier; NC, noncarrier.

Table 2
Cross-sectional studies of plasma Aβ40 and Aβ42

<table>
<thead>
<tr>
<th>Year</th>
<th>Study</th>
<th>Outcome</th>
<th>N (Ca/Co)</th>
<th>Result</th>
<th>Interval</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>Fukumoto et al., 2003</td>
<td>AD, MCI</td>
<td>146 prev AD, 37 MCI, 96 PD, 92 Co</td>
<td>No difference Ca/Co</td>
<td>Cross-sectional</td>
<td>Plasma Aβ related to age, not disease</td>
</tr>
<tr>
<td>2004</td>
<td>Assini et al., 2004</td>
<td>Amnestic MCI</td>
<td>88 MCI/72 Co</td>
<td>High Aβ42 vs. male and female Co; no difference in Aβ40</td>
<td>Cross-sectional</td>
<td>Preclinical stage. Higher incidence of AD in women</td>
</tr>
<tr>
<td>2005</td>
<td>Irizarry et al., 2005</td>
<td>AD, MCI, WMH, CAA</td>
<td>54 Ca/42 Co</td>
<td>Aβ 40 associated with WMH in AD/MCI/CAA</td>
<td>Cross-sectional</td>
<td>Prevalent AD</td>
</tr>
<tr>
<td>2005</td>
<td>Sobów et al., 2005</td>
<td>AD/MCI</td>
<td>54 AD/39 MCI/35 Co</td>
<td>High Aβ42</td>
<td>—</td>
<td>Highest in MCI</td>
</tr>
<tr>
<td>2006</td>
<td>Pesaresi et al., 2006</td>
<td>AD</td>
<td>146 Cs/89 MCI/89 NC</td>
<td>Lower Aβ42 in AD, not MCI</td>
<td>Repeat 18 mo in 20 Ss</td>
<td>Cross-sectional and repeated; no change in levels over follow-up</td>
</tr>
<tr>
<td>2010</td>
<td>Lewczuk et al., 2010</td>
<td>AD</td>
<td>243 controls/577 AD</td>
<td>Lower Aβ42/Aβ 40 in cases</td>
<td>Cross-sectional</td>
<td>Used CSF to stratify dementia</td>
</tr>
</tbody>
</table>

Results of several cross-sectional studies using a case-control design to determine whether or not plasma Aβ40 and Aβ42 can be useful as a diagnostic biomarker.

Key: Aβ, amyloid beta; AD, Alzheimer’s disease; CSF, cerebrospinal fluid; MCI, mild cognitive impairment; PD, Parkinson’s disease.
Table 3
Prospective studies of plasma Aβ40 and Aβ42

<table>
<thead>
<tr>
<th>Year</th>
<th>Study</th>
<th>Outcome</th>
<th>n (Ca/Co)</th>
<th>Result</th>
<th>Interval</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>Mayeux et al., 2003</td>
<td>AD</td>
<td>86 Inc/365: repeat in 307</td>
<td>High baseline Aβ42, not Aβ40, and declining Aβ42</td>
<td>3 y</td>
<td>Repeat measures: onset associated with decline in Aβ42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>34 Nondemented (60–74 y)</td>
<td>High initial Aβ42/decline Aβ 42 and decline in MMSE</td>
<td>4.2 y</td>
<td>Repeat measures. All cognitively intact</td>
</tr>
<tr>
<td>2005</td>
<td>Pomara et al., 2005</td>
<td>MMSE; SRT total, delayed recall</td>
<td></td>
<td>Aβ42 increased with cognitive decline, NC → MCI and NC → AD</td>
<td>2.5 y</td>
<td>Repeat measures VITA. ≥75 y, Vienna</td>
</tr>
<tr>
<td>2008</td>
<td>Blasko et al., 2008</td>
<td>NC → MCI → AD</td>
<td>585 Nondemented, 487 followed; 90 AD/98 all demented</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>van Oijen et al., 2006</td>
<td>AD</td>
<td>392 Inc/1373 Co</td>
<td>Low Aβ42/Aβ40 predicts onset</td>
<td>8.6 y</td>
<td>Case-cohort</td>
</tr>
<tr>
<td>2007</td>
<td>Graff-Radford et al., 2007</td>
<td>MCI/AD</td>
<td>53 Inc/510 NC; 379 with 2 evaluations</td>
<td>Low Aβ42/Aβ40 predicts onset and cognitive decline</td>
<td>12 mo; 3–7 y</td>
<td>Combined MCI/AD outcome (17 AD and 36 MCI)</td>
</tr>
<tr>
<td>2008</td>
<td>Sundelof et al., 2008</td>
<td>AD/VAD/all dementia</td>
<td>44 Ca/606 Co in 77-y cohort</td>
<td>Low Aβ40 and low Aβ42 (marginal)</td>
<td>5.3 y</td>
<td>Repeat measures (not analyzed). Association only in those 77 y</td>
</tr>
<tr>
<td>2010</td>
<td>Hansson et al., 2010</td>
<td>MCI → AD</td>
<td>48 Ca/51 stable MCI/38 Co</td>
<td>No difference</td>
<td>4–7 y</td>
<td>Cohort 1 low Aβ42/Aβ40 in those with CSF markers of AD</td>
</tr>
<tr>
<td>2006</td>
<td>Hansson</td>
<td>MCI AD</td>
<td>15 Ca/95 MCI/41 Co</td>
<td>No difference</td>
<td>2–4 y</td>
<td>Cohort 2</td>
</tr>
<tr>
<td>2008</td>
<td>Lopez et al., 2008</td>
<td>AD</td>
<td>88 Inc/69 MCI/117 MCI</td>
<td>High Aβ40 and Aβ42</td>
<td></td>
<td>CHS study</td>
</tr>
<tr>
<td>2008</td>
<td>Locascio et al., 2008</td>
<td>Cognitive (BDS) and ADL scores</td>
<td>122 with AD from memory clinic</td>
<td>Earliest: low Aβ40 and Aβ42 with more rapid decline</td>
<td>0.2–12 y</td>
<td>Retrospective and prospective. Duration of AD at base = 3.2 y (0.2–12 y)</td>
</tr>
<tr>
<td>2008</td>
<td>Schupf et al., 2008</td>
<td>AD</td>
<td>104 Inc/1021 Co</td>
<td>High baseline Aβ42, decline in Aβ42 or Aβ42/Aβ40</td>
<td>4.6 y</td>
<td>Repeat measures; onset associated with decline in Aβ42 and Aβ42/Aβ40 ratio</td>
</tr>
<tr>
<td>2009</td>
<td>Okereke et al., 2009</td>
<td>Cognitive (TICS) scores</td>
<td>481 Nondemented women</td>
<td>Low Aβ42/Aβ40 in midlife and decreases over next 10 y</td>
<td>10 y</td>
<td>Telephone-based interviews</td>
</tr>
<tr>
<td>2009</td>
<td>Lambert et al., 2009</td>
<td>Dementia (AD, VAD, other types)</td>
<td>985 Nondemented 233 Demented</td>
<td>High Aβ42/Aβ40 associated with lower risk</td>
<td>4 y</td>
<td>Prospective, clinical assessment</td>
</tr>
</tbody>
</table>

This table depicts the results of several prospective studies designed to examine the ability of plasma Aβ40 and Aβ42 to predict transition from no dementia to mild cognitive impairment or dementia. Key: Aβ, amyloid beta; AD, Alzheimer’s disease; ADL, activities of daily living; CSF, cerebrospinal fluid; MCI, mild cognitive impairment; MMSE, Mini Mental State Examination.
late onset Alzheimer’s disease (LOAD) without known mutations or genetic variants have also been found to have increased plasma Aβ42 (Ertekin-Taner et al., 2008). Adults with Down syndrome who produce more plasma Aβ42 due to trisomy and triplication of the gene for amyloid precursor protein (APP) also show high plasma levels of Aβ42 before onset of dementia (Schunk et al., 2007). Among studies of sporadic LOAD, there has been little consistency in the findings.

Cross section studies (Table 2) comparing patients with clinically diagnosed LOAD with controls showed no differences in either plasma Aβ40 or Aβ42. However, 2 studies that included patients with mild cognitive impairment (MCI) found high plasma Aβ42, but not Aβ40. Prospective studies (Table 3) have generally shown that higher plasma Aβ42 levels are associated with increased risk of LOAD or cognitive decline. Most of the studies cited in Table 3 also indicate that the elevation in levels of plasma Aβ42 are present before or just at the onset of clinically diagnosed disease. Others have included repeated measurements of plasma Aβ42, finding that levels decreased as disease progressed. Not all studies agree. At least 2 prospective studies examining plasma Aβ42 showed no difference in patients with MCI who progressed to LOAD (Hansson et al., 2006, 2007). Two others showed that low plasma Aβ42 or a low ratio of plasma Aβ42/Aβ40 was associated with more rapid cognitive decline or with disease (Graf-Radford et al., 2007; Locascio et al., 2008). These findings are consistent with our hypothesis that high plasma Aβ42 is an antecedent indicator of risk for LOAD and declines with onset and progression of dementia. One explanation for the differences in the results is the time of sampling. A decline in plasma Aβ42 may already have begun when patients develop signs of MCI. Alternatively, the methods used to analyze plasma Aβ42 and Aβ40 may explain some of the differences in results, though at least 2 studies used the same methods for measurement of plasma Aβ with different results (Mayeux et al., 2003; van Oijen et al., 2006). Despite the variability in levels before disease onset, there is a general consensus that plasma Aβ42 levels, and perhaps plasma Aβ40 levels as well, decrease with disease progression.

### 3. Aβ42 as a risk biomarker for LOAD

Compared with asymptomatic individuals with low plasma Aβ42 levels at initial assessment, Schupf et al. (2008) reported that those with high Aβ42 levels had more than a 3-fold increased risk of developing LOAD over an average of 4 and a half years. At the follow-up assessment when blood sampling was repeated, a decrease in plasma Aβ42 levels, but not Aβ40 levels, was related to the development of LOAD. The likelihood of having converted to LOAD 18 to 24 months before the second blood draw was 3 times higher when plasma Aβ42 levels had decreased by more than a half a standard deviation or when the plasma Aβ42/Aβ40 ratio decreased by more than a half a standard deviation. Thus, over time, decreasing levels of plasma Aβ42 or a decline in the Aβ42/Aβ40 ratio are sensitive indicators of recent conversion to LOAD. The authors have posited that the decline in plasma Aβ42 reflects compartmentalization and deposition of Aβ peptides in brain.

These results confirmed and extended previous findings among nondemented individuals who subsequently developed LOAD (Mayeux et al., 1999, 2003; Schupf et al., 2008) and are consistent with studies in women with MCI and among asymptomatic first-degree relatives of patients at high risk of developing LOAD (Assini et al., 2004; Ertekin-Taner et al., 2008). However, these results are in distinct contrast with others that report (1) no relation between plasma Aβ peptide levels and risk of LOAD (Fukumoto et al., 2003), (2) an association between low plasma Aβ40 and LOAD (Sundelöf et al., 2008; van Oijen et al., 2006), or (3) a relation between a low plasma Aβ42/Aβ40 ratio and subsequent cognitive impairment or LOAD (Graff-Radford et al., 2003, 2007). A number of factors may account for these inconsistencies. One important factor is likely to be the timing of sample collection in relation to the preclinical period or to stage of disease onset and progression. Few studies have examined risk associated with change in plasma Aβ peptide levels or change in Aβ42/Aβ40 ratio over time. In the study by Schupf et al. (2008), conversion to LOAD was strongly related to a decline in Aβ42 levels and in the Aβ42/Aβ40 ratio. Similarly, in a study of healthy nondemented elderly individuals, higher initial Aβ42 levels and greater reductions in Aβ42 levels over an approximate 4-year period were associated with greater cognitive decline (Pomara et al., 2005). In the cerebrospinal fluid (CSF), low levels of Aβ42 and Aβ42/Aβ40 ratios in patients with MCI are associated with higher brain amyloid load (Fagan et al., 2006, 2007) and predict conversion to LOAD (Bliennow and Hampel, 2003; Bliennow and Vanmechelen, 2003; Hansson et al., 2006). This suggests that a decline in Aβ42 levels and in Aβ42/Aβ40 ratios can herald the onset of LOAD, possibly reflecting sequestration of Aβ42 in senile plaques or the formation of semisoluble oligomers (Lesné et al., 2006, 2008).

### 4. Other plasma biomarkers

Offspring of patients with AD are more likely to have lower mean plasma apolipoprotein E (APOE) levels compared with offspring of controls. Individuals with 1 or more APOEε4 alleles have lower APOE levels compared with those with other APOE alleles (van Vliet et al., 2009). Blood levels of progranulin are useful as an indicator of progranulin-related frontotemporal lobar degeneration, both null mutations and missense mutations as well as blood levels of progranulin have also been observed in patients with Alzheimer’s disease (Sleegers et al., 2010). Increased
plasma concentration of clusterin was predictive of greater fibrillar amyloid-beta burden in the medial temporal lobe at autopsy but has not been investigated as a diagnostic or risk biomarker (Thambisetty et al., 2010, 2011).

5. Conclusions and recommendation

There is insufficient evidence to permit a conclusion regarding the use of plasma Aβ40, Aβ42, or the ratio of Aβ42/Aβ40 in the diagnosis or assessment of risk of AD using cross-sectional or single measurements. There is suggestive evidence that changes in plasma Aβ40, Aβ42, or the ratio of Aβ42/Aβ40 may be associated with—and therefore useful in identifying—individuals at risk for developing AD. Standardization of the measurement Aβ40 and Aβ42 is required to determine whether plasma measurement will be useful in risk assessment of Alzheimer’s disease.

There are insufficient data to permit a conclusion regarding the use of plasma or serum levels of APOE, clusterin, or progranulin in the assessment of risk or as a diagnostic biomarker in Alzheimer’s disease.

6. Genetic variants as biomarkers of AD and LOAD

Predisposing genetic variants could, in the future, be used to predict who will develop dementia. Individuals genetically predisposed to dementia may benefit from therapeutic interventions in the early stages of the disease. Early intervention could significantly prevent or delay the onset, which in turn would improve quality of life of the patient and their relatives and would significantly reduce the public health burden.

The majority of familial early onset Alzheimer’s disease (EOAD) is caused by mutations in APP, PSEN1, and PSEN2. These genes have almost complete penetrance (> 85%) and a clear autosomal dominant pattern of inheritance. Thus, the presence of a mutation is virtually synonymous with disease. Complicating the wide-spread use of genetic testing for familial early onset Alzheimer’s disease is the relative rarity of cases and the fact that there are multiple mutations in these genes producing the same phenotype. Thus, genetic testing requires full sequencing of the gene.

Current knowledge suggests that a variety of mechanisms underlie the various neuropathological and clinical changes, and that these have different genetic and environmental components. Thus, it is likely that late-onset cognitive impairment is a complex genetic disorder characterized by an interaction of multiple genes and the environment leading to genetic variants that have incomplete penetrance and a low-magnitude associated risk. Consistent with this notion is the fact that to date only APOE has been firmly identified as a genetic risk factor, although segregation analyses conducted in families of patients with LOAD support the presence of additional genetic variants (Daw et al., 2000). With a population attributable risk that is estimated at 20%–50%, the APOEε4 allele increases risk of cognitive impairment, LOAD, and age-of onset of cognitive impairment in a dose-dependent fashion: 1 ε4 allele is associated with a 2- to 3-fold increased risk, having 2 copies is associated with a 5- to 10-fold increase. Similar effect sizes have been observed for progression of cognitive impairment to dementia.

SORL1 was identified as a candidate gene in which 2 haplotypes in the 3’ and 5’ regions of the gene (Rogaeva et al., 2007) found to be associated with LOAD, with effect sizes ranging from odds ratios of 1.4 to 2.2. Subsequent studies have confirmed the association (Reitz et al., 2011). Variants in APOE and SORL1 lack sufficient sensitivity and specificity to be used as diagnostics. Variants in both increase risk of cognitive impairment in a non-Mendelian fashion, are not fully penetrant, and are neither necessary nor sufficient by themselves to cause impairment. The same is likely to be true for the remaining yet-to-be-identified genetic factors associated with cognitive decline.

There have been approximately 16 genome-wide association studies (GWAS) published to date Table 4. Most of these studies have included unrelated patients with LOAD and controls, but there are 3 studies that used family data (Bertram et al., 2008; Poduslo et al., 2009; Wijsman et al., 2011). Most genome-wide association studies have confirmed the APOE association with LOAD. Two large studies have identified variants in clusterin (CLU), phosphatidylinositol binding clathrin assembly protein (PICALM), and complement component (3b/4b) receptor 1 (CRI) as being associated with LOAD risk (Harold et al., 2009; Lambert et al., 2009). CLU, or apolipoprotein J (APOJ), is a lipoprotein found to be part of amyloid plaques. CLU specifically binds soluble Aβ in cerebrospinal fluid to form complexes able to cross the blood-brain barrier. It has also been noted that reduced levels of APOE and increased levels of CLU are correlated with the number of ε4 alleles, suggesting a compensatory induction of CLU in the brain of LOAD individuals with the APOE ε4 allele and low brain levels of APOE. CRI encodes the complement C3b protein, which is likely to contribute to Aβ clearance (Wyss-Coray et al., 2002). PICALM is involved in clathrin-mediated endocytosis (CME), an essential step in the intracellular trafficking of proteins and lipids such as nutrients, growth factors, and neurotransmitters. Of relevance to LOAD, PICALM also appears to be involved in directing the trafficking of vesicle-associated membrane protein 2 (VAMP2), which is a soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) protein that has a prominent role in the fusion of synaptic vesicles to the presynaptic membrane during neurotransmitter release, a process that is crucial to neuronal function. The gene encoding bridging integrator 1 (BIN1) was also noted initially at a lower threshold of significance (Harold et al., 2009) with stronger results in a recent replication analysis (Seshadri et al., 2010). BIN1, a gene expressed in the central nervous sys-
Table 4
Results of recent genome-wide association studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Population</th>
<th>Platform</th>
<th>SNPs, n</th>
<th>AD cases</th>
<th>Controls</th>
<th>Identified genes/regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reiman et al., 2007</td>
<td>2007</td>
<td>USA, Netherlands</td>
<td>Affymetrix (500K)</td>
<td>312,316</td>
<td>446, 415</td>
<td>290, 260</td>
<td>APOE, GAB2</td>
</tr>
<tr>
<td>Li et al., 2008</td>
<td>2008</td>
<td>Canada</td>
<td>Affymetrix (500K)</td>
<td>469,438</td>
<td>753, 418</td>
<td>736, 249</td>
<td>APOE, TOMM40, APOC1, GOLPH2</td>
</tr>
<tr>
<td>Poduslo et al., 2009</td>
<td>2008</td>
<td>USA</td>
<td>Affymetrix (500K)</td>
<td>489,218</td>
<td>9, 199</td>
<td>10, 225</td>
<td>APOE, TRPC4AP</td>
</tr>
<tr>
<td>Bertram et al., 2008</td>
<td>2008</td>
<td>USA (NIMH)</td>
<td>Affymetrix (500K)</td>
<td>484,522</td>
<td>941, 1767</td>
<td>404, 838</td>
<td>APOE, APOC1, ACE, CHRNA4, GAB2, TF, GWAS14q31.2</td>
</tr>
<tr>
<td>Beecham et al., 2009</td>
<td>2009</td>
<td>USA (CAP)</td>
<td>Illumina (550K)</td>
<td>532,000</td>
<td>492, 238</td>
<td>496, 220</td>
<td>APOE, PCDH11X</td>
</tr>
<tr>
<td>Carrasquillo et al., 2009</td>
<td>2009</td>
<td>USA (Mayo)</td>
<td>Illumina (300K)</td>
<td>313,504</td>
<td>844, 1547</td>
<td>1255, 1209</td>
<td>APOE, PCDH11X</td>
</tr>
<tr>
<td>Potkin et al., 2009</td>
<td>2009</td>
<td>USA (ADNI)</td>
<td>Illumina (610K)</td>
<td>516,645</td>
<td>172,</td>
<td>209,</td>
<td>APOE, TOMM40, EFNA5, CAND1, MAGI2, ARSB, PRUNE2</td>
</tr>
<tr>
<td>Lambert et al., 2009</td>
<td>2009</td>
<td>Europe (EADI1)</td>
<td>Illumina (610K)</td>
<td>537,029</td>
<td>2032, 3978</td>
<td>5328, 3297</td>
<td>APOE, CLU</td>
</tr>
<tr>
<td>Harold et al., 2009</td>
<td>2009</td>
<td>Europe and USA (GERAD1)</td>
<td>Illumina (various)</td>
<td>529,205</td>
<td>3941, 2023</td>
<td>7848, 2340</td>
<td>APOE, CLU, PICALM</td>
</tr>
<tr>
<td>Feulner et al., 2010</td>
<td>2010</td>
<td>Europe and USA (CHARGE, EADI1, GERAD1)</td>
<td>Illumina (550K)</td>
<td>555,000</td>
<td>491,</td>
<td>479,</td>
<td>APOE, TOMM40, MAPT, SORL, APOE, CLU, PICALM, CR1, EXOC3L2, BIN1</td>
</tr>
<tr>
<td>Seshadri et al., 2010</td>
<td>2010</td>
<td>USA (Alzheimer’s Disease Genetics Consortium)</td>
<td>Affymetrix and Illumina</td>
<td>2,540,000 (imputed)</td>
<td>3006, 6505</td>
<td>22,604, 13,532</td>
<td>MS4A, ABCA7, CD33, EPHA1, CD2AP</td>
</tr>
<tr>
<td>Naj et al., 2011</td>
<td>2011</td>
<td>USA</td>
<td>Affymetrix and Illumina</td>
<td>2,324,889 (imputed)</td>
<td>8039, 3531</td>
<td>7366, 3565</td>
<td>MS4A, ABCA7, CD33, EPHA1, CD2AP</td>
</tr>
<tr>
<td>Hollingworth et al., 2011</td>
<td>2011</td>
<td>Europe and USA (CHARGE, EADI1, GERAD1)</td>
<td>Affymetrix and Illumina</td>
<td>496,763</td>
<td>6688, 8286</td>
<td>13,685, 21,258</td>
<td>MS4A, ABCA7, CD33, EPHA1, CD2AP</td>
</tr>
</tbody>
</table>

Key: AD, Alzheimer’s disease; GWAS, genome-wide association study; NIMH, National Institute of Mental Health; SNPs, single nucleotide polymorphisms.
tem, is reported to activate a caspase-independent apoptotic process.

Genome-wide significant results at \textit{MS4A4A} on chromosome 11, \textit{CD2AP} on chromosome 6, \textit{EPHA1} on chromosome 7, and \textit{CD33} on chromosome 19 were observed in independent meta-analyses by the Alzheimer’s Disease Genetic Consortium and another European-American Consortium (Hollingworth et al., 2011; Naj et al., 2011). The true genetic effector for these genes has not been identified and their roles in the pathogenesis are unknown.

Can any of these genes be used as biomarkers for diagnosis or as measures of disease risk? At this point, only mutations in genes associated with early onset familial Alzheimer’s disease could be considered diagnostic. These mutations are virtually 100% penetrant, and therefore those who carry a mutation will develop disease. Of interest, some patients with these mutations also have elevated plasma levels of A\beta42. Genetic variants associated with late-onset Alzheimer’s disease are not useful as diagnostic measures. For example, APOE-ε4 is strongly associated with disease risk, but it has been shown to provide little help in the diagnosis. While there are many candidate genes in which variants are, or may prove to be, associated with increased risk of subsequent disease, none are currently used as biomarkers of risk.

7. Conclusions and recommendation

There is sufficient direct evidence to support the utility of pathogenic mutations identified in presenilin1, presenilin2, and APP in the diagnosis of Alzheimer’s disease. There is sufficient evidence to support the use of variation at the APOE locus in the assessment of risk of Alzheimer’s disease, but insufficient evidence to permit a conclusion regarding the use of variation at the APOE locus for diagnosis. There is insufficient evidence to permit a conclusion regarding the use of genetic variation in any other gene identified in the assessment of risk or in the diagnosis of Alzheimer’s disease.

Disclosure statement

The authors disclose no conflicts of interest.

Acknowledgements

This work was supported by federal grants from the National Institute on Aging of the National Institutes of Health (P01AG07232, R37AG15473, P50 AG08702) and by grants from the Alzheimer Association, the Blanchette Hooker Rockefeller Fund, the Robertson Gift from the Banbury Fund, and the Merrill Lynch Foundation.

References


Imaging

Amyloid imaging as a biomarker for cerebral β-amyloidosis and risk prediction for Alzheimer dementia

William E. Klunk*

University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Abstract

Since the introduction of amyloid imaging nearly 10 years ago, this technique has gained widespread use and acceptance. More recently, published reports have begun to appear in which amyloid imaging is used to detect the effects of anti-amyloid therapies. This review will consider the issues involved in the use of amyloid imaging in the development and evaluation of drugs for the treatment of Alzheimer’s disease. Current evidence regarding the postmortem correlates of in vivo amyloid imaging data are considered. The application of amyloid imaging to screening subjects for trials and use as an outcome measure is discussed in light of longitudinal changes in the in vivo amyloid signal. While the bulk of this review is directed at symptomatic patients with dementia, consideration is given to the use of amyloid imaging in nondemented subjects as well. Similarities and differences of cerebral amyloid assessment by amyloid imaging and cerebrospinal fluid (CSF) measurements are delineated and an agenda for further research to improve the applicability of amyloid positron emission tomography (PET) to clinical trials is proposed.

© 2011 Published by Elsevier Inc.

Keywords: Amyloid imaging; Drug development; Alzheimer’s disease

1. Introduction

This review will focus on the application of amyloid imaging as a tool in the development and evaluation of drugs for the treatment of Alzheimer’s disease. To avoid confusion, we will use the term “pathophysiology of Alzheimer’s disease” when referring to the full spectrum of underlying biological abnormalities that begin before symptoms and extend into the clinically evident phases. The main focus will be on the use of amyloid imaging in symptomatic Alzheimer’s disease (which we will refer to as Alzheimer’s disease (AD) dementia), but attention also will be given to prodromal and preclinical manifestations, including mild cognitive impairment (MCI)—particularly in the context of predicting progression to clinically “probable AD dementia” (McKhann et al., 1984). From the outset we must be careful to stress that amyloid imaging, as the name implies, is intended to detect brain pathophysiology, but not to make a clinical diagnosis. It can be an important adjunct to a clinical evaluation in making more accurate clinical diagnoses—much as postmortem pathology can ultimately confirm the presence of pathologically proven AD dementia if coupled to a typical clinical history. Used in isolation, amyloid imaging cannot diagnose AD, MCI, or detect normal or abnormal aging.

Amyloid imaging agents typically detect beta-sheet rich fibrillar deposits of amyloid β-protein (Aβ) in plaques and cerebrovascular amyloid (CAA). For example, tracer binding to plaques and CAA in postmortem brain tissue can be abolished by destruction of the beta-sheet fibrillar structure by formic acid treatment (Ikonomovic et al., 2008). Fibrillar Aβ is a major component of compact/cored plaques, whether or not they are neuritic. In addition, fibrillar Aβ can be found in varying degrees in plaques that have been loosely characterized as “diffuse,” and these diffuse plaques can be detected to some extent by amyloid imaging agents (Burack et al., 2010; Ikonomovic et al., 2008; Lockhart et
al., 2007). However, diffuse plaques that are sometimes described as “amorphous” or “fleecy” such as those found in the cerebellum contain little beta-sheet structure and are not detectable by typical amyloid imaging agents (Ikonomovic et al., 2008). The fibrillar Aβ deposits detected by prototypical amyloid imaging agents may be unique to human brain, as even compact-appearing plaques in squirrel monkey and amyloid precursor protein (APP) transgenic mice do not produce significant binding (Klunk et al., 2005a; Rosen et al., 2009; Toyama et al., 2005).

With regard to the amyloid imaging agents, the focus of this review will be on the most widely evaluated positron emission tomography (PET) tracer, Pittsburgh Compound B (PiB) (Klunk et al., 2004). At the time of this writing, there have been single, small published studies using each of the F-18-labeled tracers, [F-18]florbetaben (18F-BAY94-9172 or AV-1; Rowe et al., 2008), [F-18]florbetapir (AV-45; Wong et al., 2010) and [F-18]flutemetamol (3’F-PiB or GE-067; Nelissen et al., 2009) in AD dementia patients. Another F-18-labeled agent has been used in preclinical studies, but no human data have been published (Jureus et al., 2010). There is insufficient published evidence available to evaluate any of these F-18 tracers for their potential use in drug trials at present. While the findings discussed below for PiB PET may ultimately be found to extend to these F-18-labeled tracers as well, this cannot be assumed until appropriate studies have been repeated with each individual tracer or until pharmacological equivalency to PiB has been established by direct comparison in the same subjects. This becomes especially true in studies aimed at detecting the first signs of in vivo amyloid deposition in cognitively normal subjects, when the lower signal-to-noise ratio of the F-18-labeled tracers may become important. Another F-18-labeled tracer, [F-18]FD-DNP, has fundamentally different properties from all of these tracers and will not be discussed here (Small et al., 2006; Thompson et al., 2009; Tolboom et al., 2009a, 2009b, 2009d, 2010).

In this review, PiB imaging will be discussed in the context of how it might be used in therapeutic, clinical trials. For example, it will be assumed that early trials will be conducted in highly specialized referral centers, and although the population of AD dementia subjects in these centers is not necessarily representative of a general AD dementia population as might be captured in an epidemiological study, it is likely to be the target population for AD dementia drug trials in the near future. Essentially all of the PiB PET studies discussed below have been performed in such “specialized center” settings.

Although this review focuses on amyloid imaging in isolation, it is clear that this technique cannot fulfill all of the biomarker needs of any clinical trial and will need to be considered as a part of a broader biomarker arsenal. Other promising biomarkers are discussed separately in this position statement. Furthermore, there is significant overlap between the utility of amyloid imaging and measurement of cerebrospinal fluid (CSF) Aβ42 as a screening tool (but not as a trial outcome measure), and so an attempt will be made to address the areas where these 2 biomarkers may be equivalent and areas where 1 technique may hold unique advantages.

This review is organized around the most likely applications of amyloid imaging to clinical trials, i.e., use as a screening tool or as an outcome measure, and in trials of AD dementia or predementia syndromes associated with the pathophysiology of Alzheimer’s disease. However, within this structure, discussion of the typical aspects of biomarker characterization/validation will be incorporated, including: (1) cross-sectional association between PiB retention and clinical diagnosis; (2) longitudinal change of PiB retention as a marker of AD dementia progression; (3) prediction of progression of prodromal and preclinical syndromes associated with the pathophysiology of Alzheimer disease; and (4) postmortem correlation/validation studies. Among these 4, the most basic and important characterization/validation of most biomarkers is correlation with postmortem assessment of pathology, so this will be discussed first.

2. Postmortem pathological characterization/validation of amyloid imaging

Amyloid imaging is foremost a method for the detection of Aβ pathology more than it is a surrogate of any other aspect of the pathophysiology of Alzheimer disease or the clinical manifestations of this disorder. However, from the outset, it is recognized that no in vivo measure of pathology is likely to be as sensitive as modern postmortem histological and biochemical measures of pathology detection. For comparison with in vivo amyloid imaging, it is important to choose the method of postmortem analysis that best reflects that in vivo target(s). In the case of in vivo amyloid imaging, the target is Aβ deposition—in all of its forms. This would include all forms of Aβ plaques (e.g., diffuse, cored, neuritic, etc.) as well as CAA. Of course, the targets do not include other common forms of pathology associated with AD dementia, in particular, tau pathology in the form of neurofibrillary tangles, dystrophic neurites, and neuropil threads. Thus any postmortem grading system or quantitative measure that includes tau pathology should not be a component of a postmortem validation of amyloid imaging. This would include the Braak and Braak staging system of neurofibrillary tangles (Braak and Braak, 1991) and the NIA-Reagan criteria (1997), because the latter incorporates Braak tangle staging in the determination of the “likelihood” of AD dementia. The optimal postmortem correlate for in vivo amyloid imaging may be a specific measure of total Aβ pathology by the use of sensitive and specific anti-Aβ antibodies. These antibodies can be applied in quantitative biochemical (e.g., enzyme-linked immunosorbent assays or ELISA) or immunohistochemical (IHC) analyses of Aβ load. Other biochemical methods to quantify Aβ
could apply as well, but these often require specialized equipment and expertise that may be less available.

The literature contains reports of 24 cases that had been studied with PiB PET prior to autopsy (n = 14) or after biopsy (n = 10). These studies, described in Table 1, meet our aforementioned methodological goals to varying degrees but they do provide important insights about the sensitivity and specificity of in vivo amyloid imaging for detecting the presence of postmortem Aβ pathology (Bacskaï et al., 2007; Burack et al., 2010; Cairns et al., 2009; Ikonomovic et al., 2008; Sojkova et al., 2011; Villemagne et al., 2009). First, in all of the 12 cases with positive PiB PET scans in vivo, postmortem analyses confirmed the presence of significant Aβ deposition (Table 1). Thus, the specificity of PiB PET in this small sample was 100%. This is not surprising, given the discussion above that the postmortem measures are expected to be more sensitive than PET measures. Also consistent with this relative sensitivity is the finding that 3/12 cases that were PiB-negative in vivo showed evidence of some Aβ pathology postmortem (Table 1). The Sojkova et al. (2011) Case-B is difficult to interpret: while there were moderate numbers of neuritic plaques by IHC in the precuneus showed no Aβ deposits (and there was no CAA). Thus, it is unclear if this case represents a mismatch between in vivo PiB PET and postmortem pathology. Two cases do appear to be clear mismatches. Cairns et al. (2009) have reported a negative PiB scan in the presence of biochemically and immunohistochemically detectable Aβ at levels expected to be detectable in vivo. In a biopsy study, Leinonen et al. reported “Case #6” with high numbers of plaques by IHC (although sparse neuritic plaques) but a negative PiB scan (Leinonen et al., 2008). These may represent an example of nonfibrillar amyloid deposition (i.e., a type of diffuse plaques) or other alterations in the tertiary structure of Aβ, as has been re-

Table 1

<table>
<thead>
<tr>
<th>Clinical diagnosis</th>
<th>PiB status</th>
<th>Scan-PM interval</th>
<th>Autopsy</th>
<th>Aβ ELISA or IHC</th>
<th>Path diagnosis</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Braak stage</td>
<td>Diffuse plaques</td>
<td>Neuritic plaques</td>
<td>CAA</td>
</tr>
<tr>
<td>DLB</td>
<td>Pos</td>
<td>3 mo</td>
<td>IV</td>
<td>Moderate</td>
<td>Sparse</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DLB, possible AD</td>
</tr>
<tr>
<td>AD</td>
<td>Pos</td>
<td>10 mo</td>
<td>V/VI</td>
<td>Frequent</td>
<td>Frequent</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Definite AD</td>
</tr>
<tr>
<td>CJD</td>
<td>Neg</td>
<td>2 wk</td>
<td>NA</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CJD</td>
</tr>
<tr>
<td>CJD</td>
<td>Neg</td>
<td>4 wk</td>
<td>NA</td>
<td>Sparse, frontal only</td>
<td>None</td>
<td>Low-none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>Very mild AD</td>
<td>Neg</td>
<td>2.5 y</td>
<td>III</td>
<td>Moderate</td>
<td>Mild</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Possible AD</td>
</tr>
<tr>
<td>PDD</td>
<td>Pos</td>
<td>&lt; 15 mo</td>
<td>V/VI</td>
<td>Frequent</td>
<td>Sparse</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PDD, possible AD</td>
</tr>
<tr>
<td>PDD</td>
<td>Neg</td>
<td>&lt; 15 mo</td>
<td>NA</td>
<td>Sparse</td>
<td>Mild</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PDD, possible AD</td>
</tr>
<tr>
<td>A: CDR = 0</td>
<td>Neg</td>
<td>1.7 y</td>
<td>IV</td>
<td>NA</td>
<td>None</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Normal brain</td>
</tr>
<tr>
<td>B: CDR = 0</td>
<td>Neg</td>
<td>2.4 y</td>
<td>III</td>
<td>NA</td>
<td>Moderate</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Possible AD</td>
</tr>
<tr>
<td>C: CDR = 0</td>
<td>Pos</td>
<td>2.4 y</td>
<td>IV</td>
<td>NA</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>D: CDR = 0.5</td>
<td>Pos</td>
<td>1.1 y</td>
<td>III</td>
<td>NA</td>
<td>Moderate</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Possible AD</td>
</tr>
<tr>
<td>E: CDR = 0</td>
<td>Pos</td>
<td>1.4 y</td>
<td>IV</td>
<td>Frequent</td>
<td>Sparse</td>
<td>Present</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Normal brain</td>
</tr>
<tr>
<td>F: CDR = 2–4.5</td>
<td>Pos</td>
<td>2 mo</td>
<td>NA</td>
<td>Moderate</td>
<td>Present</td>
<td>Present</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Possible AD</td>
</tr>
<tr>
<td>1: CDR = 0.5</td>
<td>Neg</td>
<td>36 mo</td>
<td>HP−y</td>
<td>NA</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>2: CDR = 0</td>
<td>Neg</td>
<td>23 mo</td>
<td>HP−y</td>
<td>NA</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>3: CDR = 0</td>
<td>Neg</td>
<td>15 mo</td>
<td>HP−y</td>
<td>NA</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>4: CDR = 1</td>
<td>Neg</td>
<td>2.0</td>
<td>HP−y</td>
<td>NA</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>5: CDR = 1</td>
<td>Neg</td>
<td>5.0</td>
<td>HP−y</td>
<td>NA</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>6: CDR = 1</td>
<td>Neg</td>
<td>20 mo</td>
<td>HP−y</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>39^{b}</td>
</tr>
<tr>
<td>7: CDR = 1</td>
<td>Pos</td>
<td>12 mo</td>
<td>HP−y</td>
<td>NA</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>42^{b}</td>
</tr>
<tr>
<td>8: CDR = 0.5</td>
<td>Neg</td>
<td>26 mo</td>
<td>HP+e</td>
<td>NA</td>
<td>2</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45^{a}</td>
</tr>
<tr>
<td>9: CDR = 2</td>
<td>Pos</td>
<td>27 mo</td>
<td>HP+e</td>
<td>NA</td>
<td>3</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>66^{a}</td>
</tr>
<tr>
<td>10: CDR = 1</td>
<td>Pos</td>
<td>28 mo</td>
<td>HP+e</td>
<td>NA</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80^{a}</td>
</tr>
</tbody>
</table>

Key: Aβ, amyloid β-protein; AD, Alzheimer’s disease; CAA, cerebrovascular amyloid; CDR, Clinical Dementia Rating; CERAD, Consortium to Establish a Registry for Alzheimer’s Disease; CJD, Creutzfeldt-Jakob disease; DLB, dementia with Lewy bodies; ELISA, enzyme-linked immunosorbent assay; IHC, immunohistochemistry; NA, not assessed; Neg, negative; NPH, normal pressure hydrocephalus; PDD, Parkinson’s disease dementia; PiB, Pittsburgh Compound B; PM, postmortem; Pos, positive.

* Interval between PiB scan and postmortem analysis.
* Level of Aβ by enzyme-linked immunosorbent assay or anti-Aβ IHC.
* Possible or definite AD dementia by CERAD criteria (Mirra et al., 1991).
* Not assessed.
* Subject F had 3 assessments.
* Biopsy study of subjects with suspected NPH.
* Positive or negative for hyperphosphorylated tau (HP).
* Aβ plaques per ×100 field detected with (IHC).
ported in transgenic mice (Klunk et al., 2005b; Toyama et al., 2005). Thus, as expected, the sensitivity of amyloid imaging for all forms of Aβ deposits will be somewhat less than 100%. However, it should be noted that none of the PiB-negative cases with postmortem Aβ deposits met criteria for definite AD dementia, so the sensitivity of PiB for Aβ deposition in pathologically proven definite AD dementia is likely to be closer to 100% than its sensitivity for any form of Aβ deposition. A similar finding has recently been published for the F-18 agent, florbetapir (Clark et al., 2011).

2.1. Conclusion

Because the numbers are relatively small, the 7 primary, peer-reviewed studies discussed above (6 using standard neuropathological criteria) in 14 autopsy and 10 biopsy cases currently provide only “sufficient evidence of an association between PiB PET and postmortem assessment of Aβ pathology.”

3. The use of amyloid imaging in therapeutic trials in AD dementia

Two broad uses of most biomarkers in clinical trials are for entry screening and as an outcome measure. Screening into a clinical trial is typically based on cross-sectional (1-time) collection of biomarker data. Outcome measures require acquisition of pre- and posttreatment (i.e., longitudinal) data. The breadth of applicability of amyloid imaging in clinical trials will be different whether it is used as a screening tool or as an outcome measurement of drug efficacy. For example, the use of amyloid-positivity as an inclusion criterion may be applicable to almost all clinical trials directed at AD dementia, MCI that is “prodromal to AD dementia,” as well as primary prevention trials directed at cognitively normal individuals in whom the process of cerebral β-amyloidosis has begun. The purpose of this screening is to increase the homogeneity of the clinical trial population by including only those with Aβ deposition. Thus amyloid imaging as a screen could be useful regardless of the mechanism of action of the putative therapeutic. In contrast, amyloid imaging as an outcome measure is likely to be most applicable to therapeutics designed to significantly affect fibrillar Aβ levels over time.

3.1. Screening and cross-sectional association between clinical diagnosis and amyloid imaging

Although a diagnosis of clinically “probable AD dementia” made using standard criteria (McKhann et al., 1984) in the setting of a specialized center such as the Alzheimer’s Disease Centers in the USA has been confirmed by autopsy in over 95% of cases (Mayeux et al., 1998), this number can drop to near 70% in less specialized settings (Knopman et al., 2001). Inclusion of only amyloid-positive AD dementia subjects in clinical trials is likely to increase the homogeneity of the study population due to exclusion of non-Alzheimer dementias. This would be most important in trials of antiamyloid therapies, but is likely to be important for any AD dementia trial. Table 2 shows that across 14 specialty centers plus the Alzheimer’s Disease Neuroimaging Initiative (ADNI), and using the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer’s Disease and Related Disorders Association (ADRA) criteria for the diagnosis of AD, 328 of 341 (96%) clinically diagnosed AD dementia patients were amyloid-positive. This is consistent with the clinical diagnostic accuracy of Alzheimer’s Disease Centers previously found by autopsy (Mayeux et al., 1998). This finding suggests that the use of amyloid imaging can extend a very high rate of diagnostic accuracy to even less specialized centers when used as a screening tool in conjunction with a clinical diagnosis of dementia.

Although this sensitivity for clinical AD dementia is encouraging, the specificity of amyloid imaging for the clinical diagnosis of AD dementia deserves some attention in the context of using a biomarker for screening entry into AD dementia clinical trials. Table 2 suggests that the specificity of amyloid imaging for the “diagnosis” of AD dementia is approximately 76% when including only AD dementia and control subjects and would be worse with MCI subjects included. However, it must be kept in mind that amyloid imaging will not be used in isolation to make a diagnosis of AD dementia, MCI, or “normal aging.” Amyloid imaging will be used to assess the underlying pathophysiology of subjects who have already been clinically evaluated and given a preliminary diagnosis. In this sense, the 24% PiB-positive subjects in the cognitively normal group is not a “false positive”; rather, based on the autopsy studies discussed above, these are more likely to be true positives for the presence of Aβ deposition, and the same will apply to the MCI subjects discussed below.

It is evident from Table 2 that the absolute value (and dynamic range) of PiB retention is dependent on the particular center conducting the study. This partly relates to trivial issues like the use of different units to express the outcome (i.e., DVR, SUVR/tissue ratios, BPND, MCBP, etc.), but also relates to some true differences caused by scanner differences, use of single regions or global means, use of atrophy correction, the size and location of the reference region, and other factors. Even so, the values do tend to fall into a relatively narrow range. In addition, the cutoff values used in these studies all seem to identify similar subjects as amyloid-positive and amyloid-negative. This simplifies the task of standardization across studies. Over these 15 peer-reviewed studies of 341 AD dementia and 651 cognitively normal subjects, the difference in PiB retention observed in AD dementia subjects and cognitively normal controls was highly significant (p < 0.001), and the effect sizes were very large (3.2 ± 1.4).
Table 2
PiB PET studies comparing AD dementia patients with cognitively normal controls

<table>
<thead>
<tr>
<th>Institution</th>
<th>AD dementia</th>
<th>Cognitive normal</th>
<th>Effect size</th>
<th>p value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PiB(+) / PiB(−), n</td>
<td>Quantitative PiB</td>
<td>PiB(+) / PiB(−), n</td>
<td>Quantitative PiB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>London</td>
<td>13 / 0</td>
<td>1.98 ± 0.30</td>
<td>1 / 13</td>
<td>1.08 ± 0.04</td>
<td>5.29 &lt; 0.001</td>
</tr>
<tr>
<td>Korea</td>
<td>8 / 2</td>
<td>1.66 ± 0.28</td>
<td>0 / 10</td>
<td>1.16 ± 0.06</td>
<td>3.24 &lt; 0.001</td>
</tr>
<tr>
<td>Tuebingen</td>
<td>6 / 0</td>
<td>1.80 ± 0.19</td>
<td>0 / 10</td>
<td>1.18 ± 0.07</td>
<td>4.77 &lt; 0.001</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>22 / 0</td>
<td>2.25 ± 0.22</td>
<td>11 / 32</td>
<td>1.27 ± 0.20</td>
<td>4.67 &lt; 0.001</td>
</tr>
<tr>
<td>Munich</td>
<td>32 / 0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Finland</td>
<td>16 / 1</td>
<td>1.73 ± 0.32</td>
<td>0 / 11</td>
<td>1.06 ± 0.10</td>
<td>3.19 &lt; 0.001</td>
</tr>
<tr>
<td>Boston</td>
<td>35 / 0</td>
<td>1.69</td>
<td>21 / 17</td>
<td>1.19</td>
<td>NA</td>
</tr>
<tr>
<td>Mayo</td>
<td>13 / 0</td>
<td>2.3 (median)</td>
<td>6 / 14</td>
<td>1.3 (median)</td>
<td>NA</td>
</tr>
<tr>
<td>Amsterdam</td>
<td>21 / 0</td>
<td>1.86 ± 0.12</td>
<td>1 / 19</td>
<td>1.23 ± 0.39</td>
<td>2.47 &lt; 0.001</td>
</tr>
<tr>
<td>Columbia</td>
<td>17 / 1</td>
<td>1.77</td>
<td>1 / 17</td>
<td>1.11</td>
<td>NA</td>
</tr>
<tr>
<td>Sweden</td>
<td>32 / 5</td>
<td>2.17 ± 0.51</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Australia</td>
<td>52 / 1</td>
<td>2.46 ± 0.49</td>
<td>58 / 119</td>
<td>1.49 ± 0.44</td>
<td>2.09 &lt; 0.001</td>
</tr>
<tr>
<td>Berkeley/UCSF</td>
<td>38 / 1</td>
<td>1.58 ± 0.27</td>
<td>3 / 27</td>
<td>1.13 ± 0.18</td>
<td>2.00 &lt; 0.001</td>
</tr>
<tr>
<td>St. Louis</td>
<td>8 / 0</td>
<td>NA</td>
<td>44 / 197</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Total</td>
<td>328 / 13</td>
<td>96% positive</td>
<td>155 / 496</td>
<td>24% positive</td>
<td></td>
</tr>
</tbody>
</table>

Key: +: positive; −: negative; AD, Alzheimer’s disease; ADNI, Alzheimer’s Disease Neuroimaging Initiative; BPND, binding potential; CDR, Clinical Dementia Rating; DVR, distribution volume ratio; MCBP, mean cortical binding potential; NINDS-ADRD, National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer’s Disease and Related Disorders Association; PET, positron emission tomography; PiB, Pittsburgh Compound B; SUVR, standardized uptake value ratio; UCSF, University of California at San Francisco.

* Determined using cutoffs given in the report or the level giving the best combination of sensitivity and specificity for this study.
* AD dementia versus control quantitative data.
* SUVR or tissue ratio (cerebellar reference).
* DVR (cerebellar reference).
* Visual read.
* NA: not assessed or not applicable.
* Expressed as BPND + 1.

3.1.1. Conclusion

The primary, peer-reviewed studies discussed above, using standard clinical diagnostic criteria for AD dementia in 992 cases (combined AD dementia and controls; χ² = 469; p < 0.0001) provide “sufficient evidence of a direct relationship between PiB PET signal and the clinical diagnosis of AD dementia.”

There are other, more “fine-grained” approaches that have been employed to provide even stronger support for the link between a biomarker and a clinical diagnosis. One of these is demonstration of statistical differences in group means between control versus MCI and MCI versus AD dementia. Although this approach is commonly used, application of it to amyloid imaging tends to obscure the major advantage of this particular biomarker: distinguishing amyloid-positive subtypes of controls, MCI, and AD dementia subjects. It makes the least sense to compare mean values of PiB retention in a group of MCI subjects, when most MCI cohorts are composed of similar-sized subgroups of 2 very different populations: those with detectable amyloid (and presumably prodromal AD dementia; see below) and those without. The strength of amyloid imaging is to distinguish the amyloid-positive subtypes within these diagnostic groups, not to distinguish between the diagnostic groups. That having been recognized, Table 3 shows several studies in which amyloid imaging has shown statistically significant differences in comparisons of control versus MCI or MCI versus AD dementia (Devanand et al., 2010; Forsberg et al., 2008, 2010; Jack et al., 2008; Kemppainen et al., 2007; Koivunen et al., 2008; Li et al., 2008; Lowe et al., 2009; Mormino et al., 2009; Okello et al., 2009b; Pike et al., 2007; Rowe et al., 2010; Tolboom et al., 2009d), although when the control group contains subjects with very high levels of amyloid, the control versus MCI distinction may not be apparent (Forsberg et al., 2008; Jack et al., 2008; Lowe et al., 2009).

3.1.2. Conclusion

Although these 13 primary, peer-reviewed studies using standard diagnostic criteria and comprising 960 cases (combined controls, MCI, and AD) show the ability to discriminate between control versus MCI in 8/12 studies and MCI
versus AD dementia in 9/10 studies, the significant overlap at the individual subject level between the MCI group and both the control and the AD dementia groups suggests that there is only “sufficient evidence of an association between PiB PET and the distinction between control versus MCI and MCI versus AD dementia.”

Another fine-grained approach is the correlation of a continuous measure of the biomarker (e.g., PiB SUVr) with continuous measures of cognition or clinical function (e.g., episodic memory scores or Clinical Dementia Rating [CDR] sum of boxes [SOB]). This sort of correlation is not a strength of amyloid imaging, although several studies do report significant correlations of this type. Five studies have shown a significant correlation between degrees of PiB retention and levels of cognitive performance when combining controls, MCI, and AD dementia subjects (Edison et al., 2007; Forsberg et al., 2010; Pike et al., 2007; Rentz et al., 2010; Tolboom et al., 2009b), but 1 has not (Jagust et al., 2009). Several studies (3 in controls, 1 in MCI, and 3 in AD dementia) have found a significant correlation between PiB and cognition in single diagnostic groups (Darreh-Shori et al., 2010; Engler et al., 2006; Grimmer et al., 2009a; Mormino et al., 2009; Pike et al., 2007; Rentz et al., 2010)—although 1 of these studies suggests the correlation is mediated through hippocampal atrophy (Mormino et al., 2009).

At least 5 studies failed to find any correlation between continuous PiB retention measures and continuous cognition scores in 1 or more of the diagnostic groups when the latter were considered in isolation (Aizenstein et al., 2008; Forsberg et al., 2010; Furst et al., 2010; Pike et al., 2007; Rowe et al., 2010). Very demanding memory tests may be required to demonstrate this continuous correlation in cognitively normal controls (Rentz et al., 2010). Most likely, the usually weak correlations between continuous measures of amyloid imaging and cognition is due to the fact that Aβ deposition is a very early event in the full spectrum of pathophysiological changes in this disorder and does not necessarily correlate quantitatively with late events like cognition and clinical function (for a review, see Jack et al., 2010). When done side-by-side, other later stage biomarkers (e.g., CSF tau protein levels, hypometabolism, and brain atrophy) tend to correlate better quantitatively with degree of cognitive impairment than does PiB (Engler et al., 2006; Jack et al., 2009; Jagust et al., 2009; Mormino et al., 2009; Storandt et al., 2009).

3.1.3. Conclusion

Given the fairly weak correlations and the contradictory finding, these studies can provide only “limited/suggestive evidence of an association between continuous PiB PET levels and continuous measures of cognitive/clinical performance.”

3.2. Outcome measurement and longitudinal change of PiB retention as a marker of disease progression

Longitudinal change in a biomarker is often considered a surrogate for biological progression of disease. In the case of AD, it is increasingly believed that the noise in biomarkers over time may be less than that in cognitive or functional measures of disease progression. In turn, it is believed that this decreased variability will facilitate detection of drug-induced changes in disease progression, i.e., a disease-modifying effect. Therefore, change in a biomarker over time (i.e., the natural history of that biomarker) is increasingly being considered as an outcome measure for clinical trials. Driven largely by the evaluation of a variety of analytical approaches to the measurement of magnetic resonance imaging (MRI) volumetry and cerebral metabolism measured by fluorodeoxyglucose (FDG) PET, a metric has evolved that is the sample size required for a drug-induced 25% reduction in the rate of change in a biomarker over some specified period of time. This metric has become a staple of biomarker comparisons in the ADNI (Cummins, 2010; Weiner et al., 2010).
However, this 25% rate-reduction metric may not be well suited for amyloid imaging for 2 reasons. First, PiB retention increases slowly over time. Aβ deposition is believed to begin 10–15 years prior to the diagnosis of AD dementia and, as will be discussed below, continues to progress slowly during the clinical course of AD dementia. Obviously, an amyloid-free individual destined to develop typical AD dementia must have an increase in PiB retention in order to progress from the PiB retention typical of controls (approximately 1.1 SUVR units) to that typically found in AD dementia (> 2.0 SUVR units). However, this accumulation often occurs over 10–20 years, suggesting a rate of increase of 0.05–0.10 SUVR units per year. Given that the test-retest variability is on this same order, detecting a 25% reduction in this rate will be difficult. Second, and more importantly, this slow rate of change becomes moot when considering the fact that achieving a 25% reduction in the rate of Aβ accumulation over time may not be clinically meaningful unless the drug is begun when there is still little or no Aβ deposition. Any clinically relevant antiamyloid drug will likely need to actually decrease the Aβ fibrillar load (i.e., PiB retention) from baseline (although other forms such as soluble oligomers may be targeted by some Aβ-lowering treatments). This means a >100% reduction in the rate of increase of PiB signal over time. This is not an unreasonable goal, and those contemplating the use of amyloid imaging as an outcome biomarker in trials have the unique advantage of being able to refer to published data showing that a significant reduction in fibrillar Aβ load could be detected over 78 weeks with only 20 mild-to-moderate AD dementia patients in the treatment arm and 8 in the placebo arm (Rinne et al., 2010). Using the passive immunotherapy, bapineuzumab, in a Phase 2 trial, Rinne et al. reported a decrease of 0.9 PiB SUVR units in the bapineuzumab-treated patients over 78 weeks that was significant both when compared with the patients’ own baseline PiB retention or to the increase of 0.15 SUVR units observed in the placebo group over the same 1.5-year period of time. This outcome is thus equivalent to a 160% decrease in the rate and represents a 25% reduction in the absolute amyloid load (not rate) of the treated group compared with the placebo group. As will be seen below, the increase observed in the bapineuzumab placebo group is typical of that observed in AD dementia natural history studies. It is the characterization of this natural history as a foundation for drug trials that makes longitudinal studies of amyloid imaging measures important.

### 3.2.1. Conclusion

This single study, performed using standard methodology and showing detection of a drug-induced change in the absolute level of amyloid load constitutes “limited/suggestive evidence of an association between PiB PET and drug-induced changes in disease biology.”

Given the newness of amyloid imaging as a biomarker, there are still relatively few longitudinal data in the published literature. The few small studies that have been published suggest that, in individuals who are amyloid-positive at baseline, a relatively slow increase in amyloid deposition occurs across the full spectrum of the illness from preclinical stages to symptomatic AD dementia. Individuals who are amyloid-negative at baseline tend to show little change over time—although some become amyloid-positive and then progress. Amyloid-positive individuals rarely, if ever, revert to amyloid-negative status. While this is apparent at the individual level, group increases are not always observed at the AD dementia stage. The first longitudinal follow-up study of the 16 AD dementia subjects included in the original report of PiB PET imaging (Klunk et al., 2004), showed no significant group change over 2 years of follow-up in any brain area examined (Engler et al., 2006). Similar findings were reported for group-level determinations in 14 AD dementia patients studied at the Turku PET Center over 2 years (Scheinin et al., 2009). However, closer inspection of these data showed that a majority of the AD dementia patients in the Engler et al. study tended to show a combination of increased PiB retention and decreased cerebral metabolism (Klunk et al., 2006). Similarly, while only the medial frontal cortex showed a statistically significant (4.3%) group increase in the Turku study (Scheinin et al., 2009), 10 of the 14 subjects tended to show an increase in PiB retention. Similar results can be seen in the ADNI natural history data, where group changes were not significant but 3 of the 12 AD dementia patients showed a significant increase in PiB retention over 1 year (Jagust et al., 2010). In a group of 21 cognitively normal, 32 amnestic MCI, and 8 AD dementia subjects, Jack et al. (2009) found that the annual rate of change in global PiB retention ratio was significantly greater than zero over all subjects (p < 0.001), and individually among cognitively normal (p = 0.002), and amnestic MCI subjects (p = 0.008), with a trend in Alzheimer’s disease (p = 0.11). Overall, the rate of change was small (0.03–0.06 SUVR units per year) across these 3 groups but tended to be higher in the subjects who were amyloid (PiB) positive at baseline (Jack et al., 2009). This rate matches the expected rate of change mentioned above. Grimmer et al. followed 24 AD dementia patients and found a significant increase in PiB retention of approximately 0.14 SUVR units over 24 months (8.7 ± 14.3%; annual rate = 3.92%) (Grimmer et al., 2010). Grimmer et al. found that the increase was dependent on apolipoprotein-E (ApoE) ε4 gene dose, with the 5 homozygotes showing a 0.31 ± 0.27 SUVR increase over the 2 years. These latter 2 studies suggest that the rate of amyloid accumulation is relatively constant over the course of the disease—at least through the phase of mild AD dementia. This result suggests that the course of PiB-detectable Aβ deposition over time is best described by either a linear model or a sigmoid with a very gradual incline that does not level-off until the late stages of AD dementia.
Two studies have examined the relationship between the rate of change in PiB retention and the rate of cognitive decline on CDR and Mini Mental State Examination (MMSE) and differ on their results, so that insufficient evidence exists to comment on this specific relationship (Grimmer et al., 2010; Jack et al., 2009).

3.2.2. Conclusion
Given some contradictory evidence about the ability to detect change in PiB retention over time, the reports discussed above constitute only “limited/suggestive evidence of an association between PiB PET signal and disease progression.”

4. The use of amyloid imaging in trials of nondemented subjects

There is growing consensus that it will be necessary to study drugs earlier than the stage of clinical dementia in order to find robust treatment effects for the pathophysiology of Alzheimer’s disease. The first difficulty in conducting these early-intervention trials is that clinical diagnosis becomes less and less accurate as we move into these prodromal stages and clinical evaluation as currently conducted becomes useless in presymptomatic phases. The second difficulty is that it may take prohibitively long periods of time to determine a drug-effect on clinical measures alone when subjects are enrolled at very early stages. Thus, biomarkers are likely to play their most valuable role in clinical trials of “not-yet-demented” subjects. In the context of such trials, biomarkers can play 3 important roles. The first 2 roles are for screening and as an outcome measure, as discussed above for trials in AD dementia. The third role is staging nondemented subjects to select those that are likely to show a significant clinical change over a relatively short period of time (e.g., conversion from a diagnosis of MCI to AD dementia). A question that must be asked about a biomarker in this context is can the biomarker accurately identify subjects who are destined to progress to the next clinical stage? Therefore, the next 2 sections will review cross-sectional studies that have measured PiB retention at baseline in MCI or cognitively normal subjects (i.e., relevant for screening purposes) and longitudinal studies that have assessed whether baseline PiB retention can predict future cognitive course (relevant for predicting impending clinical progression). The issues of using amyloid imaging as an outcome measure in predementia subjects are essentially identical to those discussed above for AD dementia and will not be discussed again here.

4.1. Screening and cross sectional association between clinical diagnosis and amyloid imaging

Many studies have reported the results of PiB PET imaging in MCI subjects (Bourgeat et al., 2010; Butters et al., 2008; Cohen et al., 2009; Devanand et al., 2010; Forsberg et al., 2008, 2010; Fripp et al., 2008; Jack et al., 2008, 2009; Jagust et al., 2009, 2010; Kemppainen et al., 2007; Koivunen et al., 2008; Li et al., 2008; Lopresti et al., 2005; Lowe et al., 2009; Okello et al., 2009a, 2009b; Pike et al., 2007; Price et al., 2005; Raji et al., 2008; Rowe et al., 2007, 2010; Tolboom et al., 2009b, 2009d; Wolk et al., 2009; Zhou et al., 2007). However, as some of these studies represent progressive accumulations of subjects and variations in the analysis method, Table 4 describes only the most recent relevant report from each group, in order to avoid counting single subjects more than once in this analysis.

Table 4
PiB PET studies in MCI: prevalence and conversion

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of MCI subjects</th>
<th>Number of converters</th>
<th>Duration (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>PiB(+)</td>
<td>PiB(−)</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rowe et al., 2010</td>
<td>57</td>
<td>39</td>
<td>18</td>
</tr>
<tr>
<td>Devanand et al., 2010</td>
<td>24</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>Tolboom et al., 2009d</td>
<td>13</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Lowe et al., 2009</td>
<td>23</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Longitudinal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forsberg et al., 2008</td>
<td>21</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Koivunen et al., 2008</td>
<td>15</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>Wolk et al., 2009</td>
<td>23</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Okello et al., 2009b</td>
<td>31</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>Jagust et al., 2010a</td>
<td>65</td>
<td>47</td>
<td>18</td>
</tr>
<tr>
<td>Total longitudinal</td>
<td>155</td>
<td>101</td>
<td>54</td>
</tr>
<tr>
<td>Total all studiesb</td>
<td>272</td>
<td>161</td>
<td>111</td>
</tr>
</tbody>
</table>

Key: +, positive; −, negative; ADNI, Alzheimer’s Disease Neuroimaging Initiative; MCI, mild cognitive impairment; PET, positron emission tomography; PiB, Pittsburgh Compound B; SUVR, standardized uptake value ratio.

a ADNI data.
b Only the most recent report from each research group was used to avoid subjects being reported multiples times. When only graphical data were published without mean ± SD, these values were derived from the graph. If specific assignment of PiB-positivity was not made in the report, visual interpretation of graphic data using 1.5 SUVR or an equivalent measure in other units was performed.
One thing that is quickly apparent when working with amyloid imaging is the bimodal nature of the scans. Visually, they tend to be clearly positive or clearly negative, suggesting that people fall into 1 or the other of 2 distinct populations (amyloid-positive and amyloid-negative) (Ng et al., 2007; Rabinovici et al., 2007; Suotunen et al., 2010; Tolboom et al., 2010). The use of visual reads as a screening tool could be very useful for clinical trials, as this is relatively easy to standardize. This bimodal character also is apparent in quantitative data. Fig. 1A shows a histogram of global cortical PiB retention across more than 300 subjects of all diagnoses combined (A); in 50 MCI subjects (B); in 90 cognitively normal control subjects (C); and in 41 AD dementia patients (D). The same 2 normal curves are superimposed on all 4 populations and the typical cutoff for PiB-positivity is shown with a dashed, vertical line. Abbreviations: AD, Alzheimer’s disease; MCI, mild cognitive impairment; PiB, Pittsburgh Compound B; SUVR, standardized uptake value ratio.

One thing that is quickly apparent when working with amyloid imaging is the bimodal nature of the scans. Visually, they tend to be clearly positive or clearly negative, suggesting that people fall into 1 or the other of 2 distinct populations (amyloid-positive and amyloid-negative) (Ng et al., 2007; Rabinovici et al., 2007; Suotunen et al., 2010; Tolboom et al., 2010). The use of visual reads as a screening tool could be very useful for clinical trials, as this is relatively easy to standardize. This bimodal character also is apparent in quantitative data. Fig. 1A shows a histogram of global cortical PiB retention across more than 300 subjects of all diagnoses combined (A); in 50 MCI subjects (B); in 90 cognitively normal control subjects (C); and in 41 AD dementia patients (D). The same 2 normal curves are superimposed on all 4 populations and the typical cutoff for PiB-positivity is shown with a dashed, vertical line. Abbreviations: AD, Alzheimer’s disease; MCI, mild cognitive impairment; PiB, Pittsburgh Compound B; SUVR, standardized uptake value ratio.

The standard PiB-positive cutoff is drawn with a dashed, vertical line, but it is clear that there will be some overlap of the 2 populations wherever a cutoff is drawn. Fig. 1C is a similar representation of 90 cognitively normal control subjects and Fig. 1D shows 41 mild-to-moderate AD dementia patients. Approximately 80% of the controls reside within the amyloid-negative population and only 1 of 90 has reached even the mean of the amyloid-positive population. In contrast > 95% of the AD dementia patients reside within the amyloid-positive population and 80% are above the mean of that population. Fig. 1B shows a histogram of 50 MCI subjects. The MCI subjects do not represent a third, separate, intermediate population (as they would on cognitive measures) but they are comprised of subjects that tend to fit into 1 or the other of the 2 populations that comprise...
our total cohort. This is a graphical representation of the fact that somewhere near half (this proportion varies with the age and ApoE ε4 frequency of a particular population) of MCI subjects are AD dementia-like as regards PiB signal and the rest are control-like. Of the 272 MCI subjects from the 9 studies included in the analysis in Table 4, 59% are PiB-positive. If broken down into amnestic and nonamnestic subtypes, 63% of 242 amnestic MCI subjects were amyloid-positive and 27% of the 30 nonamnestic subjects were amyloid-positive. These data show that PiB imaging is very well-suited to dichotomize MCI patients based on the underlying pathophysiology. This could be extremely useful for screening into a clinical trial. However, amyloid imaging in isolation from clinical and cognitive data are poorly suited for the identification of a subject who is likely to receive a clinical diagnosis of MCI, because a roughly equal portion of MCI cases are PiB-positive and PiB-negative. But this is not how this biomarker would be used in a clinical trial. The reasons trial designers would want to know the amyloid-status of an MCI subject would be if knowledge of that status could help them: (1) decrease the heterogeneity of their trial population (applies to most trials); (2) identify a cohort that is likely to respond to a drug with a certain mechanism (applies mainly to antiamyloid trials); and (3) assemble a cohort that is likely to convert to an endpoint of AD dementia in a relatively short period of time (applies to most secondary prevention trials). The first 2 points relate to the discussion above. The third point (prediction of progression) will be discussed in the following section.

4.1.1. Conclusion

The primary, peer-reviewed studies discussed above, using standard diagnostic criteria for MCI and AD dementia in 272 cases provide “sufficient evidence for the lack of a quantitative association between PiB PET and the clinical diagnosis of MCI.” Nevertheless, PiB PET could be useful for screening MCI subjects into amyloid-positive versus amyloid-negative subtypes for inclusion in clinical trials.

Very similar issues apply to screening presymptomatic subjects into clinical trials. Table 2 shows that 24% of 651 cognitively normal control subjects studied with PiB PET are amyloid-positive. Thus, when comparing only AD dementia and control subjects, a PiB-negative scan was 76% sensitive and 96% specific for identifying controls. Fig. 1C shows that most of these are likely to be on the low end of the amyloid-positive spectrum (< 2.2 SUVR) and can be distinguished from the vast majority of AD dementia patients (> 2.2 SUVR). This suggests that a cutoff higher than that typically used to detect any amyloid deposition (i.e., the “PiB-negative” cutoff) could better differentiate between cognitively normal controls and clinically diagnosed AD dementia patients. Using the typical “PiB-negative” cutoff of SUVR = 1.6 for the Pittsburgh data gives very similar sensitivity (80%) and specificity (98%) for a PiB-negative scan to identify controls. However, if we use an SUVR of 2.2 as the “PiB-AD” cutoff for the Pittsburgh data, a PiB-negative scan becomes 96% sensitive and 90% specific for identifying controls. Thus, it may be best to use a “PiB-negative” cutoff when the goal is to identify subjects with any evidence of Aβ deposition and a “PiB-AD” cutoff when the goal is to identify subjects who are AD-like. The point to be made here is simply that, although PiB PET can identify fibrillar Aβ deposition in approximately 25% of cognitively normal controls, this deposition is typically low and very distinguishable from that seen in AD dementia.

It also is important to recognize that the percentage of amyloid-positive subjects in a cognitively normal population is highly dependent on both age and the presence of an ApoE ε4 allele. In a study of 241 cognitively normal subjects with elevated PiB retention rose in an age-dependent manner from 0% at ages 45–49 years to 19% at 60–69 years to 30.3% at 80–88 years (Morris et al., 2010). Rowe et al. studied 177 cognitively normal controls and also found an age effect such that elevated PiB retention was seen in 18% at age 60–69 years and 65% older than age 80 (Rowe et al., 2010). Morris et al. also showed that there was a gene dose effect for the ApoE ε4 genotype, with greater PiB retention with increased numbers of ε4 alleles such that 8.2% of age 60–69 ε4 noncarriers were PiB-positive while 75% of age 80–89 ε4 carriers were PiB-positive (Morris et al., 2010). In addition to the 177 cognitively normal controls, Rowe et al. also studied 57 MCI and 53 AD dementia subjects and reported that ε4 carriers had higher PiB retention in the control and MCI groups, but not in the AD dementia group (Rowe et al., 2010).

4.1.2. Conclusion

The primary, peer-reviewed studies shown in Table 2, using standard diagnostic criteria for normal cognition and AD, in 992 cases (651 controls and 341 AD; χ² = 469; p < 0.0001) provide “sufficient evidence of a direct relationship between PiB PET and the clinical diagnosis of cognitively normal.” However, this diagnosis remains a neuropsychological/clinical one and PiB PET is best suited for screening cognitively normal subjects into amyloid-positive and amyloid-negative subtypes for inclusion in clinical trials.

In addition to the clinical categories of control, MCI, and AD, one also must consider the issue of differential diagnosis among different dementias when considering the usefulness of a biomarker in clinical trial design. Amyloid imaging will not distinguish mixed dementias when only 1 component of the pathology is Aβ deposition. That is, amyloid imaging is likely to be good at ruling-in Aβ pathology, but cannot rule out non-Aβ pathology. Thus, a majority of cases of dementia with Lewy bodies (DLB) will show AD dementia-like levels and patterns of Aβ deposition (Edison et al., 2008; Gomperts et al., 2008; Maetzler et al., 2009; Rowe et al., 2007) and the clinical symptoms or additional imaging with dopamine transporter tracers may help distinguish these cases if necessary (McKeith et al.,
levels of fibrillar Aβ as approximately one-quarter of normal elderly show low Aβ deposition without cognitive impairment, a similar proportion of elderly subjects with dementia from other causes could show these same low amounts of amyloid, even if the amyloid is not contributing to the clinical dementia. The use of a higher “PiB-AD” cutoff as discussed above may help screen out amyloid-positive subjects with dementia due to causes other than AD dementia who might have incidental/low amyloid deposition.

4.2. Prediction of future cognitive/clinical course

Only a portion of patients with MCI progress to clinical AD dementia over 5–10 years (Petersen et al., 1999; Ritchie et al., 2001; Visser et al., 2006) and a recent meta-analysis concluded that most people with MCI will not progress to dementia even after 10 years of follow-up (Mitchell and Shiri-Feshki, 2009). In a longitudinal study of 134 MCI cases followed for 4 or more years, Hansson et al. (2006) reported that 43% developed clinical AD, 42% remained cognitively stable (but could, of course, develop AD dementia in the future), and 15% developed other dementias (mostly vascular). Two community-based studies have shown over one-third of patients diagnosed with MCI at baseline may eventually return to normal cognition (Ganguli et al., 2004; Larrieu et al., 2002). Obviously, it would be of great value to be able to predict which MCI subjects were destined to progress to a clinical diagnosis of AD dementia. The 5 studies listed at the bottom of Table 4 describe longitudinal follow-up of 155 MCI subjects (141 amnestic) who were followed between 1 and 3 years after their baseline PiB PET scan (Forberg et al., 2008; Jagust et al., 2010; Koivunen et al., 2008; Okello et al., 2009b; Wolk et al., 2009). Of these 155 MCI subjects, 57 (37%) progressed to a clinical diagnosis of AD dementia over 1 to 3 years. The distribution of converters was far from random across these amyloid-positive and amyloid-negative groups. Of the 57 converters, 53 came from the 101 amyloid-positive subjects (representing a 53% conversion rate) and only 4 came from the 54 amyloid-negative subjects (7% conversion rate) (χ² = 30.7; p < 0.0001). It remains to be seen whether these latter 4 amyloid-negative converters were misdiagnosed with AD dementia or represent false negatives for PiB PET. Conversion rates from amyloid-positive subjects in the amnestic and nonamnestic categories could not be determined from the data published, but nonamnestic, amyloid-positive MCI subjects did show conversion to AD dementia in at least 2 studies (Wolk et al., 2009).

Other studies also have reported conversions from MCI to AD dementia. Jack et al. reported 9 amnestic MCI subjects studied at Mayo Clinic who had PiB PET scans (Jack et al., 2009). Three of these subjects converted to AD dementia within 1 year and 1 “reverted” to normal cognition, but it was not reported if these subjects were PiB-positive or PiB-negative in that report. Interestingly, Wolk et al. report 3 reversions to normal cognition and all 3 were PiB-negative.

4.2.1. Conclusion

The primary, peer-reviewed studies discussed above (using standard diagnostic criteria for MCI and AD) included 155 subjects and the data were so overwhelmingly significant that they constitute “sufficient evidence for a direct relationship between PiB PET and the likelihood of conversion from a clinical diagnosis of MCI to a clinical diagnosis of AD dementia over 3 years.”

There are very few similar prospective data on clinical conversions from cognitively normal controls to either MCI or AD dementia. This is not surprising given that the pre-symptomatic lag phase between initiation of Aβ deposition and emergence of clinical symptoms may be 10–15 years and most subjects have been followed for no more than 5 years. Several studies have looked retrospectively at data gathered in cohorts of subjects who were cognitively normal at baseline and then followed to the time of a PiB PET scan. Villemagne et al. reported a retrospective study of the cognitive course of 34 subjects who started with normal cognition in 1996 and were followed with 7–9 yearly visits prior to agreeing to a PiB PET scan (Villemagne et al., 2008). Ten of these 34 were classified as cognitive “decliners” by raters blinded to the PiB status. Three of these 10 were further diagnosed with MCI and 1 additional subject with AD; the other 6 decliners remained in the cognitively normal range. Seven of these 10 decliners (including all 3 MCI cases and the AD dementia case) were PiB-positive (70%) compared with 4/24 (17%) stable subjects. Although clinical conversion was not addressed, 2 other studies have similarly shown that the rate of cognitive decline in subjects...
who were cognitively normal at baseline is related to PiB retention. Storandt et al. followed the cognitive status of 135 individuals from 1985 to the time of PiB PET in 2004 (Storandt et al., 2009). They found that PiB retention was unrelated to the current cognitive performance in 2004, but was related to decline in working and visuospatial memory over the previous 19 years in the 29 subjects who were amyloid-positive (but not in the amyloid-negative group, as expected). Resnick et al. studied 57 participants for an average of 10.8 years who received a PiB PET scan near the end of that period (Resnick et al., 2010). They found greater declines over time in mental status and verbal learning and memory, but not visual memory, that were significantly associated with higher PiB retention. One of the subjects in this study progressed from normal cognition to MCI who was PiB-positive (Resnick et al., 2010). Similarly, Reiman et al. reported a single case of an ApoE ε4 homozygote who converted from normal cognition to MCI 7 months before a PiB PET scan and was found to be PiB-positive (Reiman et al., 2009).

Only 1 study has reported prospective, longitudinal cognitive outcomes in normal subjects imaged with PiB (Morris et al., 2009). In that study, 159 participants, with CDR = 0 at the time of their baseline PiB PET scan, were followed for 0.8–5.5 years. Nine of these subjects converted to a diagnosis of “dementia of the Alzheimer’s type (DAT) at the CDR 0.5 stage.” PiB retention was a stronger predictor of time to DAT (hazard ratio = 4.82; 95% confidence interval, 1.22–19.01; \( p = 0.02 \)) than age (hazard ratio = 1.14; 95% confidence interval, 1.02–1.28; \( p = 0.03 \)). Education, ApoE ε4 allele status and gender were not significant predictors. Unfortunately, the individual PiB-status of the 9 DAT-converters was not reported in this study. Jack et al. reported that 1 of 10 control subjects from a prospective Mayo study converted to MCI over 1 year, but did not report the PiB status (Jack et al., 2009).

### 4.2.2. Conclusion

Although there is 1 study with a relatively large number of subjects (\( n = 159 \)), there were only 9 conversions to DAT CDR 0.5 in this study. The other studies are retrospective or report on conversions of single subjects that were not the focus of the main study. Therefore, these data constitute only “limited/suggestive evidence of an association between PiB PET positivity and the likelihood of conversion from normal cognition to a clinical diagnosis of MCI.”

### 5. Considerations of CSF and PET methods for determination of brain Aβ

As mentioned in the Introduction, there is overlap in the utility of CSF and PET measures of amyloid deposition for clinical trials. Both are primarily measures of brain Aβ pathology. This overlap is mainly in the area of selecting amyloid-positive subjects for trials, i.e., screening. Most available data seem to indicate that CSF Aβ42 decreases relatively quickly to its final level very early in the course of the pathophysiological spectrum of AD dementia—probably presymptomatically (Blennow and Hampel, 2003; Fagan et al., 2007, 2009; Hansson et al., 2006). That is, the change in CSF appears to be almost a step-function, and longitudinal studies have not shown a progressive decrease in CSF Aβ42 over time (Buchhave et al., 2009). This is not surprising given the fact that typical concentrations of Aβ found in insoluble deposits in AD dementia cortex are approximately 5000 μg/L (approximately 1 μM) (Klunk et al., 2005b; Nisland et al., 2000), while typical soluble Aβ concentrations in the cortex are on the order of 50 μg/L (Klunk et al., 2005b) and CSF Aβ concentrations are approximately 0.5 μg/L (Fagan et al., 2006)—or 0.01% of insoluble cortical Aβ. Thus, it is not surprising that relatively little cortical Aβ would need to deposit before a new equilibrium would be established with CSF. This conclusion has 2 implications in clinical trials: (1) for screening purposes, CSF Aβ42 may drop before fibrillar Aβ is detectable by PET; and (2) as an outcome measure, CSF Aβ42 is not likely to normalize until the vast majority of cortical Aβ deposits are removed. This implies that CSF Aβ42 and PiB PET would be roughly equivalent as screening tools for AD dementia and MCI trials and that CSF Aβ42 may have advantages in detecting amyloid-positive controls—although this has yet to be demonstrated. The more dynamic nature of amyloid signal by PET and the fact that PiB retention correlates directly with fibrillar Aβ load (Ikonomovic et al., 2008) makes this a more suitable outcome measure. Indeed, the ability of PiB PET to show an amyloid-lowering effect of passive immunotherapy in humans has now been reported (Rinne et al., 2010). While we often reduce imaging data to a single number (e.g., mean cortical PiB retention), we must remember that a major advantage of any imaging technique is the wealth of regional information that is provides. Whereas amyloid PET can quantify amyloid load throughout the brain, it is not clear what pool of brain Aβ42 is represented by changes in CSF Aβ42. One study has suggested that CSF Aβ42 is most tightly correlated with PiB retention in brain regions immediately adjacent to CSF spaces (Grimmer et al., 2009b). The rich regional information in an amyloid PET scan also allows differentiation not only by quantitation but also regional specificity. This is especially important because it allows visual reads of amyloid PET scans to be highly accurate in distinguishing normal from abnormal scans (Ng et al., 2007; Rabinovici et al., 2007; Suotunen et al., 2010; Tolboom et al., 2010). Visual reads are relatively easy to standardize because the technical variables in quantifying the amyloid PET signal are not a factor. Of course, visual reads would apply almost exclusively to use in screening and do not lend themselves to detection of small changes. Therefore, CSF Aβ42 and PiB PET may be equivalent screening measures for entry into clinical trials in AD dementia and MCI.
Differences in the costs, practicalities and risks of the 2 procedures for the application at hand would determine which is better suited to a particular trial. CSF Aβ42 could have an advantage in identifying more amyloid-positive controls than PiB PET. Amyloid PET has the advantage of the easily standardizable visual read, but the greatest advantage of amyloid imaging for clinical trials is as a quantitative outcome measure for drugs expected to decrease fibrillar Aβ load. If amyloid PET is to be used as an outcome biomarker, it is necessary to obtain a pretreatment scan for comparison, so it seems logical to use this as the screening tool as well if amyloid PET will be used as the outcome measure. However, it is sometimes inappropriate to use the same measure for screening as is used for an outcome measure. In these cases, it may be appropriate to use CSF Aβ42 as the screening tool and amyloid PET as the outcome measure.

6. A research agenda to improve the applicability of amyloid PET to clinical trials

As for any biomarker, standardization of its application across many centers around the world, and across varying degrees of expertise, is critical for utility in clinical trials. As stated above, visual reads for a screening into amyloid-positive and amyloid-negative subtypes is relatively easy to standardize (Ng et al., 2007; Rabinovici et al., 2007; Suotunen et al., 2010; Tolboom et al., 2010). However, improved software to analyze and display amyloid PET scans in a standardized manner could aid in widespread applicability. With respect to quantitative assessment of amyloid PET signal, simplified dynamic methods of analysis such as Logan DVR (Lopresti et al., 2005) and simplified reference tissue model (SRTM)/SRTM2 (Tolboom et al., 2009; Zhou et al., 2007) may produce the most accurate and reproducible results, but practical considerations have frequently led to the application of tissue ratios and SUVR of short/late scans (Lopresti et al., 2005; McNamee et al., 2009), and these have proven to be good substitutes for the dynamic methods. Standardization of these quantitative ratio methods depend first on proper choice of the reference region (e.g., cerebellum or pons). This decision should be made carefully at the beginning of the trial, and it is important to not only choose the reference region carefully, but also carefully choose the exact method for delineating the region. This involves choices of normalization and automation. Decisions regarding atrophy correction or tissue segmentation (which often produces an equivalent outcome) must be made for each trial. The electronic nature of all PET data allows one to easily send the data to a central processing site, so that any variability in the application of the “analysis pipeline” can be minimized.

Another way to minimize variability is to use the same amyloid PET tracer throughout the trial. While the current discussion has centered on [C-11]PiB because of the lack of sufficient literature on the F-18 tracers, it is clear that widespread applicability of amyloid PET to the 90% of PET scanners that do not have access to a cyclotron could be enhanced by the use of an F-18 amyloid tracer. These F-18 tracers appear to come with somewhat inferior signal-to-noise qualities, and this could add to variability, so the desirability of application to many PET sites must be weighed against the trade-offs. Clearly, we need to see many more published studies with these new tracers, including direct comparisons with PiB PET in the same subjects, before we can fully judge the capabilities of such promising F-18-labeled amyloid imaging agents.

Disclosure statement

GE Healthcare holds a license agreement with the University of Pittsburgh based on the technology described in this manuscript. Dr. Klunk is a coinventor of PiB and, as such, has a financial interest in this license agreement. GE Healthcare provided no grant support for this study and had no role in the design or interpretation of results or preparation of this manuscript.

Acknowledgements

Supported by the National Institutes of Health grants: P50 AG005133, R37 AG025516, P01 AG025204.

References


The potential of functional MRI as a biomarker in early Alzheimer’s disease

Reisa Sperling

Center for Alzheimer Research and Treatment, Brigham and Women’s Hospital, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Abstract

Functional magnetic resonance imaging (fMRI) is a relative newcomer in the field of biomarkers for Alzheimer’s disease (AD). fMRI has several potential advantages, particularly for clinical trials, as it is a noninvasive imaging technique that does not require the injection of contrast agent or radiation exposure and thus can be repeated many times during a longitudinal study. fMRI has relatively high spatial and reasonable temporal resolution, and can be acquired in the same session as structural magnetic resonance imaging. Perhaps most importantly, fMRI may provide useful information about the functional integrity of brain networks supporting memory and other cognitive domains, including the neural correlates of specific behavioral events, such as successful versus failed memory formation.

© 2011 Elsevier Inc. All rights reserved.

Keywords: Alzheimer’s disease; Biomarker; Cognitive impairment; Dementia; fMRI; Functional magnetic resonance imaging

1. Functional MRI as a biomarker in early Alzheimer’s disease

Functional magnetic resonance imaging (fMRI) is a relative newcomer in the field of biomarkers for Alzheimer’s disease (AD). fMRI has several potential advantages, particularly for clinical trials, as it is a noninvasive imaging technique that does not require the injection of contrast agent or radiation exposure and thus can be repeated many times during a longitudinal study (Atri et al., 2011). fMRI has relatively high spatial and reasonable temporal resolution, and can be acquired in the same session as structural magnetic resonance imaging (MRI). Perhaps most importantly, fMRI may provide useful information about the functional integrity of brain networks supporting memory and other cognitive domains, including the neural correlates of specific behavioral events, such as successful versus failed memory formation (Brewer et al., 1998; Miller et al., 2008a; Sperling et al., 2003b; Wagner et al., 1998). However, there are very limited published data on fMRI test-retest or cross-scanner platform reproducibility, or correlation with longitudinal clinical outcome, and the majority of fMRI studies performed to date have enrolled small, highly selected cohorts within single academic centers.

2. BOLD fMRI techniques

Blood oxygen level dependent (BOLD) fMRI is an indirect measure of neuronal activity, thought to reflect the integrated synaptic activity of neurons via magnetic resonance (MR) signal changes due to changes in blood flow, blood volume, and the blood oxyhemoglobin/deoxyhemoglobin ratio, inferred from measuring changes in BOLD MR signal (Kwong et al., 1992; Logothetis et al., 2001; Ogawa et al., 1990). Task fMRI studies typically compare MR signal during 1 condition with MR signal during a control task or baseline condition, either in blocks of stimuli (e.g., novel vs. familiar stimuli) or in event-related designs (e.g., stimuli that were correctly remembered compared with those that were forgotten). In addition to functional activation studies, there has been considerable interest in the intrinsic connectivity of brain networks during the resting state using BOLD fMRI techniques, often referred to as...
functional connectivity or fc-MRI. These techniques examine the correlation between the intrinsic oscillations or time course of BOLD signal between brain regions, and have revealed a number of brain networks that demonstrate coherence in the spontaneous activity of distributed nodes (Vincent et al., 2006).

3. fMRI studies in AD dementia

The majority of fMRI studies in AD dementia utilized episodic memory tasks to focus on the pattern of fMRI activation in hippocampus and related structures in the medial temporal lobe (MTL). These studies consistently report decreased hippocamal or parahippocamal activity during the encoding of new information (Golby et al., 2005; Grön et al., 2002; Hämäläinen et al., 2007; Kato et al., 2001; Machulda et al., 2003; Rémy et al., 2004; Rombouts et al., 2000; Small et al., 1999; Sperling et al., 2003a). AD-related alterations in the pattern of fMRI activation in neocortex have also been reported. A recent quantitative meta-analysis of both fMRI and fluorodeoxyglucose (FDG)-positron emission tomography (PET) memory activation studies of AD identified several regions as being more likely to show greater encoding-related activation in healthy older individuals than in persons with Alzheimer dementia (Schwindt and Black, 2009). These regions include the hippocampal formation, ventrolateral prefrontal cortex, precuneus, cingulate gyrus, and lingual gyrus. Interestingly, evidence of increased neural activity, particularly in prefrontal regions, has been observed in persons with AD dementia during task performance (Celone et al., 2006; Grady et al., 2003; Sperling et al., 2003a; Wierenga et al., 2011).

4. fMRI studies in “at-risk” subjects

Task fMRI studies in individuals at risk for AD dementia, including subjects with mild cognitive impairment (MCI) and genetic-at-risk have yielded much less consistent findings. Several studies have reported decreased MTL activity in MCI compared with healthy persons (Johnson et al., 2006; Machulda et al., 2003; Petrella et al., 2006; Small et al., 1999). A number of studies in symptomatic individuals at risk for AD dementia have also reported decreased MTL activity (Borghesani et al., 2008; Lind et al., 2006a, 2006b; Mondadori et al., 2007; Smith et al., 1999; Trivedi et al., 2006), but other studies report increased MTL activity in both individuals with MCI (Celone et al., 2006; Dickerson et al., 2004, 2005; Hämäläinen et al., 2007; Heun et al., 2007; Kircher et al., 2007) and in asymptomatic persons with genetic or family history risk factors (Bondi et al., 2005; Bookheimer et al., 2000; Filippini et al., 2009; Fleisher et al., 2005; Han et al., 2007; Quiroz et al., 2010; Seidenberg et al., 2009; Smith et al., 2002; Wishart et al., 2004). A common feature of the studies reporting increased fMRI activity is that these studies primarily included subjects who were still able to perform the fMRI tasks reasonably well. In particular, some event-related fMRI studies found that the hyperactivity was observed specifically during successful memory trials, providing support for the early hypothesis that the increased activity may serve as a compensatory mechanism in the setting of early Alzheimer pathology (Dickerson and Sperling, 2008; Sperling et al., 2009). However, more recent work also suggests that the hyperactivity may be a harbinger of impending hippocampal failure and rapid clinical decline (Sperling et al., 2010). Cross-sectional studies suggest that the hyperactivity may be present only at early stages of MCI followed by a loss of activation as cognitive impairment worsens which is similar to the pattern seen in individuals with Alzheimer dementia (Celone et al., 2006). Longitudinal clinical follow-up studies suggest that hyperactivity at baseline is a predictor of both rapid cognitive decline (Bookheimer et al., 2000; Dickerson et al., 2004; Miller et al., 2008b) and loss of hippocampal function (O’Brien et al., 2010).

The mechanistic underpinnings of MTL hyperactivation remain unclear. Potential mechanisms that may contribute to this phenomenon include cholinergic or other neurotransmitter upregulation (DeKosky et al., 2002); aberrant sprouting of cholinergic fibers (Hashimoto et al., 2003), inefficiency in synaptic transmission (Stern et al., 2004), increased calcium influx or excitotoxicity (Busche et al., 2008; Palop et al., 2007), or alterations in glutamatergic receptor (Rammes et al., 2011). Further research to determine the specificity of hyperactivation to stage of disease and task performance, the relationship to baseline perfusion and metabolism, and the association with imaging markers of molecular pathology, including amyloid deposition and neurotransmitter systems, is clearly needed to elucidate this phenomenon.

5. Functional alterations in large scale networks

Both lesion studies and functional imaging evidence suggests that memory function is subserved by a network of brain regions that involves the hippocampal memory system and a set of cortical regions, including the precuneus, posterior cingulate, lateral parietal, lateral temporal, and medial prefrontal regions. Collectively known as the “default network,” these regions typically decrease activity during memory encoding and other cognitively demanding tasks focused on processing of external stimuli (Buckner et al., 2008; Raichle et al., 2001). These default network regions that typically demonstrate beneficial deactivations during encoding actually activate during successful memory retrieval (Differences et al., 2006; Vannini et al., 2011). Interestingly, a consistent failure to modulate default network activity during encoding has been reported in both AD dementia and in individuals at risk for AD (Celone et al., 2006; Fleisher et al., 2009; Lustig et al., 2003; Petrella et al., 2007; Pihlajamäki et al., 2008; Pihlajamaki et al., 2010).
BOLD fMRI techniques can also be used to investigate spontaneous brain activity and the interregional correlations in neural activity during the resting state, clearly documenting that the brain is organized into multiple large-scale brain networks; which persist during sleep and anesthesia (Damoiseaux et al., 2006; Vincent et al., 2007). These networks support specific sensory and motor systems, as well as specific cognitive processes (Vincent et al., 2006). Of particular interest in AD, is the intrinsic connectivity of the default network. Both “seed-based” connectivity and independent component analytic (ICA) techniques have demonstrated robust intrinsic connectivity between cortical nodes of the default network, with somewhat less consistent results in connectivity with the hippocampus. Multiple groups have reported impaired intrinsic functional connectivity in the default network during the resting state in individuals with MCI and AD dementia (Bai et al., 2008; Greicius et al., 2004; Rombouts et al., 2005, 2009; Sorg et al., 2007) that is greater than the general age-related disruption of large-scale networks (Andrews-Hanna et al., 2007; Damoiseaux et al., 2008). A recent study that applied connectivity measures to task fMRI data found that disrupted default network connectivity in MCI subjects was predictive of “conversion” to AD dementia over several years (Petrella et al., 2011). Another recently developed analytic technique that probes whole brain functional connectivity or “cortical hubs” may also prove useful in AD. Recent studies suggest that the topography of cortical hubs in young subjects overlaps the anatomy of amyloid-β deposition detected on PET amyloid imaging (Buckner et al., 2009), and that whole brain connectivity is disrupted in amnestic MCI patients (Bai et al., 2011; Drzezga et al., 2011).

The default network regions that demonstrate aberrant task-related fMRI activity and dysconnectivity in MCI and AD dementia correspond to regions with high amyloid burden in AD patients (Buckner et al., 2005, 2009; Klunk et al., 2004). Recent studies demonstrate evidence of disrupted default network activity during memory tasks (Sperling et al., 2009) and at rest in cognitively normal older individuals with evidence of amyloid deposition on PET imaging (Drzezga et al., 2011; Hedden et al., 2009; Sheline et al., 2010; Sperling et al., 2009) suggesting a combination of molecular and functional imaging techniques markers may be particularly useful to track response to trials of antiamyloid or therapies in preclinical stages of AD (Sperling et al., 2011).

6. Potential for fMRI in AD drug development

fMRI, either during cognitive paradigms or during resting state, may hold the greatest potential in the rapid evaluation of novel pharmacological strategies to treat AD. Several studies in healthy young and older subjects suggest that fMRI can detect acute pharmacological effects on memory networks (Kukolja et al., 2009; Sperling et al., 2002; Thiel et al., 2001). To date, only a few small fMRI studies have demonstrated enhanced brain activation after acute or prolonged treatment with cholinesterase inhibitors in MCI and AD, although these studies were not conducted as typical double-blind, placebo-controlled trials (Bokde et al., 2009; Goekoop et al., 2004; Rombouts et al., 2002; Saykin et al., 2004; Shanks et al., 2007; Venneri et al., 2009). There are a number of challenges in performing longitudinal task fMRI studies in patients with AD because as dementia severity increases, individuals are less likely to be able to perform cognitive tasks or to avoid head motion while in the scanner. As mentioned above, resting fc-MRI studies may be more much more feasible in longer-term studies in symptomatic stages of AD, although unfortunately, all fMRI techniques are very sensitive to head motion. fMRI has recently been incorporated into a small number of investigator-initiated add-on studies to ongoing Phase II and Phase III trials, which should provide information regarding the potential utility of these techniques in clinical trials.

Additional validation studies of fMRI in at-risk and AD dementia patients are critically needed. The short-term reproducibility of BOLD signal changes within young healthy individuals during memory encoding tasks and resting fc-MRI is only moderately high (Meindl et al., 2010; Sperling et al., 2002; Zuo et al., 2010) and very few reproducibility studies in older and cognitively impaired subjects have been published to date (Clement and Belleville, 2009; Putcha et al., 2011). Resting functional connectivity MRI techniques may be particularly advantageous for use in multicenter AD clinical trials and natural history studies, as no special equipment is required, individuals do not have to be able to perform a cognitive task, and a single 6-minute run added to the end of a safety or volumetric MRI protocol may provide reproducible patterns of fMRI connectivity over time and across scanner platforms (Van Dijk and Sperling, 2011). One study suggested that resting connectivity fMRI techniques may even demonstrate a larger “effect size” than task fMRI in at-risk populations (Fleisher et al., 2009). Longitudinal functional imaging studies are needed to track the evolution of alterations in the fMRI activation pattern over the course of the cognitive continuum from healthy aging to preclinical AD, MCI, and ultimately, AD dementia. It is also important to evaluate the contribution of structural atrophy to changes observed with functional imaging techniques in neurodegenerative diseases. Ideally, studies employing combinations of imaging modalities, such as structural MRI, fMRI, fluorodeoxyglucose-PET, and PET amyloid imaging techniques, will serve to further our understanding the interrelationships of these markers and their relative value in tracking change along the clinical continuum of AD (Jack et al., 2010). Such data may come in part from the Dominantly Inherited Alzheimer Network (DIAN) study of autosomal dominant AD that incorporates fc-MRI into its standard acquisition and from the continuation of the Alzheimer’s Disease Neuroimaging Initiative (ADNI)-2.
that includes fc-MRI on a limited number of scanners, and in other at-risk cohorts around the world. In summary, although both task and resting fMRI have been valuable in elucidating the neural basis of AD-related memory dysfunction, additional work to validate fMRI as a potential biomarker for use in clinical trials is critically needed. It is likely that task fMRI may have the greatest utility in early “proof of concept” studies, to detect an efficacy signal over a relatively short time frame. Recent work using fc-MRI during the resting state, which does not require special equipment or ability to perform a task, suggests that these techniques may be particularly amenable to use in multicenter clinical trials. As the field moves toward diagnosis and intervention at earlier stages of the AD process, even prior to clinically evident symptoms, the combination of amyloid biomarkers and fMRI may prove increasingly useful to detect evidence of early AD-related brain dysfunction and to monitor response to pharmacological treatment.

Disclosure statement

The author discloses no conflicts of interest.

References


Fluorodeoxyglucose positron emission tomography: emerging roles in the evaluation of putative Alzheimer’s disease-modifying treatments

Eric M. Reiman*, on behalf of the Alzheimer’s Disease Biomarkers Working Group for the Alliance for Aging Research

Abstract

Alzheimer’s disease (AD) is associated with characteristic and progressive reductions in fluorodeoxyglucose positron emission tomography (FDG PET) measurements of the regional cerebral metabolic rate for glucose. These reductions begin years before the onset of symptoms, are correlated with clinical severity, and may help predict an affected patient’s clinical course and neuropathological diagnosis. Like several other AD biomarkers, FDG PET has the potential to accelerate the evaluation of AD-modifying treatments, particularly in the earliest clinical and preclinical stages. This article considers FDG PET’s role in the detection and tracking of AD, its emerging roles in the evaluation of disease-slowing treatments, some of the issues involved in the acquisition, analysis, and interpretation of FDG PET data, and the evidence needed to help qualify FDG PET and other biomarkers for use in the accelerated approval of AD-slowing treatments. It recommends scientific strategies and public policies to further establish the role of FDG PET and other AD biomarkers in therapeutic trials and find demonstrably effective disease-modifying and presymptomatic AD treatments as quickly as possible.

© 2011 Elsevier Inc. All rights reserved.

Keywords: Alzheimer’s disease; Brain imaging; Biomarkers; Clinical trials; Positron emission tomography; Magnetic resonance imaging; Cerebrospinal fluid; Dementia; Mild cognitive impairment; Preclinical; Presymptomatic; Treatment; Prevention

1. Introduction

As noted in the other Alzheimer’s Disease (AD) Biomarkers Working Group articles, there is an urgent need to find demonstrably effective treatments to slow the progression of AD and a growing number of promising but unproven AD-modifying treatments that need to be evaluated. Right now, it takes too many research participants, too much money, and too much time to evaluate investigational AD-modifying treatments using clinical endpoints, particularly in the earliest clinical and preclinical stages of AD, when some of these treatments are likely to have their most profound benefit. The field urgently needs both the means and accelerated regulatory approval pathway to evaluate these treatments in the most rapid, cost-effective, and sufficiently rigorous way. Among other things, biomarker measurements of AD pathology and progression have the potential to reduce the number of research participants in randomized clinical trials (RCTs) and reduce the duration of these trials, particularly in the earliest clinical and preclinical stages of AD. The most promising biomarkers for the evaluation of putative AD-slowing treatments include volumetric magnetic resonance imaging (MRI) measurements of brain shrinkage, fluorodeoxyglucose positron emission tomography (FDG PET) measurements of regional reductions in the cerebral metabolic rate for glucose (CMRgl), PET measurements of fibrillar amyloid-β (Aβ) burden, and cerebrospinal fluid (CSF) measurements of low Aβ42 concentrations, high total tau and phospho-tau concentrations, and high total- or phospho-tau-to-Aβ42 ratios. These measurements have increasingly important, complementary, and converging roles in the evaluation of AD-slowing treatments.

In this review, we briefly consider FDG PET’s established role in the detection and tracking of AD, its emerging
roles in the evaluation of AD-slowing treatments in each of these stages, and some of the issues and uncertainties that need to be considered in the acquisition, analysis, and interpretation of FDG PET data for these purposes. We also consider the evidence needed to qualify FDG PET measurements for use as a surrogate endpoint, i.e., a biomarker considered reasonably likely to predict a clinical outcome—in the accelerated approval of AD-slowing treatments. Finally, we offer scientific and public policy recommendations to further establish the role of FDG PET and other AD biomarkers in the evaluation of AD-modifying treatments, galvanize the evaluation of investigational treatments in not only the clinical but preclinical stages of AD, and find demonstrably effective clinical and preclinical AD treatments as quickly as possible. For a more detailed discussion of these issues, please see Reiman and Langbaum (2009), and Reiman et al. (2010).

2. An established biomarker for the early detection and tracking of AD

FDG PET is the best established functional brain imaging technique for the detection and tracking of AD. AD is associated with preferential CMRgl reductions in precuneus and posterior cingulate, parietal and temporal cortex, beginning before the onset of symptoms, and extending into frontal cortex and whole brain in the later symptomatic stages of the disorder (Alexander et al., 2002). (A smaller number of studies have reported preferential CMRgl reductions in entorhinal cortex, hippocampal and medial temporal regions of interest [ROIs].) The CMRgl reductions are thought to reflect reductions in the activity or density of terminal neuronal fields or perisynaptic glials cells, mitochondrial or other metabolic dysfunctions, or a combination of these factors. (They are not solely attributable to the combined effects of brain atrophy and partial volume averaging.) Thus, FDG PET measurements are thought to provide information about AD-related synaptic dysfunction or loss, a downstream event in the pathogenesis of AD thought to be most closely related to cognitive impairment, thus complementing the information provided by other downstream biomarker measurements (e.g., MRI measurements of brain shrinkage and increases in CSF total tau and phospho-tau), and information about fibrillar Aβ burden, which may be an earlier event that is less well correlated quantitatively with cognitive decline, at least in the symptomatic stages of AD.

In patients with Alzheimer’s dementia, the CMRgl reductions are correlated with clinical severity, predict subsequent clinical decline, and the neuropathological diagnosis of AD with about 84%–93% sensitivity and about 73% specificity (Jagust et al., 2007; Silverman et al., 2001), and continue to decline over time. In patients with mild cognitive impairment (MCI), the CMRgl reductions, alone or in conjunction with other information (such as apolipoprotein E [APOE] ε4 carrier status or smaller hippocampal volumes) have shown the potential to predict subsequent rates of progression to Alzheimer’s dementia. Characteristic and progressive CMRgl reductions have also been reported in cognitively normal late-middle-aged carriers of the APOE ε4 allele, the major genetic risk factor for late-onset AD; baseline reductions and 2-year declines were correlated with APOE ε4 gene dose (reflecting 3 levels of genetic risk for AD), were apparent before evidence of hippocampal shrinkage, and were found at roughly the same ages in which PET and CSF evidence of Aβ pathology have been reported in other studies. While some of the CMRgl reductions have been reported in young adult APOE ε4 carriers almost 5 decades before the estimated average age at clinical onset, may even be developmental, and may not progress further until middle age, these reductions predict some of the brain regions associated with the earliest progressive CMRgl decline and fibrillar Aβ burden at older ages. CMRgl declines in cognitively normal carriers of certain early-onset AD-causing mutations and in cognitively normal older people who subsequently show cognitive decline, even after controlling for their APOE genotype. Analyses from a small single-center study and the larger multicenter AD Neuroimaging Initiative (ADNI) support the possibility that FDG PET could be used as an endpoint to evaluate the efficacy of putative AD-modifying treatments in a fraction of the Alzheimer’s dementia patients, MCI patients, and cognitively normal subjects at genetic risk for AD who would be needed to evaluate the treatment using clinical endpoints. Indeed, one could argue that FDG PET and other biomarkers are critically needed to provide a sufficiently rapid and cost-effective way to evaluate these treatments in the preclinical stages of AD.

3. Emerging roles in therapeutic trials

3.1. Use as a therapeutic trial endpoint

Like several of the other AD biomarkers described in our Working Group articles, FDG PET could be used as an endpoint to reduce the number of Alzheimer’s dementia, MCI, and cognitively normal at-risk subjects in trials and the time to evaluate putative AD-slowing treatment effects. For instance, we used longitudinal natural history data from ADNI to estimate a need for about 70 Alzheimer’s dementia patient completers per group to detect a 25% treatment effect (in this case, a slowing of CMRgl decline in an empirically predetermined statistical region of interest) with 80% power and 2-tailed \( p = 0.05 \) in a 12-month, multicenter, parallel-group, placebo-controlled RCT (Chen et al., 2010). This number (70) is roughly comparable to the number of completers needed using the best MRI-based measurements of regional or whole brain shrinkage, and it is a fraction of the approximately 600 completers per group estimated to detect a 25% treatment effect using the com-
monly used AD Assessment Scale—Cognitive (ADAS-Cog). We used ADNI longitudinal data to estimate a need for about 220 MCI patient completers per group to detect a 25% treatment effect with 80% power and 2-tailed $p = 0.05$ in a 12-month, multicenter, parallel-group, placebo-controlled RCT, again roughly comparable to the number of patients needed using the best MRI-based measurements of regional or whole brain shrinkage, and a fraction of the approximately 4400 completers estimated to detect a 25% treatment effect using the AD Assessment Scale—Cognitive (Chen et al., 2010). We used longitudinal data from a study of initially late-middle-aged cognitively normal APOE $e4$ homozygotes, heterozygotes, and noncarriers to estimate the need for fewer than 200 cognitively normal APOE $e4$ homozygote or heterozygote completers per group to detect a 25% treatment effect with 80% power and 2-tailed $p = 0.05$ in a 24-month parallel-group, placebo-controlled RCT (Reiman et al., 2001; Reiman et al., unpublished data), thus providing the potential to evaluate a range of putative AD-modifying treatments in the clinical stages of AD without having to study many thousands of research subjects or wait many years to detect a clinical benefit.

It is important to note several caveats and lingering questions, almost all of which could be addressed as FDG PET and the other promising biomarkers are embedded into clinical trials. First, the sample-size estimates noted above were based on data from a limited number of subjects and have a large confidence interval. Second, the estimates are at least partly based on the specific image analysis technique used, and researchers continue to develop, test, and compare numerous image analysis techniques in terms of their statistical power and freedom from the type I error associated with multiple regional comparisons. Third, with the increasing use of FDG PET in clinical trials, it will be possible to determine the extent to which different treatments can slow the decline in FDG PET and other biomarker measurements (some of which may be harder to budge than others) and the extent to which the treatment’s biomarker effects are reasonably likely to predict a clinical outcome—the evidence needed for regulatory agencies to qualify FDG PET for use (as an unvalidated but reasonably likely surrogate endpoint) in the accelerated approval of AD treatments. Fourth, it is important to anticipate and prepare for the possibility that a treatment might have a confounding effect on FDG PET or other biomarkers of interest unrelated to disease slowing (e.g., an effect on synaptic activity, metabolism, or density unrelated to synaptic loss in the case of FDG PET or an effect on brain swelling unrelated to brain atrophy in the case of volumetric MRI). To help address this issue, we have proposed the acquisition of additional FDG or MRI images shortly after a treatment is started or discontinued to help address treatment effects unrelated to disease-slowing as well as the use of complementary biomarkers to help overcome any modality-specific confounding effect.

### 3.2. Use in subject selection and subgroup analyses

In addition to using information from sequential scans as an endpoint in therapeutic trials, FDG PET could also be used to help select research participants for enrollment or subgroup analyses in clinical trials of putative AD-modifying treatments. For instance, FDG PET measurements could be used alone or in combination with other information to select those MCI patients at highest risk for clinical progression to Alzheimer’s dementia, further reducing the sample size and treatment duration needed to detect treatment effect (Chen et al., 2011; Landau et al., 2010). It may also have a role in identifying those patients most likely to benefit from an AD-slowing treatment, reducing attrition in the drug product’s development. Consider, for instance, the evaluation of an Aβ-modifying therapeutic agent in cognitively normal people with PET or CSF evidence of Aβ pathology: one could make a plausible case that those individuals who also show FDG evidence of significant synaptic pathology may show a preferential response to treatment during the trial due to a higher risk of subsequent clinical decline or, conversely, a preferential resistance to the treatment due to the extent to which the downstream events have already ravaged the brain.

### 4. Methodological issues

In order to maximize statistical power, the ability to compare findings from different studies, and freedom from potentially confounding effects (like a sensory or task-dependent increase in local neuronal activity), ADNI researchers developed standard operating procedures for the acquisition of FDG PET scans, some of which have been further developed for use in clinical trials. For instance, they have proposed the acquisition of images in the resting state (eyes open and directed forward in a dark room with minimal sensory stimulation) to minimize potentially confounding effects of sensory and motor activity of regional CMRgl, use of phantom data to qualify PET systems for use in clinical trials, standardized radiotracer uptake and dynamic scanning periods, manufacturer-dependent image reconstruction and attenuation-correction algorithms, real time quality assessment and quality control, the centralized standardization of PET images to a common spatial resolution, and strategies to minimize the likelihood of scanner changes during the study. Researchers continue to develop, test, and compare different image analysis techniques in terms of their ability to predict clinical progression, track CMRgl decline, and evaluate AD-modifying treatments with the best statistical power and freedom from the type I error associated with multiple comparisons. As previously noted, we recommend the use of additional scans and complementary biomarker endpoints to help address the potentially confounding effects of treatment on different biomarker measurements.
5. Scientific and public policy recommendations

Despite the expense, we recommend the use of multiple brain imaging and CSF biomarker endpoints in clinical trials for these reasons: (1) to provide converging evidence in support of a treatment’s AD-modifying effects; (2) to help overcome the possibility of modality-specific confounding effects on any individual biomarker measurement (e.g., it might have been useful in characterizing the AD-modifying effects of the investigational Aβ immunization therapy AN1792 in the face of its possibly confounding effects on volumetric MRI measurements of brain shrinkage); (3) to help address different questions and anticipate the possibility that a combination of biomarker effects (e.g., Aβ-modifying plus 1 or more downstream biomarker effects) may be needed to predict a treatment’s clinical benefit; and (4) to provide additional evidence to qualify 1 or more of these biomarker measurements for use in therapeutic trials, including those trials initiated before symptoms, when the use of clinical endpoints is not practical and when some of the treatments now in development may be likely to have their most profound clinical benefit. Because FDG PET has been shown to characterize AD-related CMRgl declines in the preclinical stages of AD, its inclusion in clinical trials would not only advance the evaluation of the particular investigational treatment but also help provide the evidence needed to suggest its use for the accelerated approval of preclinical AD treatments.

We recommend testing potentially disease-modifying treatments in the earliest clinical stage or even the preclinical stage of the disorder, not only to be able to observe a benefit but also to provide evidence that a treatment’s biomarker effects are reasonably likely to predict a clinical benefit. The field urgently needs both the means and the accelerated regulatory approval pathway to find demonstrably effective clinical and preclinical AD treatments as quickly as possible.

Disclosure statement

The author discloses no conflicts of interest.

Acknowledgements

The author has received relevant grant support from the National Institute on Aging (RO1 AG031581 and P30 AG19610), the state of Arizona, and the Banner Alzheimer’s Foundation.

References


Alliance for Aging Research AD Biomarkers Work Group: structural MRI
Clifford R. Jack, Jr.
Mayo Foundation, Rochester, MN, USA

Abstract
Biomarkers of Alzheimer’s disease (AD) are increasingly important. All modern AD therapeutic trials employ AD biomarkers in some capacity. In addition, AD biomarkers are an essential component of recently updated diagnostic criteria for AD from the National Institute on Aging—Alzheimer’s Association. Biomarkers serve as proxies for specific pathophysiological features of disease. The 5 most well established AD biomarkers include both brain imaging and cerebrospinal fluid (CSF) measures—cerebrospinal fluid Abeta and tau, amyloid positron emission tomography (PET), fluorodeoxyglucose (FDG) positron emission tomography, and structural magnetic resonance imaging (MRI). This article reviews evidence supporting the position that MRI is a biomarker of neurodegenerative atrophy. Topics covered include methods of extracting quantitative and semiquantitative information from structural MRI; imaging-autopsy correlation; and evidence supporting diagnostic and prognostic value of MRI measures. Finally, the place of MRI in a hypothetical model of temporal ordering of AD biomarkers is reviewed.

© 2011 Published by Elsevier Inc.

Keywords: Alzheimer’s disease; Biomarkers; MRI

1. How are measures of neurodegenerative atrophy extracted from structural MRI images?

The topographic pattern of neurodegenerative atrophy in Alzheimer’s disease (AD) captured by anatomic magnetic resonance imaging (MRI) mirrors that of neurofibrillary pathology (Braak and Braak, 1991; Whitwell et al., 2007, 2008a). Atrophy begins in and is ultimately most severe in the medial temporal lobe, particularly the entorhinal cortex and hippocampus, which is why these structures have been targeted in many MRI studies for diagnostic purposes. Atrophy later spreads to the inferior temporal lobe and paralimbic cortical areas. The transition from mild cognitive impairment (MCI) to full dementia is felt to be due to spread of degenerative atrophy to multimodal association neocortices. Below is a brief survey of methods to extract and/or visualize this information from 3-D MRI scans of cross-sectional and longitudinal studies (modified from Vemuri and Jack, 2010).

1.1. Cross-sectional methods

1.1.1. Visual assessment of scans

Visual assessment of the degree of atrophy in the medial temporal lobe is often used to assess disease severity and to add confidence in a clinical diagnosis of AD (Scheltens et al., 1992). Fig. 1 shows the medial temporal lobe in cognitively normal elderly (CN), amnestic mild cognitive impairment (aMCI), and AD. While simple visual assessment is easily implemented and widely available, atrophy is a continuous process and this method does not lend itself to accurate or reproducible assessment of fine incremental grades of atrophy.

1.1.2. Quantitative ROI-based techniques

Manual tracing and quantifying the volume of medial temporal lobe structures, e.g., the hippocampus or entorhinal cortex has been traditionally employed and provides an accurate quantitative measure of atrophy but is time-consuming (Fox et al., 1996; Jack et al., 1992).
1.1.3. Automated and semiautomated techniques

Methods have been developed to automatically parcelate gray matter density (Tzourio-Mazoyer et al., 2002) or the thickness of cortical surfaces (Dale et al., 1999; Fischl et al., 2004) into regions of interest. This is computationally intensive but is reproducible and does not require manual intervention.

An advantage of measuring something like the hippocampus is that the measurements describe a known anatomic structure that (in the case of the hippocampus) is closely related to the pathological expression of the disease and is also functionally related to 1 of the cardinal early clinical symptoms—memory impairment. The disadvantage of using a single structure or region of interest (ROI) to consolidate 3-D information is that it is topographically limited and does not make use of all the available information in a 3-D MRI.

1.1.4. Quantitative voxel-based

These methods assess atrophy over the entire 3-D MRI scan.

1.1.4.1. Voxel-based analytic techniques. Methods such as voxel-based morphometry (VBM) (Ashburner and Friston, 2000) are a popular and useful way to test for group-wise differences in the topography of atrophy. However, the statistical testing portion of voxel-based morphometry is not designed to provide diagnostic information at the single subject level.

1.1.4.2. Automated individual subject diagnosis. Several investigators have developed multivariate analysis and machine learning-based algorithms which use the entire 3-D MRI data to form a disease model against which individual subjects may be compared. A new incoming scan is scored based on the degree and the pattern of atrophy in comparison with the scans of a large database of well characterized subjects (Alexander and Moeller, 1994; Csermisky et al., 2000; Davatzikos et al., 2009; Fan et al., 2005; Koppel et al., 2008; Stonnington et al., 2008; Vemuri et al., 2008a; Welch et al., 2002). Such measures capture the severity of neuronal pathology, i.e., Braak staging, better than hippocampal volumes (Vemuri et al., 2008b).

1.2. Longitudinal methods

While change over time can be determined by simply measuring a volume independently on each scan in a series and performing arithmetic subtraction of the volumes, more sophisticated techniques have been developed to extract tissue loss information from serial MRI scans. In these techniques all MRI scans within a subject’s time series are registered to each other and brain loss between scans is quantified as a measure of neurodegenerative disease progression.

1.2.1. Global atrophy quantification

One of the earliest methods developed to quantify the global change in brain volume between 2 scans was the boundary shift integral (BSI) (Fox and Freeborough, 1997; Freeborough and Fox, 1997). BSI determines the total volume through which the surface of the brain has moved between scans acquired at 2 time points, i.e., as the brain volume decreases and the volume of the ventricles increases.

1.2.2. Tensor-based morphometry (TBM)

Unlike BSI which only analyzes spatial shift in the brain surfaces, TBM provides 3-D patterns of voxel-level brain degeneration (Chételat et al., 2005; Thompson and Apostolova, 2007).

2. Evidence validating MRI as a neurodegenerative biomarker in AD

Evidence validating MRI as a neurodegenerative AD biomarker is reviewed below. Studies are classified on several criteria, including the method of measurement, numbers of subjects, and source of subjects. The ideal source is an epidemiological or population-based cohort. The next best option is a community-based sample. The least desirable but most common source of data are referral samples, which have the highest risk of biases. Evidence validating MRI as an AD biomarker takes the form of several different types of studies: cross-sectional clinical-MRI correlations; prediction of future clinical change; correlating change-over-time on serial MRI with concurrent change on clinical indexes; and MRI-autopsy correlation.
2.1. Cross sectional clinical-MRI correlations

Many studies have been published describing the accuracy, sensitivity, specificity, or area under receiver operating characteristic curve (AUROC) with which clinically diagnosed AD subjects can be separated from cognitively normal elderly control subjects. This is the simplest type of data to acquire and hence this is the most frequent type of study found in the literature. This is the weakest category of validation data, because the gold standard against which the MRI is compared is a clinical diagnosis, which can be wrong. A clinical diagnosis is also available in the absence of any biomarker data. Accuracy ranges from 85% to 100%. Different methods have been employed as described above. The literature is too vast to describe each publication, but Table 1 contains some representative examples of studies demonstrating cross-sectional separation of clinically diagnosed AD versus controls. Results vary depending on measurement method, source of subjects, and statistical endpoints.

A related class of studies is those that demonstrate cross-sectional separation of clinically diagnosed controls versus subjects with mild cognitive impairment. Mild cognitive impairment may have been defined using the formal diagnostic criteria for MCI outlined by Petersen (2004) or may have been defined using other criteria. Table 2 contains some representative examples of studies demonstrating cross-sectional separation of clinically diagnosed controls versus subjects with mild cognitive impairment using quantitative MRI measures. Results vary depending on measurement method, source of subjects, and statistical endpoints.

2.2. Autopsy-MRI correlation

MRI-autopsy studies have convincingly validated that quantitative measurements of brain volume loss correlate with pathological indexes of neurodegenerative severity. Hippocampal volumes measured from antemortem MRI scans correlate with Braak neurofibrillary tangle pathologic staging in both demented and nondemented subjects (Gosche et al., 2002; Jack et al., 2002). Antemortem hippocampal volume as well as rates of brain and hippocampal atrophy from MRI correlate with hippocampal neurofibrillary tangle density (Csernansky et al., 2004; Silbert et al., 2003) at autopsy. Excellent correlation is found between hippocampal volume measures obtained on either antemortem MRI (Zarow et al., 2005) or postmortem MRI (Bobinski et al., 2000) and hippocampal neuron cell counts in autopsy specimens. On the basis of these imaging-to-pathology correlation studies, quantitative measures from structural MRI, such as hippocampal volume, are inferred to represent an approximate surrogate of the stage/severity of neuronal pathology—neuron loss, neuron shrinkage, and synapse loss—that occurs in AD. Voxel-wise studies of gray matter density also show useful diagnostic accuracy, especially in specific brain regions such as the hippocampus. The literature is too vast to describe each publication, but Table 2 contains some representative examples of studies demonstrating cross-sectional separation of clinically diagnosed mild cognitive impairment versus controls.

Table 1

<table>
<thead>
<tr>
<th>Study</th>
<th>Subjects</th>
<th>Source of subjects</th>
<th>Measurement method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desikan et al., 2009</td>
<td>CN 94, AD 65</td>
<td>Referral sample</td>
<td>Ctx thickness, ERC + Hipp + SupMarg gyrus</td>
<td>AUROC 1.0</td>
</tr>
<tr>
<td>Gerardin et al., 2009</td>
<td>CN 25, AD 23</td>
<td>Referral sample ADNI</td>
<td>Hippocampal shape metric</td>
<td>Sensitivity 96%, specificity 92%</td>
</tr>
<tr>
<td>Hinrichs et al., 2009</td>
<td>CN 94, AD 89</td>
<td>Referral sample ADNI</td>
<td>Multi voxel classifier</td>
<td>AUROC 0.88</td>
</tr>
<tr>
<td>Jack et al., 1992</td>
<td>CN 22, AD 20</td>
<td>Community sample</td>
<td>Manual hippocampal volume adjusted for head size and age</td>
<td>Sensitivity 95%, specificity 95%, accuracy 89%, AUROC 0.92</td>
</tr>
<tr>
<td>Killiany et al., 2000</td>
<td>CN 24, AD 16</td>
<td>Referral sample</td>
<td>ERC, banks of superior temporal sulcus, anterior cingulate</td>
<td>Accuracy 100%</td>
</tr>
<tr>
<td>Kohannim et al., 2010</td>
<td>CN 213, MCI 158</td>
<td>Referral sample ADNI</td>
<td>Multi voxel classifier</td>
<td>AUROC 0.89</td>
</tr>
<tr>
<td>McEvoy et al., 2009</td>
<td>CN 139, AD 84</td>
<td>Referral sample ADNI</td>
<td>Ct thickness; medial and lateral temporal, isthmus cingulated orbitofrontal</td>
<td>Sensitivity 83%, specificity 93%</td>
</tr>
<tr>
<td>Walhovd et al., 2010</td>
<td>42 CN, 38 AD</td>
<td>Referral sample ADNI</td>
<td>Ct thickness</td>
<td>Accuracy 85%</td>
</tr>
</tbody>
</table>

Key: AD, Alzheimer’s disease; ADNI, Alzheimer’s Disease Neuroimaging Initiative; AUROC, area under receiver operating characteristic curve; CN, cognitively normal.

Table 2

<table>
<thead>
<tr>
<th>Study</th>
<th>Subjects</th>
<th>Source of subjects</th>
<th>Measurement method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desikan et al., 2009</td>
<td>CN 94, MCI 57</td>
<td>Referral sample</td>
<td>Ctx thickness of ERC + Hipp + SupMarg gyrus</td>
<td>AUROC 0.95, sensitivity 90%, specificity 91%</td>
</tr>
<tr>
<td>Gerardin et al., 2009</td>
<td>CN 25, MCI 23</td>
<td>Referral sample ADNI</td>
<td>Hippocampal shape metric</td>
<td>Sensitivity 83%, specificity 84%</td>
</tr>
<tr>
<td>Kohannim et al., 2010</td>
<td>CN 213, MCI 264</td>
<td>Referral sample ADNI</td>
<td>Multi voxel classifier</td>
<td>AUROC 0.84</td>
</tr>
<tr>
<td>Xu et al., 2000</td>
<td>CN 30, MCI 30</td>
<td>Community sample</td>
<td>Hippocampal W score</td>
<td>Sensitivity 63%, specificity 80%</td>
</tr>
</tbody>
</table>

Key: ADNI, Alzheimer’s Disease Neuroimaging Initiative; AUROC, area under receiver operating characteristic curve.
loss demonstrate that the topographic distribution of gray matter loss closely mirrors Braak and Braak spatial distribution of neurofibrillary pathology in subjects who have had antemortem MRI and have come to autopsy (Fig. 2) (Whitwell et al., 2008). Fully automated multivoxel analysis methods demonstrate close correlation between quantitative antemortem MRI and Braak staging, as depicted in Fig. 3 with STructural Abnormality iNDex (STAND) scores.

We point out that while MRI measures of atrophy do scale with pathological indexes of neurodegeneration, brain atrophy is not specific for AD. It occurs in other conditions that may be associated with cognitive impairment, such as cerebrovascular disease, hippocampal sclerosis, frontotemporal lobar degeneration, and head trauma (Jack et al., 2002; Jagust et al., 2008; Zarow et al., 2005).

3. Modeling of the longitudinal trajectory of AD with biomarkers—where does structural MRI fit?

Because different AD biomarkers provide information about different AD-related pathological processes, it stands to reason that comprehensive in vivo assessment of the disease requires information from different classes of biomarkers. Based on the assumptions that MRI provides an index of neurodegenerative pathologic burden (above) and Pittsburgh Compound B (PIB) positron emission tomography (PET) a measure of amyloid plaque burden, a model of AD has been proposed in which the rate of amyloid deposition and the rate of neurodegeneration later in life are dissociated. The presence of brain amyloidosis is necessary but not sufficient to produce cognitive decline; the neurodegenerative component of AD pathology is the immediate substrate of cognitive impairment, and the rate of cognitive decline is driven by the rate of neurodegeneration. In this proposed model, amyloid deposition is dynamic early in the disease process (presymptomatically) while neurodegeneration is dynamic in the mid- to late-stage. This amyloid and neurodegeneration model (Jack et al., 2009) is reproduced in Fig. 4. In the model, the lifetime course of the disease is divided into clinically defined presymptomatic, early symptomatic (MCI), and dementia phases. Neurodegeneration, detected by atrophy on volumetric MRI, is indicated by a dashed line. Cognitive function is indicated by a dot-dash line. Amyloid deposition, detected by PIB, is indicated by a solid line later in the course of AD (i.e., that portion of the disease for which PIB data are now available). The time course of amyloid deposition early in life is represented as 2 possible theoretical trajectories (dotted lines), reflecting uncertainty about the time course of early PIB signal.

An expanded version of this disease biomarker model (Jack et al., 2010a) incorporates the 5 most well validated AD biomarkers into a comprehensive sequence of pathological events as subjects progress from cognitively normal
in middle age to dementia in older age. There are presently 5 well-accepted biomarkers of AD. Both cerebrospinal fluid (CSF) Aβ42 and amyloid PET imaging are biomarkers of Aβ plaque deposition. CSF tau is an indicator of tau pathology and associated neuronal injury. Fluorodeoxyglucose (FDG) PET measures AD-mediated neuronal dysfunction, while structural MRI measures AD-mediated neurodegeneration. This model rests on the assumption that these 5 AD biomarkers become abnormal in a sequential manner, but their time courses also overlap. The hypothesis is that amyloid PET imaging and CSF Aβ42 become abnormal first, perhaps as much as 20 years before the first clinical symptoms appear. CSF tau and FDG PET become abnormal later and structural MRI is the last of the 5 major biomarkers to become abnormal. CSF tau, FDG PET, and structural MRI correlate with clinical symptom severity while CSF Aβ42 and amyloid PET imaging may not. The hypothesis is that together these 5 biomarkers of AD are able to stage the complete trajectory of AD, which may span as much as 20–30 years or more in affected individuals. Fig. 5 illustrates this expanded model (Jack et al., 2010a).

### 4. Use of MRI in therapeutic trials

MRI is used in several different ways in therapeutic trials. Therapeutic modification of the natural rate of atrophy has been used as an outcome measure in a number of AD and MCI trials. As a measure of the severity or stage of neurodegeneration, MRI has been used as a covariate in analyses, much the same way disease severity on clinical scales like the Mini Mental State Examination (MMSE) or AD Assessment Scale—Cognitive (ADAS—Cog) is used. In theory MRI can also be used to stratify trial subjects at baseline on the basis of disease severity. Although the discussion above has focused on structural MRI as a measure of the severity of AD-related neurodegeneration, MRI also is commonly used for inclusion/exclusion purposes in therapeutic trials. For example, hemispheric cerebral infarction, tumor, normal pressure hydrocephalus (NPH), prior surgery, major head trauma, and cerebral hemorrhage are common exclusionary findings on screening MRI. Micro hemorrhages that exceed a prespecified number are also a common exclusionary finding in antiamyloid trials. The major barrier to the use of volumetric MRI as an outcome measure in clinical trials has been lack of standardization of MRI methods, particularly methods for extracting quantitative information from scans. This lack of standardization leads to different results (Tables 1–4), which in turn undermines the credibility of the method in the minds of regulators. Although initiatives such as the Alzheimer’s Disease Neuroimaging Initiative (ADNI) have focused on standard-
<table>
<thead>
<tr>
<th>Study</th>
<th>Subjects</th>
<th>Source of subjects</th>
<th>Measurement method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predicting progression from mild cognitive impairment to AD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bakkour et al., 2009</td>
<td>49 CDR 0.5</td>
<td>Referral sample</td>
<td>Cortical thickness in temporal and parietal ROIs</td>
<td>Predict MCI progression to AD, 83% sensitivity and 65% specificity</td>
</tr>
<tr>
<td>Brys et al., 2009</td>
<td>24 MCI</td>
<td>Referral sample</td>
<td>Medial temporal lobe gray matter concentration</td>
<td>Accuracy, predict MCI progression to AD: 74%</td>
</tr>
<tr>
<td>Convit et al., 2000</td>
<td>46 Normal or MCI</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Declining subjects had 11.3% of reduction in HC compared with nondecliners</td>
</tr>
<tr>
<td>DeCarli et al., 2007</td>
<td>190 MCI</td>
<td>ADCS Vit E donepezil trial</td>
<td>Visual assessment of hippocampal atrophy</td>
<td>Atrophy score &gt; 2.0 increased likelihood of progression, HR: 2.30</td>
</tr>
<tr>
<td>Desikan et al., 2009</td>
<td>129 MCI</td>
<td>Referral sample ADNI</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD, adjusted HR: 0.73 (0.51–1.04)</td>
</tr>
<tr>
<td>Desikan et al., 2008</td>
<td>47 MCI</td>
<td>Referral sample</td>
<td>Temporal-parietal regions of interest</td>
<td>Combination of entorhinal cortex (HR = 0.60) and the inferior parietal lobule (HR = 0.62) was best predictor of time to progression to AD</td>
</tr>
<tr>
<td>Devanand et al., 2007</td>
<td>139 MCI</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD, HR: 2.84 (1.47–5.49)</td>
</tr>
<tr>
<td>Eckerström et al., 2008</td>
<td>42 MCI</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Hippocampal volumes smaller in converters to AD versus nonconverters</td>
</tr>
<tr>
<td>Fleisher et al., 2008</td>
<td>129 aMCI</td>
<td>ADCS Vit E donepezil trial</td>
<td>Ventricular volumes and hippocampal volumes</td>
<td>Ventricular volumes and hippocampal volumes predicted progression to AD</td>
</tr>
<tr>
<td>Galluzzi et al., 2010</td>
<td>90 MCI</td>
<td>Referral sample</td>
<td>Medial temporal atrophy</td>
<td>Predict MCI progression to AD, AUC: 0.73</td>
</tr>
<tr>
<td>Galton et al., 2005</td>
<td>31 CDR 0.5</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Converters had a greater atrophy compared with nonconverters.</td>
</tr>
<tr>
<td>Henneman et al., 2009</td>
<td>39 MCI</td>
<td>Referral sample</td>
<td>Hippocampal volume adjusted for age, sex, baseline MMSE</td>
<td>Predict MCI progression to AD, HR: 10.4 (3.1–34.8)</td>
</tr>
<tr>
<td>Herukka et al., 2008</td>
<td>21 MCI</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD, right HC: 15.8 (1.4–174.2)</td>
</tr>
<tr>
<td>Jack et al., 2010b</td>
<td>218 MCI</td>
<td>ADNI plus Mayo community sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD, HR 2.6 (1.8–3.8) 25% versus 75%</td>
</tr>
<tr>
<td>Jack et al., 2008</td>
<td>131 MCI</td>
<td>ADCS Vit E donepezil trial</td>
<td>Volumes of hippocampus, entorhinal cortex, brain, ventricle</td>
<td>Rates of change in all volumes were greater in converters than nonconverters</td>
</tr>
<tr>
<td>Jack et al., 2005</td>
<td>72 MCI</td>
<td>Community sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD, HC volume OR: 1.51 (1.1–2.0)</td>
</tr>
<tr>
<td>Jack et al., 2000</td>
<td>43 MCI</td>
<td>Community sample</td>
<td>Hippocampal volume</td>
<td>Rates of hippocampal atrophy were greater in converters than nonconverters</td>
</tr>
<tr>
<td>Jack et al., 1999</td>
<td>80 MCI</td>
<td>Community sample</td>
<td>Hippocampal W score</td>
<td>Relative risk 0.69—for each 1 unit increase in W score (less atrophy) risk of progression to AD decreased by 31%</td>
</tr>
<tr>
<td>Kantarci et al., 2005</td>
<td>21 MCI</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD OR: 2.5 (1.0–6.2)</td>
</tr>
<tr>
<td>Killiany et al., 2002</td>
<td>94 CDR 0.5</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD OR: 1.5 (1.0–2.31)</td>
</tr>
<tr>
<td>Landau et al., 2010</td>
<td>85 MCI</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD OR: 2.49 (1.02–5.96)</td>
</tr>
<tr>
<td>Leung et al., 2010</td>
<td>335 MCI</td>
<td>ADNI</td>
<td>Hippocampal volume</td>
<td>Rates higher in converters compared with stable and reverter groups</td>
</tr>
<tr>
<td>Risacher et al., 2009</td>
<td>227 MCI</td>
<td>ADNI</td>
<td>Hippocampal volume</td>
<td>Effect size for separating MCI stable versus converter Cohen’s d = 0.60</td>
</tr>
<tr>
<td>Stoub et al., 2010</td>
<td>29 aMCI</td>
<td>Referral sample</td>
<td>Entorhinal cortex and hippocampus</td>
<td>Atrophy rate of entorhinal cortex and hippocampus in controls less than MCI converters</td>
</tr>
<tr>
<td>Tapia et al., 2008</td>
<td>60 MCI</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD OR: total HC 0.815 (0.69–0.97)</td>
</tr>
<tr>
<td>Vemuri et al., 2009</td>
<td>192 MCI</td>
<td>Referral sample ADNI</td>
<td>STAND score</td>
<td>HR for time to conversion from MCI to AD 25th versus 75th percentile 2.6</td>
</tr>
<tr>
<td>Visser et al., 1999</td>
<td>13 MCI</td>
<td>Community sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD OR: 0.21 (0.05–0.99)</td>
</tr>
<tr>
<td>Visser et al., 2002</td>
<td>30 MCI</td>
<td>Community sample</td>
<td>Hippocampal volume</td>
<td>Hippocampal volume predicts MCI progression to AD</td>
</tr>
<tr>
<td>Wang et al., 2009</td>
<td>58 aMCI</td>
<td>Referral sample</td>
<td>Hippocampal volume</td>
<td>Predict MCI progression to AD left HC HR: 0.38 (0.10–0.88)</td>
</tr>
</tbody>
</table>

Key: AD, Alzheimer’s disease; ADNI, Alzheimer’s Disease Neuroimaging Initiative; aMCI, amnestic mild cognitive impairment; AUC, area under the curve; CDR, Clinical Dementia Rating; HR, hazard ratio; MCI, mild cognitive impairment; MMSE, Mini Mental State Examination; OR, odds ratio; ROIs, regions of interest; STAND, STructural Abnormality iNDex;
izing imaging methods, to date universally accepted standards for MRI image quantification have not emerged.

At the present time, AD biomarkers have not yet been validated as surrogate endpoints for regulatory purposes. However the impact of interventions on these biomarkers has been evaluated in a few trials and was found to be potentially useful in capturing the pharmacodynamic effects of an agent. The efficacy of donepezil, an acetylcholinesterase inhibitor, was evaluated using serial anatomic MRI (Hashimoto et al., 2005; Jack et al., 2008; Krishnan et al., 2003) and was found to possibly be neuroprotective based on some evidence of decreased rates of atrophy in the treatment versus placebo arms. In a different study, antibody responders immunized to amyloid-beta (A) had more rapid volume loss than placebo patients during a Phase IIa immunotherapy trial that was prematurely terminated due to meningoencephalitis in a small subset of patients (Fox et al., 2005).

5. Predicting the risk of progression in MCI and CN

About 12%–15% of MCI subjects annually progress to AD (Fischer et al., 2007; Petersen, 2007); however, clinical criteria alone cannot identify with certainty which subjects will progress more rapidly than others. For this reason, predictive information from imaging has been sought to supplement clinical prognostic indicators. Studies demonstrating the ability of MRI to predict future progression have taken several forms. Studies using time-to-event methods are appropriate when follow-up times vary among subjects in the cohort which is most commonly the case. Such studies typically employ Cox proportional hazards models in which cutoffs stratify a baseline MRI measurement into risk groups and the results are reported as hazard ratios (HR) (Jack et al., 1999). This type of analysis relates an imaging measurement to the time to progression from a diagnosis of MCI to AD, not to the lifetime risk of developing AD. A related method of analysis employs a rate of change at baseline as the predictor rather than a brain volume measurement at 1 point in time (Jack et al., 2005). If all subjects in the study have the same follow-up time, then simply comparing baseline MRI between progressors and nonprogressors is appropriate. Unfortunately, several studies have simply compared baseline MRI measures between progressors and nonprogressors when follow-up times were not the same across subjects in the cohort. Inferences about imaging as a predictor may be invalid in this situation because subjects classified as progressors may simply be those who have longer follow-up times than subjects classified as nonprogressors. Table 3 illustrates examples of studies evaluating the ability of baseline MRI measures to predict time to progression from MCI to AD. Results vary depending on measurement method, source of subjects, and statistical endpoints.

Table 4
Sample sizes per arm needed to power treatment study in AD/MCI

<table>
<thead>
<tr>
<th>Citation</th>
<th>Subjects</th>
<th>Source of subjects</th>
<th>Measurement method</th>
<th>Sample size required to detect treatment effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fox et al., 2000</td>
<td>18 AD</td>
<td>Referral sample</td>
<td>Classic BSI</td>
<td>207 per arm assuming: 1-year trial, 20% effect size, 90% power, 10% dropout, 10% unusable scans</td>
</tr>
<tr>
<td>Holland et al., 2009</td>
<td>129 AD;299 MCI</td>
<td>Referral sample ADNI</td>
<td>Ctx thickness ERC ROI</td>
<td>Assuming 24-mo trial, 25% effect size, 80% power, scans every 6 mo .45 per arm for AD; 135 per arm MCI</td>
</tr>
<tr>
<td>Hua et al., 2010</td>
<td>50 AD;MCI 122</td>
<td>Referral sample ADNI</td>
<td>TBM temporal lobe</td>
<td>Assuming 12-mo trial, 25% effect size, 80% power; 43 AD per arm; 82 MCI per arm</td>
</tr>
<tr>
<td>Jack et al., 2003</td>
<td>192 AD</td>
<td>Referral sample, terminated multisite therapeutic trial</td>
<td>Hippocampus</td>
<td>Assuming 12-mo trial, 50% effect size, 90% power at 0.05; 21 per arm for AD</td>
</tr>
<tr>
<td>Leung et al., 2010</td>
<td>81 AD</td>
<td>Referral sample ADNI</td>
<td>KN-BSI</td>
<td>Assuming 12-mo trial, 25% effect size, 80% power; 81 AD per arm</td>
</tr>
<tr>
<td>Schott et al., 2006</td>
<td>46 AD</td>
<td>Referral sample ADNI</td>
<td>BSI</td>
<td>Assuming 12-mo trial, 20% effect size, 90% power, 2-sided significance at 0.05, 4 ideally spaced scans; 138 AD per arm</td>
</tr>
<tr>
<td>Schuff et al., 2009</td>
<td>96 AD;226 MCI</td>
<td>Referral sample ADNI</td>
<td>Hippocampal volume (SNT), model includes 3 scans, Markov chain, APOE</td>
<td>Assuming 12-mo trial, 25% effect size, 90% power; 186 AD per arm; 341 MCI per arm</td>
</tr>
<tr>
<td>Vemuri et al., 2010</td>
<td>71 AD;149 MCI</td>
<td>Referral sample ADNI</td>
<td>Ventricular-BSI</td>
<td>Assuming 12-mo trial, 25% effect size, 80% power, 2-sided 2 sample t test at 0.05; 100 AD per arm; 186 MCI per arm</td>
</tr>
<tr>
<td>Wolz et al., 2010</td>
<td>126 AD;279 MCI</td>
<td>Referral sample ADNI</td>
<td>Simultaneous 4-D graph segmentation</td>
<td>Assuming 12-mo trial, 25% effect size, 80% power, 2-sided 2 sample t test at 0.05; 67 AD per arm; 206 MCI per arm</td>
</tr>
</tbody>
</table>

Key: AD, Alzheimer’s disease; ADNI, Alzheimer’s Disease Neuroimaging Initiative; APOE, apolipoprotein E; BSI, boundary shift integral; MCI, mild cognitive impairment; ROI, region of interest; TBM, tensor-based morphometry.
6. Measuring longitudinal disease progression with serial MRI scans

The idea of using change-over-time measures of brain volume on serial MRI was introduced by Freeborough and Fox (1997). This approach has appeal as a means of measuring disease progression that is independent of clinical assessment. It has found utility in assessments of individual subjects; in longitudinal observational studies; and as an outcome measurement in therapeutic trials. The potential of change-over-time measures as outcomes in therapeutic trials is particularly appealing because longitudinal MRI measures have considerably better precision and therefore can be powered with much smaller sample sizes than traditional clinical assessment tools. A number of different methodological approaches have been employed ranging from simple manual tracing to sophisticated TBM methods. Several investigators have shown that the lower variance in the serial MRI measurements compared with clinical measures of cognition and function could potentially permit performing clinical trials with smaller sample sizes than would be possible using traditional clinical instruments (Fox et al., 2000; Hua et al., 2008; Jack et al., 2003; Schott et al., 2006; Vemuri et al., 2010). Table 4 illustrates examples of sample sizes needed to power AD or MCI trials. Results vary depending on measurement method, assumptions about the trial design, and statistical methods.

References


Cognitive
Changes in cognition
Marilyn S. Albert*
Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Abstract
The clinical hallmark of Alzheimer’s disease (AD) is a gradual decline in cognitive function. For the majority of patients the initial symptom is an impairment in episodic memory, i.e., the ability to learn and retain new information. This is followed by impairments in other cognitive domains (e.g., executive function, language, spatial ability). This impairment in episodic memory is evident among individuals with mild cognitive impairment (MCI) and can be used to predict likelihood of progression to dementia, particularly in association with AD biomarkers. Additionally, cognitively normal individuals who are likely to progress to mild impairment tend to perform more poorly on tests of episodic memory than do those who remain stable. This cognitive presentation is consistent with the pathology of AD, showing neuronal loss in medial temporal lobe structures essential for normal memory. Similarly, there are correlations between magnetic resonance imaging (MRI) measures of medial temporal lobe structures and memory performance among individuals with mild cognitive impairment. There are recent reports that amyloid accumulation may also be associated with memory performance in cognitively normal individuals.

© 2011 Elsevier Inc. All rights reserved.

Keywords: Alzheimer’s disease; Dementia; Cognition; Cognitive testing; Cognitive function; Memory; Biomarkers

1. Introduction
The clinical hallmark of Alzheimer’s disease (AD) is a gradual decline in cognitive function. This decline occurs over years and ultimately leads to overt dementia, in which multiple cognitive domains are sufficiently impaired such that the individual is no longer capable of functioning independently. The pattern of decline in cognition is relatively consistent in the majority of patients. This reflects the fact that AD is a neurodegenerative process that follows a relatively predictable pattern, with specific brain regions affected early (e.g., the medial temporal lobe), while others are spared until late in the course of disease (e.g., sensorimotor areas). The role of biomarkers for AD ultimately will be to measure the neurobiological changes associated with this process in a sufficiently reliable manner that they can be used to detect, track, and predict the disease course over time.

2. Patients with AD dementia
The initial neuropsychological studies pertaining to patients with Alzheimer dementia focused on identifying the cognitive domains that are impaired among mildly demented patients, based on the assumption that this would provide information about the cognitive phenotype of AD. These studies led to a general consensus that the majority of patients have a deficit in the learning and retention of new information. This difficulty, known as an episodic memory impairment, is evident in day-to-day activities where retention over a delay is needed (such as remembering conversations and appointments) and on tasks that require an individual to learn something new (e.g., a shopping list or a news event) and then retain it over a delay. In the laboratory, episodic memory is assessed by a range of tasks including the learning and retention of stories, word lists, and paired-associates. Most of the initial studies compared Alzheimer dementia patients with healthy older controls,
and demonstrated a striking episodic memory impairment in the patients (Butters et al., 1988; Flicker et al., 1984; Hart et al., 1988; Moss et al., 1986; Petersen et al., 1994; Sahakian et al., 1990; Storandt and Hill, 1989; Welsh et al., 1991, 1992; Wilson et al., 1983).

Comparisons between Alzheimer dementia patients and patients with other causes of dementia (e.g., frontotemporal dementia, Huntington’s disease, Lewy body dementia), subsequently showed that the ability to retain information over a delay is more impaired in patients with AD than in other dementing disorders (Connor et al., 1998; Hamilton et al., 2004; Hodges et al., 1990; Lange et al., 1995; Milberg and Albert, 1989; Moss and Albert, 1988; Raskovsky et al., 2002; Tröster et al., 1993).

In addition to memory problems, mildly demented patients with Alzheimer dementia were shown to be substantially impaired with regard to a set of abilities collectively known as “executive functions.” This includes impairments on tasks that involve coordinating 2 concurrent tasks (Baddeley et al., 1986), as well as tasks requiring shifting between stimulus dimensions (Baudic et al., 2006; Filoteo et al., 1992; Parasaruraman et al., 1992; Sahakian et al., 1990; Sebastian et al., 2008). Mild-to-moderately demented Alzheimer patients also demonstrate executive function deficits (Becker, 1988; Bondi et al., 2002; Lafleche et al., 1990; Morris and Baddeley, 1988; Nestor et al., 1991; Stokholm et al., 2008). The evidence suggests, based on studies that compared very mildly demented Alzheimer patients with controls on tasks assessing a range of cognitive domains, that executive function deficits precede deficits in language and spatial function in the majority of individuals (e.g., Amieva et al., 2004; Grady et al., 1988; Lafleche and Albert, 1995).

Mild-to-moderately impaired Alzheimer dementia patients also have impairments in language function. Some investigators have argued that these deficits are the result of a broader impairment in semantic memory, defined as “that system which processes, stores and retrieves information about the meaning of words, concepts and facts” (Warrington, 1975). Semantic memory abnormalities in patients with Alzheimer dementia have been documented using a range of tasks that include category fluency (Chan et al., 1993; Hodges et al., 1992; Martin and Fedio, 1983; Tröster et al., 1989), category membership (Grossman et al., 1998), confrontation naming (Grossman et al., 1998; Hodges et al., 1992; Martin and Fedio, 1983), and similarity judgments (Chan et al., 1993, 1995, 1997). In addition, several studies of word priming (Glosser et al., 1998; Milberg et al., 1999; Salmon et al., 1988) have reported significant deficits in Alzheimer dementia patients, although other studies have failed to find this effect (Nebes and Brady, 1988). For a recent review of this area, see Altmann and McClung (2008).

Visuospatial function is impaired in the course of Alzheimer dementia. On simple copying tasks, such as drawing a clock or a triangle, mildly demented Alzheimer patients do not differ from controls (Karrasch et al., 2005; Rouleau et al., 1992). However, visuospatial impairments are common among mild-to-moderately impaired patients (Kurylo et al., 1994; Rouleau et al., 1996).

It should be noted that, although the findings above are characteristic of the majority of patients, at least 2 other clinical presentations have been described among patients with Alzheimer dementia: (1) individuals with gradually progressive impairments in spatial ability (sometimes referred to as having posterior cortical syndrome); and (2) individuals with gradually progressive language deficits. The former set of patients typically has AD pathology on autopsy, which is particularly striking in visual association pathways (e.g., Hof et al., 1997; Renner and Burns, 2005). The latter group of patients is more challenging to diagnose correctly, because their pattern of deficits overlaps that seen in a form of frontotemporal lobar degeneration (FTLD) known as primary progressive aphasia (PPA) (Mesulam et al., 2008).

3. Individuals with prodromal Alzheimer’s disease dementia

It is now widely accepted that there is a transitional phase between normal function and Alzheimer dementia, during which cognitive impairment is progressing. The term most commonly used to characterize individuals in this prodromal phase of disease is mild cognitive impairment (MCI) (Petersen, 2004; Petersen et al., 1999).

The feasibility of studying this transitional phase is based on the fact that the hallmark of AD is a progressive decline in cognition. To study this phenomenon, many research groups recruited nondemented individuals with mild cognitive impairments and followed them over time. The general design of the studies was to evaluate a range of cognitive functions in the participants when they were first evaluated, and then to follow them over time to determine which cognitive changes were the best predictors of progressing from mild impairment to frank dementia.

Among these studies there is considerable consensus that tests of episodic memory are significantly different among nondemented individuals with mild memory deficits who subsequently receive a diagnosis of Alzheimer dementia on follow-up, as compared with those who also have memory problems but do not progress to dementia within a few years (Albert et al., 2001; Blackwell et al., 2004; Bondi et al., 1994; Chen et al., 2000; Dierckx et al., 2009; Howieson et al., 2003; Jacobs et al., 1995; Kluger et al., 1999; Newmann et al., 1994; Petersen et al., 1994; Rabin et al., 2009; Rubin et al., 1998; Sarazin et al., 2007; Small et al., 1995; Tabert et al., 2006; Tierney et al., 1996; Tuokko et al., 1991).

A number of studies have reported that executive function abnormalities are also evident in the prodromal stage of Alzheimer dementia (Albert et al., 2001, 2007; Chen et al., 2008).
Accumulating evidence indicates that some individuals who are cognitively normal show the hallmark pathological features of AD in their brain (i.e., neuritic plaques and neurofibrillary tangles). The percentage of individuals with evidence of substantial AD pathology appears to vary with age (Bennett et al., 2006; Hulette et al., 1998; Knopman et al., 2003; Price 2009; Price and Morris, 1999; Schmitt et al., 2000). Approximately 1-third of the oldest individuals have these pathological changes. Despite abundant pathology, most investigators have, however, not reported that these individuals have significant neuronal loss. These findings suggest that there may be a “window of opportunity” to prevent Alzheimer dementia, if treatments can be initiated when individuals are cognitively normal despite evidence of some AD pathology in their brain.

Recent studies have therefore sought to determine if cognitive testing can be used to predict which cognitively normal individuals will subsequently develop cognitive decline and dementia. Though the number of studies is small, the majority have found that measures of episodic memory are significant predictors of future cognitive decline (Blacker et al., 2007; De Jager et al., 2005; Johnson et al., 2009; Kawas et al., 2003) or that tests of episodic memory are correlated with degree of pathology in normal individuals (Bennett et al., 2006; Schmitt et al., 2000). One study has reported that a composite measure of visuospatial ability, most of which were assessed by speed of performance, are significant predictors of which normal individuals will subsequently develop cognitive decline (Johnson et al., 2009). These findings suggest that the presence of AD pathology may be having a subtle influence on cognition, even though it is in the normal range.

6. Relevance to biomarkers

Studies using amyloid imaging have reported findings consistent with this hypothesis. For example, it has been reported that episodic memory performance is correlated with neocortical Abeta amyloid accumulation, as measured by PiB, in normal controls (Pike et al., 2007) and with decline in episodic memory performance (as well as working memory and visuospatial ability) (Storandt et al., 2009). Reports concerning the correlations between MRI measures of brain volume and cognitive performance in normal individuals suggest that the relationship may be complex. One study reported that episodic memory performance among controls with elevated PiB binding is associated with smaller hippocampal volume (Storandt et al., 2009), while another did not (Morimino et al., 2009). Subsequent analyses suggested that the sequential pattern of these changes may alter the associations that are seen at any 1 point in time (Morimino et al., 2009). Additional studies in this area should clarify these issues.

Disclosure statement

The author discloses no conflicts of interest.
References


Use of biomarkers to expedite clinical trials

Clinical trial methodologies for disease-modifying therapeutic approaches

Paul S. Aisen*

Department of Neurosciences, University of California San Diego, La Jolla, CA, USA

Abstract

In recent years, advances in Alzheimer’s disease (AD) biomarker research have provided powerful tools to improve trial design. In particular, biomarkers provide powerful methods for the selection of individuals with Alzheimer’s disease prior to the onset of dementia. Data suggest that neuroimaging biomarkers will be useful as endpoints for trials in very early, even asymptomatic disease, though further work is necessary to establish validity for regulatory purposes.

© 2011 Elsevier Inc. All rights reserved.

Keywords: Alzheimer’s disease; Biomarkers; Clinical trials

1. Overview

Alzheimer’s disease (AD) is not directly observable. The pathological changes, amyloid plaques, neurofibrillary tangles, and loss of neurons and synapses, can be observed histopathologically postmortem but not directly during life. Physical and neurological examinations are usually of limited or no use. The pathology causes progression from an asymptomatic state, through syndromes of mild cognitive impairment to progressive dementia and death. Symptoms cannot be adequately observed during limited clinical encounters; they are generally deduced from interviews with family members.

Biomarkers, objectively measured indicators of the disease, therefore are useful for diagnosis, longitudinal assessment, and evaluation of therapeutic response. They are essential to clinical trial design, particularly for disease-modifying interventions. Biomarkers can provide indirect indications of pathology, brain function, and symptomatology, aiding diagnosis and evaluation and providing quantitative assessment of drug effects.

2. Biomarkers for pharmacodynamics, mechanism, proof of concept, and dose selection

At early stages of clinical drug development, there are specific examples of utility of biochemical biomarkers in establishing target engagement and potentially in dose selection. The strongest examples involve the measurement of amyloid-beta (Aβ) peptides to indicate amyloid binding activity or reduction in peptide generation. For example, the activity of solenazumab, a monoclonal anti-amyloid antibody thought to reduce brain amyloid by binding and sequestering amyloid in the peripheral circulation, can be assessed by measurement of Aβ peptides in plasma. This method guided initial clinical studies in this program (Siemers et al., 2010). Measurement of cerebrospinal fluid levels of amyloid peptides can indicate the extent of peptide generation and clearance in the brain, particularly in radiolabeled amino acid infusion studies (Bateman et al., 2006). This technique supported dose-selection of the gamma secretase inhibitor semagacestat (Bateman et al., 2009).

3. Biomarkers for subject selection

It is generally estimated that 10%–20% of participants in AD trials do not have AD. Without a validated antemortem
diagnostic test, and particularly when site expertise may vary substantially, identification of subjects according to standard clinical and psychometric criteria is imperfect, leading to dilution of observable treatment effects on the disease. While not yet done for pivotal trials, addition of a biomarker assessment would be expected to significantly reduce the diagnostic inaccuracy at enrollment. For example, requiring an amyloid signal by amyloid positron emission tomography (PET) scanning or low cerebrospinal fluid (CSF) Aβ_{42} would reduce the number of individuals with cognitive symptoms caused by conditions other than AD enrolling in trials (Aisen et al., 2010).

As drug development programs move into the predementia population, this issue becomes much more important. Mild cognitive impairment is a heterogeneous clinical syndrome, with 30%–40% of individuals amyloid-negative and not destined for Alzheimer dementia. Predictors of progression include apolipoprotein E (APOE) genotype, cognitive and clinical scores, imaging measures, and cerebrospinal fluid markers; these can be used to enrich a population of individuals with mild cognitive impairment (MCI) for rapid progression (Aisen et al., 2010; Petersen et al., 2010). Particularly for an antiamyloid therapeutic programs, it seems appropriate to select individuals for trials on the basis of amyloid PET imaging and/or low cerebrospinal fluid Aβ_{42}, presumably, amyloid biomarkers not only enrich for progression, but also for potential response to antiamyloid intervention.

Apolipoprotein E genotyping can also be used for selecting predementia subjects more likely to progress to Alzheimer dementia. An advantage of this method is its low cost. However, selecting subjects on the basis of genotype excludes the large portion (30%–50%) of individuals with AD who do not carry the ε4 allele. Further, it may raise regulatory difficulties in late stage development, as it would be necessary to establish efficacy in ε4 carriers as well as lack of efficacy in noncarriers.

4. Biomarkers as covariates

Within any stage of Alzheimer dementia, biomarkers can be useful in defining the level of impairment, which in turn predicts subsequent decline, thus reducing unexplained variance in the modeling of trajectories of outcome measures. For example, baseline hippocampal volume can contribute to characterization of disease severity in individuals with mild cognitive impairment; including this as a covariate can increase study power with reduction of sample sizes by 5%–15% (Aisen et al., 2010).

5. Biomarkers to support proposed mechanism of action

Regulatory agencies currently require that pivotal trials demonstrate drug efficacy on the primary disease symptoms, and establish the clinical relevance of the effect; this has been accomplished using coprimary outcomes, specifically a cognitive performance test such as the AD Assessment Scale—Cognitive (ADAS—Cog) plus a clinician’s global impression of change. But regulators may consider treatment effects on biomarkers as indicators of impact on the underlying neurobiology (i.e., a disease-modifying effect) rather than just a symptomatic effect. For example, regulators have been willing to assess treatment effect on volumetric magnetic resonance imaging (MRI) measures as evidence that symptomatic benefit reflects an effect on the neurobiology of AD. Because no putative disease-modifier has yet had a successful pivotal trial, there are no examples of this use of biomarker data. In a small Phase II trial, however, it was possible to show an impact of antiamyloid immunotherapy on brain amyloid-load, though the small size did not allow association of this effect with cognitive performance (Rinne et al., 2010).

6. Surrogate outcome measures

In Alzheimer dementia and mild cognitive impairment, pivotal trials can rely on standard cognitive and clinical assessments to demonstrate efficacy and clinical relevance. But AD neurobiology, including the accumulation of amyloid in brain, begins many years before symptoms. Consensus holds that antiamyloid therapy and other disease-modifying interventions may have the greatest clinical impact if initiated at an early stage. Obviously cognitive and clinical assessments are not useful in this early asymptomatic phase of disease. Evaluation of drugs at this stage will require the use of biomarkers as surrogate outcome measures (Aisen, 2009). Candidate biomarkers for this purpose include volumetric magnetic resonance imaging and fluorodeoxyglucose positron emission tomography (FDG-PET) measures, as these seem to be dynamic indicators of disease progression even in the asymptomatic stage (Aisen et al., 2010; Jack et al., 2010).

However, regulatory agencies require that treatment effects on such biomarkers be reasonably likely to predict later clinical effects. To establish this validity, it is important that biomarker measures be included at all stages of clinical development. An association between a treatment effect on biomarkers and on cognitive and clinical assessments to demonstrate efficacy and clinical relevance.

Disclosure statement

The author discloses no conflicts of interest.

Acknowledgements

This work was supported by grants (U01-AG10483 and U01-AG024904) from the National Institute on Aging of the National Institutes of Health.
References


