The Use of MRI and PET for Clinical Diagnosis of Dementia and Investigation of Cognitive Impairment: A Consensus Report

Prepared by the Neuroimaging Work Group of the Alzheimer's Association

Marilyn Albert, PhD; Charles DeCarli, MD; Steven DeKosky, MD; Mony de Leon, EdD; Norman L. Foster, MD; Nick Fox, MD; Richard Frank, MD, PhD; Richard Frackowiak, PhD; Clifford Jack, MD; William J. Jagust, MD; David Knopman, MD; John C. Morris, MD; Ronald C. Petersen, MD, PhD; Philip Scheltens, MD, PhD; Gary Small, MD; Hilkka Soininen, MD, PhD; Leon Thal, MD; Lars-Olof Wahlund, PhD; William Thies, PhD; Michael W. Weiner, MD (Chair); and Zaven Khachaturian, Ph.D. (Secretariat).

1 The Neuroimaging Work Group of the Alzheimer's Association was convened to address the clinical application of brain imaging for the detection and diagnosis of cognitive impairment leading to dementia. The intent of this document is to (1) review current evidence in support of brain imaging in the detection and diagnosis of dementia, (2) suggest guidelines for the use of imaging in the clinical assessment of dementia, and (3) stimulate further systematic multisite research to validate the use of these methods in the early diagnosis and treatment of AD.
I. Introduction

Brain imaging in dementing illness has undergone revolutionary changes in the past 25 years with the wide availability of an unprecedented array of new techniques. The Neuroimaging Work Group Report presents this overview of the current state of the field and future directions.

This introduction, from a clinician’s perspective, provides a context for assessing current neuroimaging technologies as well as expected advances in (1) diagnosis of incipient cases, (2) prediction of incident cases and (3) outcome measures in therapeutic trials. Presently, neuroimaging cannot tell us whether or not a person has a cognitive disorder that is a clinical question. However, neuroimaging with computed tomography (CT) or magnetic resonance imaging (MRI) plays a key role in ruling out structural lesions of the brain in individuals with dementia. Either modality is excellent at detecting brain tumors, abscesses, strokes, and hematomas. Although mass lesions are uncommon, the discovery of such a finding dictates the need for an imaging study in the initial evaluation of every dementia patient.

Once the presence of dementia has been established, the role of imaging in the diagnosis of dementia subtypes is very much a function of the clinical diagnosis. The accuracy of the clinical diagnosis of Alzheimer’s disease (AD) is quite good. Pathological AD has a prevalence of about 70% (range 50% to above 80% depending upon whether the AD occurs in isolation or with other entities) among all dementias (see evidence Table 1 in reference 1); thus, even clinicians with limited neurological expertise should have a diagnostic accuracy, for AD at least, at about that level. A review of 13 published studies gave average values for sensitivity and specificity of the clinical diagnosis of AD of 81% and 70%, respectively (3). The overall accuracy of the clinical diagnosis of AD versus not-AD compared with the neuropathological standard based on those values for prevalence, sensitivity, and specificity, is 78%. For imaging to make a useful contribution to the diagnosis of AD, the sensitivity and specificity of imaging compared with neuropathological diagnoses must substantially exceed the clinical standard. Very few studies have addressed the accuracy of imaging studies as compared with pathological diagnoses. A clinical-pathological study of patients who had undergone Positron Emission Tomography (PET) studies showed the sensitivity of PET to be 94% with a specificity of 73% for the diagnosis of AD (3), which would yield an overall accuracy of 89%. However, the highly selective nature of the cohort makes generalization from this data risky. Comparable estimates for MR diagnoses of AD versus neuropathological diagnoses are not available, but it is clear that MR hippocampal atrophy has only modest specificity for AD, as hippocampal atrophy is seen in hippocampal sclerosis and frontotemporal dementia (3). Since clinical diagnosis achieves relatively high levels of accuracy, PET and MR, as currently performed, offer only relatively modest incremental benefits for the diagnosis of AD.

On the other hand, neuroimaging contributes to the diagnostic certainty of the frontotemporal dementias and Creutzfeldt-Jakob disease (4–7), and the diagnosis of vascular dementia requires imaging confirmation as well (2, 8). When these diagnoses are being considered in a particular individual, a clinician should have the ready opportunity to order MRI, PET, or Single Photon Emission Computerized Tomography (SPECT). It may never be possible to quantify the added value of certainty in dementia diagnoses. Yet ignoring existing resources that permit confident diagnoses, given their importance for physicians, patients, and families, is shortsighted.

In contrast to their accuracy in diagnosing prevalent dementia, current clinical techniques are poor at predicting which nondemented individuals will develop AD or other dementias in the future. Neuroimaging holds the promise of making an important and unique contribution to identifying individuals at higher risk for future dementia. Work demonstrating the potential of MRI has already begun to appear. Several groups have shown that hippocampal or entorhinal atrophy is associated with an increased likelihood of subsequent dementia due to AD (9–14). PET imaging changes may also have predictive value (15). While the practical value of such observations is limited by our lack of preventive therapies for AD, neuroimaging is poised to play a large role in future efforts at preventing dementia, as more effective preventive strategies emerge.

Perhaps the greatest value of neuroimaging will come in therapeutic trials. Current clinical trial methodology depends upon clinical assessments of cognition and behavior, which, while directly reflective of real-life phenomena, are inherently variable from day to day and examiner to examiner. Because of its precision, volumetric MRI could become the principal outcome measure for clinical trials in AD in the near future. Results of a recent study of MRI volumetric versus cognitive testing demonstrate this point clearly (16). Clinical trials in vascular dementia and frontotemporal dementia will also come to rely on MRI.

The technologies of MRI and PET have advanced rapidly. Development of novel imaging sequences or “contrast” agents that are sensitive to imaging of amyloid-beta peptide or tau pathology could change the landscape for diagnosis of prevalent cases by imaging. Specific molecular markers for MRI or PET imaging would be invaluable for predicting incident cases and in therapeutic trials. Ongoing collaboration between clinicians and neuroimaging practitioners is the key to moving closer to our ultimate goal in dementia research: to reduce the burden of dementia on individuals and on society.

---

2 Introduction prepared by David Knopman
II. The Role of MRI and CT in the Clinical Assessment of Cognitive Impairment and Dementia

Abstract and recommendations: Current American Academy of Neurology (AAN) guidelines for dementia diagnosis\(^1\) recommend imaging to identify structural brain diseases that can cause cognitive impairment. Accruing scientific evidence supports the use of imaging for the detection of early AD in individuals who have mild cognitive impairment (MCI)\(^2\), especially of the amnestic type (which may represent a prodromal form of AD). However, more research is needed to establish the clinical value of imaging in MCI. While much of this scientific evidence comes from quantitative analysis of structural brain MRI, qualitative estimates of medial temporal atrophy are highly correlated with quantitative measures and are much more suitable for clinical diagnosis.

The clinical use of structural brain imaging confers substantial assistance in the early diagnosis of AD. In this regard, the following recommendations are offered:

- Follow the American Academy of Neurology guidelines\(^3\), obtain brain imaging as part of dementia evaluation when AD is suspected and expand the AAN guideline to include individuals who have amnestic MCI.
- The impact of extensive white matter disease or a single lacunar infarct outside the thalamus in the presence of MCI remains unclear. Further research in imaging of a broad spectrum of individuals with MCI is needed to determine the exact utility of predicting a future “diagnosis of AD” and the particular type of MCI that is likely to be “prodromal AD.”
- Coronal brain imaging, preferably perpendicular to the long axis of the hippocampus, should be included in routine MRI protocols.
- Standardize imaging parameters whenever possible; consistent use of a standard imaging protocol and method of interpretation will translate into clinically meaningful results. Individual sites may need to tailor exact sequences to the type of brain imaging machine available.
- The use of a widely adopted (standardized and well-validated) protocol is essential for the clinical interpretation of medial temporal atrophy on MRI. However, presently there is a paucity of such qualitative instruments. Two scales are in use\(^4\), both of which have demonstrated (1) inter- and intra-rater reliability, (2) correlations with quantitative volumetric and neuropsychology and (3) prospective predictive value. The Work Group recommends the development and/or validation of uniform methods to measure, describe, or interpret the clinical significance of structural changes observed in brain images.

- Prospective studies are needed to evaluate the utility of MRI for differential diagnosis and early detection. There is a potential role for MRI in identifying AD by proven associations with AD pathology; future research should focus on the use of qualitative interpretation of various types of anatomical brain images from disparate populations to assess more fully the true utility of this method for the early diagnosis of dementia.
- Use of CT imaging for early diagnosis of AD should be the focus of future research for those individuals who cannot receive MRI\(^5\).

A. Rationale for Structural Imaging in the Early Detection of Dementia

The pathology of AD has a definite topographic distribution\(^6\) in regions that can be well characterized by available neuroimaging methods\(^7,8\). Hippocampal atrophy is an early marker of AD\(^9,10,11,12,13\) that correlates with impairments in memory function\(^14\). In addition, AD leads to progressive loss of brain volume throughout the cerebral cortex and other areas that is significantly greater in AD patients than age-matched controls\(^15,16,17,18,19,20,21,22\) that correlates with the rate of cognitive deterioration\(^23\).

Neuropathological studies also show that the pathological features of AD may be present for years before clinical symptoms are evident\(^24,25\) and are often present in individuals with memory impairment who are not demented, suggesting that most individuals with MCI have early AD\(^26\). Additionally, careful pathological studies find a close association between hippocampal size as imaged by MRI and the extent of AD pathology\(^27\). Given that structural imaging is a relatively easy and noninvasive method to evaluate early AD and is closely associated with the pathological features of the disease, a number of studies have evaluated the prognostic significance of finding hippocampal atrophy in high-risk populations.

B. The Prognostic Significance of Hippocampal Atrophy in MCI

Accumulating evidence from quantitative MRI studies suggests that hippocampal atrophy is present before dementia onset\(^28,29,30,31\) and progresses with conversion to clinically apparent AD\(^32\). In a large prospective study of MCI patients, Jack et al.\(^33\), found a 4-fold increase in the percentage of individuals converting to dementia within five years when initial hippocampal size was two standard deviations below age- and gender-defined norms. Similar findings were noted in a second study, although memory scores were also significant predictors\(^34\).

---

\(^{3}\) MRI & CT Subcommittee Members: Charles DeCarli, (Chair), Philip Scheltens, Norman L. Foster, Hilkka Soinenen, Lars-Olof Wahlund, John C. Morris, William J. Jagust, Gary W. Small, Nick Fox, Ron Petersen, Michael W. Weiner (ex officio) and Zaven Khachaturian (Secretary).

\(^{4}\) Presently the construct of MCI is not a diagnosis; it has no code in either the ICD or DSM documents.
Another study using qualitative estimates of hippocampal size also found similar results. These findings support the utility of structural brain imaging in MCI to predict conversion to AD within five years. Importantly, future work from population-based studies will be helpful in clarifying the utility of structural imaging in the diagnosis of early AD for populations where clinical definitions of MCI are less predictive. The general use of structural imaging, however, requires a simple and convenient method by which clinicians can have reliable and accurate estimates of hippocampal size. Recent developments in qualitative measurements appear to meet this need.

C. Qualitative versus Quantitative Measures of Hippocampal Atrophy

Image quantification is an arduous task, currently requiring considerable processing time and operator attention. Thus, it is not applicable now to the clinical evaluation of the average patient with a memory disorder. A variety of qualitative scales exist, several of which have high intra-observer reliability and correlate with neuropsychological and volumetric measures.

D. CT versus MRI

MRI is generally regarded as a superior tool for brain imaging, as compared with CT, due to the absence of ionizing radiation, increased imaging flexibility, and better tissue contrast. Unfortunately, expense, patient claustrophobia, or the presence of metal implants or medical devices common in older individuals can limit the use of MRI. A variety of CT protocols have been developed to image medial temporal structures, several of which have high intra-observer reliability and correlate with neuropsychological and volumetric measures.

III. The Role of Structural MRI in Clinical Trials in AD and MCI

Abstract and recommendations: Structural MRI has been extensively used to characterize the changes in normal aging, mild cognitive impairment, AD, and other dementias. MRI measures of brain volume, especially medial temporal lobe structures, are expected to be useful surrogates for measuring treatments that slow progression of neurodegeneration in AD. This is because purely symptomatic therapy should not affect rates of atrophy and because of the high statistical power of these measures. MR imaging should be used in clinical AD trials for characterizing structural brain changes in subjects and for ruling out other causes of dementia. It could also be used to monitor treatment effects over time, depending on the nature of the trial.

Multisite studies of MR imaging outcomes should be done at 1.5T, with a T2-weighted, or fluid attenuated inversion recovery (FLAIR) protocol, to assess signal changes due to vascular or other causes, and a 3D T1-sequence to assess atrophy. To reduce dropout and improve compliance, acquisition time should be minimized—total scanner time should be less than 20 minutes, with no single sequence more than 10 minutes. Quantitative measures are necessary and require high standards of quality control during image acquisition, including assessments of geometric fidelity, contrast, and lack of homogeneity.

Analyses should be standardized and include a measure of global loss and a measure of medial temporal lobe volume. A multisite, longitudinal structural MRI “observational, naturalistic” study of brain atrophy rates in AD, MCI, and control subjects would facilitate the establishment of standards for acquisition and image processing and provide information useful for the design of treatment studies. Recently the NIH announced an RFP for a neuroimaging initiative to perform such a study. Studies at higher fields and using magnetic resonance spectroscopy (MRS), perfusion, diffusion, amyloid imaging, and other advanced techniques should be encouraged for future applications.

A. Rationale for Use of MRI and CT in Clinical Trials

The main potential roles of MRI in trials are to (1) define the study population (provide exclusion and inclusion criteria and sample stratification) and (2) measure outcome (provide surrogate markers of progression using serial MRI). Each of these functions requires consideration of optimal acquisition and analysis for the particular subject group studied.

B. Defining Study Populations in MCI and AD

Due to the lack of specificity of clinical criteria and the heterogeneity of AD and MCI, MRI can assist in defining and homogenizing study populations by adding MRI exclusion and inclusion criteria to existing clinical criteria. Excluding patients with significant small vessel cerebrovascular disease as well as large vessel strokes will narrow, but also homogenize, the study population. Including patients with a significant degree of medial temporal atrophy (MTA) will increase the proportion of subjects that will progress from MCI to fulfill a diagnosis of AD. The challenge of MCI is to determine which patients will progress to fulfill criteria for AD, since some have fixed deficits, others will develop vascular dementia or frontotemporal dementia, and a few will turn out to be “worried well.”

MRI is a feasible addition to clinical trials because it is widely available and inexpensive relative to the total cost of a clinical trial. MRI is noninvasive, and even with repeated imaging no adverse effects are known, as long as care is taken to exclude...
subjects with pacemakers or certain ferrous-metallic implants, and appropriate ear protection is provided. Most MRI research studies require patients to have a mini–mental state examination (MMSE) score of >10/30 in order to comply with instructions and to remain still during the scan. More severely affected subjects, however, may still be scanned if short scan times (<10 minutes) are used. Although sedation (usually with a benzodiazepine) is used in clinical practice for anxious and claustrophobic patients, sedatives are generally inappropriate during a clinical trial.

C. Recommendations for MRI Inclusion and Exclusion Criteria

The Work Group recommends the use of MRI in all clinical drug trials that seek to establish AD as the pathological substrate for therapy. Imaging should be used to exclude (1) nondegenerative, nonvascular pathology such as tumor, subdural hematoma, hydrocephalus, etc., (2) cerebrovascular disease as the major pathological substrate for the cognitive problems (i.e., vascular dementia or MCI of the vascular type) and (3) other degenerative dementias presenting with focal or lobar atrophy. Either MRI or CT may be adequate for 1 above; MRI is superior in assessing 2 or 3.

Even if a given therapy is considered nonspecific for one sort of pathology and licensing is sought for a nonspecific indication such as “dementia,” the use of MRI should still be considered since (1) there is potential value in determining which subgroups show response to the therapy (assessments of vascular load or MTA may be useful covariates in outcome analyses) and (2) these data may make a contribution to research more generally.

1. Technical requirements and specifications for exclusion criteria to exclude/assess vascular and nondegenerative pathologies:

1.1. Acquisition. Scanner field strength should be at least 0.5T (preferably 1.5T) and, ideally, single field strength should be used for all patients in a particular trial to allow comparison between sites and scanners. The field of view should ensure whole brain coverage. T2-weighted or FLAIR techniques should be used, particularly to assess white matter disease, and acquisition time should be tailored to patient tolerance.

1.2. Analysis. We recommend central assessment of all scans, based on visual inspection, to exclude tumor, subdural hematomas, etc. Such inspection can be performed on either hard copy or digital imaging. If the vascular load is to be assessed, this should involve either (1) a visual grading scale, (see reference 68) or (2) an automated or semi-automated method for quantifying vascular load should be encouraged; such a method should have been fully validated prior to the study.

2. Technical requirements and specifications for inclusion criteria to increase probability of AD pathology. AD is associated with early and disproportionate median temporal lobe atrophy diffuse cerebral atrophy. The presence of atrophy of the medial temporal lobe increases the probability of AD in clinically diagnosed AD and MCI patients, but has less specificity in distinguishing AD from other dementias.

2.1 Acquisition. We recommend the use of whole brain coronal 3D T1-weighted imaging to assess lobar and medial temporal lobe atrophy.

2.2 Analysis. We recommend central assessment of MTA, (1) based on visual inspection of hard copy or digital imaging or (2) calculated using a region-of-interest based volumetric analysis.

There is now good evidence that voxel-based, automatic, computerized methods of scan analysis e.g., voxel based morphometry, can be effectively used to analyze differential and sequential changes in brain anatomy. These methods have the advantages of using all the available scan data to maximize the sensitivity for detection of change, of being objective in the sense that observer bias is removed, and of providing results that are reproducible on the same data set after image preprocessing. A formal comparison between observer-based ROI and voxel-based computerized methods in Alzheimer’s disease is very promising.

D. Progression Markers

1. Current practice. The advent of potential disease-modifying agents in AD has heightened the importance of developing imaging markers of progression. Such markers will not only increase our knowledge of disease and help provide prognostic information to patients but may also provide cost-effective ways of identifying therapies that slow AD, as opposed to providing only symptomatic benefit. Ideally, a surrogate marker of disease progression should relate directly to the extent of the underlying molecular pathology—synaptic loss, amyloid load, or abnormal tau deposition. Such measures are being sought but to date are not available in vivo. Cerebral atrophy due to neuronal loss is a downstream event, which is nevertheless central to pathological progression. MRI can measure rates of atrophy that can act as in vivo markers of structural neuronal degeneration. Atrophy progression can be assessed repeatedly and noninvasively, blind to treatment allocation and to time point within a trial. However, other factors that influence brain hydration may produce alterations in brain volume unrelated to neurodegeneration.

Therefore, these and other potential confounding factors must be considered in relation to specific interventions. Currently, there are several large multicenter clinical trials in MCI and AD that are using MRI measures of atrophy as outcome measures. The outcome measures of interest are measures of regional (medial temporal lobe) atrophy and/or whole brain atrophy rates. At time of writing, the results of these studies have not yet been reported.
2. Recommendations. Measures of rates of atrophy based on manual outlining of regions of interest (for example, the hippocampus or entorhinal cortex), and semiautomated whole brain atrophy rates from serial MRI are the first choices as outcome measures. Rates of atrophy of other structures should also be investigated as possible surrogate markers. Alternative manual and automated image analysis techniques are in development. These merit comparison with the forgoing outcome measures in future multicenter studies to determine the most powerful markers of disease progression. A central site for standardized analysis should be used; if multiple independent measures are being chosen, they may be performed at different central sites.

2.1 Medial temporal lobe atrophy

2.1a. Hippocampus (HC). The HC is the most extensively studied structure in AD; large numbers of cross-sectional and a small number of longitudinal studies(66) have shown increased rate of atrophy (4 to 6% annually) in patients with AD, relative to controls (1 to 2% annually)(66, 85). Manual tracing of the HC on digitized images is the best validated and recommended method. Alternatives that require further longitudinal evaluation include (1) visual assessment using rating scales(12, 36), (2) stereological measures(97) and (3) automated deformation-based methods either from standard template(88, 89) or from a baseline segmentation(90).

2.1b. Entorhinal cortex (EC). The EC should, on pathological grounds, be at least as sensitive as the HC as a measure of progression in AD. However, measurement reliability may be lower for the EC than for the hippocampus. It is currently unclear if there is any practical advantage in using EC measures over those from the HC; both measures appear to provide similar power for clinical trials in AD(33, 91). The role of EC compared with HC to determine atrophy rates in MCI and AD remains to be determined (92–94).

2.2. Global atrophy

2.2a. Registration-based methods allow semi-automated measurement of atrophy rates (brain boundary shift integral), which have shown annualized mean (SD) rates of atrophy in AD of 2.4 ±1.1% annually vs. controls 0.5 ±0.4%(95). These methods can incorporate correction for scanner voxel variability. Further experience is required to assess whether similar results are possible with the greater variability in MR acquisition in multisite trials. There are some published data on sample sizes(96) and this method is currently being used in multicenter trials in AD and MCI. In the future, automated registration algorithms might increase the comparability of findings between studies that are analyzed at different sites.

2.2b. Brain parenchymal fraction (BPF) has been used in multicenter multiple sclerosis (MS) trials. To date, there is little experience with this technique in dementia(97, 98).

2.2c. Ventricular CSF measurements are simple markers of global atrophy, which have been used in longitudinal studies to date(99–103).

2.2d. High-dimensional nonlinear registration methods, as novel techniques, have the potential to warp a template or a baseline image onto follow-up images, allowing “compression maps” to provide rates of atrophy in different regions (88, 98, 104). These methods hold considerable potential to reduce user input, especially with multiple scans per subject, however, more research is needed since there is currently limited validation in AD(83, 105).

2.2e. Sample size estimates depend on the degree of homogeneity of disease severity in subjects, method of data acquisition, method of image analysis, and possible treatment effects. Therefore, sample size requirements cannot be generalized but must be calculated for each trial.

E. Technical Requirements and Specifications

Field strength of 1.5T is adequate and embraces the combined attributes of wide use and availability. Although higher field strengths (3T and above) may improve contrast and resolution for a given acquisition time, the lack of widespread availability of such scanners currently makes them unsuitable for large Phase III studies. There may be advantages in their use in Phase II studies, although experience is limited at present. For atrophy measurements in AD and MCI, we recommend the use of a volumetric T1 weighted imaging (spoiled gradient-echo ‘SPGR’ and magnetization pre-prepared rapid acquisition gradient echo ‘MPRAGE’) or inversion recovery sequence, with voxel dimensions (ideally isotropic) of 1.5mm or less. Rigorous comparison of these two types of sequences is lacking.

Multisequence acquisition for segmentation purposes is an area of research that should be encouraged, as should comparisons between these different techniques in dementia studies.

F. Manufacturer and Other Consistency Considerations

While the use of a single MRI manufacturer in a multicenter study may simplify acquisition and quality control, and thereby improve consistency, this is an unrealistic strategy for large Phase III studies. The opportunity exists for MRI manufacturers to establish their commitment to providing consistent MRI acquisition, specifically for multicenter quantitative studies, and there is a need to improve the standards of quality control for quantitative as opposed to clinical work. This Work Group would encourage pursuit of these interests. We recommend that setting standards for consistent acquisition and good quality assurance should be the responsibility of and managed by an independent image-analysis coordinating center.

G. Role of Structural MRI as an Outcome Measure in Clinical Trials

According to accelerated approval regulations for serious or life-threatening illnesses, including AD, the United States Food and
A growing clinical literature indicates that patterns of FDG PET are offered: developed for the putative assessment of AD histopathology in the living human brain and therefore helpful in the differential diagnosis of dementia. Current evidence thus suggests that FDG PET may be considered as part of the evaluation of patients with dementia when symptoms are unusual, present diagnostic difficulties, or reflect diagnostic uncertainties between AD and frontotemporal dementia. FDG PET in direct comparison with clinical diagnosis, and in addition to a high-quality evaluation including MRI, has not been thoroughly evaluated and deserves further study.

- Considerable evidence also suggests that FDG PET may reveal metabolic abnormalities very early in the course of degenerative dementias, which portend the likelihood of progressive clinical decline. Therefore, PET may be of clinical utility in the evaluation of patients presenting with mild symptoms of memory loss and cognitive dysfunction by establishing a likely, though nonspecific, neurodegenerative basis for the symptoms. FDG PET has not been fully studied in this situation in comparison with other modalities, including clinical, cognitive, and MRI evaluations, and deserves further study.

- Clinical PET studies should be performed at rest with minimal ambient stimulation. Individuals specifically trained to interpret PET FDG images in patients with dementia should analyze images.

- FDG PET provides a promising marker of disease progression. It may have greater statistical power in the assessment of putative treatments than traditional outcome measures. Thus FDG PET may be useful as an ancillary outcome measure in clinical trials of putative AD treatments, and further research in this area is encouraged.

- Additional studies are recommended to evaluate novel radiotracer techniques for imaging amyloid and other histopathological features of AD in the living human brain and in relevant animal models.

- Research PET studies should use standardized protocols that specify acquisition, image analysis, and quality control procedures, which may be different from the techniques required for clinical studies.

A. Rationale for Use of PET

Patients with probable and definite AD have reductions in posterior cingulate, parietal, temporal, prefrontal, and whole-brain measurements of the cerebral metabolic rate for glucose (CMRgl), which are progressive and correlated with dementia severity. These findings predict both subsequent clinical decline and the histopathological diagnosis and are apparent prior to the onset of AD and in persons at risk for AD. The CMRgl reductions may reflect a reduction in density or activity of terminal neuronal fields or perisynaptic glial cells, a metabolic abnormality, or a combination of these factors, and do not appear to be solely attributable to the combined effects of atrophy or partial volume averaging\(^{(106)}\). Other PET tracers, notably oxygen-15\(^{(107)}\), also have been used in the study of AD, including tracers that have been recently developed for the putative assessment of AD histopathology in the living human brain.
brain. The PET subcommittee sought to answer the following questions:

- What methodological and technical considerations must be addressed to further develop PET for multicenter studies?
- Can PET increase accuracy in the differential diagnosis of dementia, particularly in its earliest clinical stages?
- Can PET aid in determining the prognosis of individuals at high risk for dementia?
- How can PET enhance the discovery of treatments and prevention therapies for memory and dementing disorders?

B. Technical Considerations for a Multisite PET Research Study

Data acquisition and image analysis methods must be considered carefully when designing experiments and proposing clinical applications. PET studies should control behavioral activity, task performance, ambient room conditions, whether eyes are open or closed, and monitor use of concurrent medication. Attempts have been made to improve the ability to distinguish groups of AD patients from normal controls by conducting PET studies during standardized behavioral tasks (for example, visual, visuospatial attention, and memory tasks). However, studies performed during mental rest currently have the best established value and are less likely to be confounded by effects of differential task performance, as well as motivation, attention, and effects of medications.

A longitudinal decline in whole brain metabolism has been found in studies of patients with probable Alzheimer’s disease (108, 109). This decline was not observed in a study of normal aging, and it is not yet clear whether this decline will be observed in patients with mild cognitive impairment. In order to generate quantitative images of CMRgl, an FDG input function should be calculated using repeated blood samples from the radial artery, use of “arterialized venous” samples, or dynamic PET scans (image-derived input carotid artery input function) (110). In order to maximize subject tolerance, minimize subject attrition, and investigate declines in normalized regional PET measurements in multicenter studies, the subcommittee recommends further evaluation of these various methods of the performance of PET studies in the absence of arterial or arterialized venous measurements. Approaches to data display include:

1. **Generation of statistical brain maps**, which appear to improve the ability to characterize abnormalities, compare measurements from different subjects, and compare findings from different studies; of course, these do not necessarily need an input function, but in that case do not quantify glucose metabolism.

2. **Use of region of interest sampling** (111), which requires a co-registered structural image for the guidance of sample placement.

3. **Statistical parametric mapping (SPM)** approaches, which, unlike ROI approaches, may not be able to study accurately small regions such as the hippocampus in atrophic brains.

One study (112) used hippocampal volumes of interest to study co-registered MRI and FDG PET images (and atrophy-corrected PET data) of normal, MCI, and AD groups, showing that the two modalities had equivalent classification value. In a longitudinal study of normal elderly that declined to MCI (115), only an entorhinal cortex volume of interest predicted which elderly would deteriorate to MCI. Further, at follow up, the MCI group demonstrated hippocampal metabolic reductions relative to the nondeclining normal group. This and other studies (106, 112, 113) indicate that the neocortical deficits observed in AD reflect true metabolic reductions and not just the result of atrophy.

Simple instructions appear to improve inter-rater reliability in the visual interpretation of PET images (114). Several algorithms have been used to generate statistical brain maps in the study of dementia, including (but not limited to) statistical parametric mapping (SPM) (115, 116), stereotactic surface projection (SSP, also known as Neurostat) (117, 118), and principal component analysis (PCA) (119). In clinical studies, the reliability of image interpretation needs to be further studied (120).

C. Increasing the Specificity of Dementia Diagnosis

PET consistently shows reduced glucose metabolism in AD in the posterior cingulate gyrus, the parietal and temporal association cortices, and in later stages of the disease spreading into the prefrontal cortex and other brain regions (121). The extent of hypometabolism correlates with severity of cognitive impairment (122) and often shows right/left hemispheric asymmetry (123, 124).

In a multicenter study of patients undergoing evaluation for dementia symptoms (114), visually interpreted FDG PET images predicted progressive dementia with a sensitivity of 93% and a specificity of 76%. FDG PET predicted the histopathological diagnosis of AD or any neurodegenerative disease with a sensitivity of 94% and specificities of 73% and 78%, respectively, comparing favorably with sensitivities of 83% to 85% and specificities of 50% to 55% with a Class I study of standard clinical assessment of early dementia without PET (125). However, this large multicenter study has several limitations, including a relatively short average follow-up period of 3 years, questions about the extent to which these patients referred from clinical research centers reflect the typical medical setting referral, the retrospective nature of the analysis, and the lack of an attempt to compare the accuracy of visual inspection with that of either a brain mapping algorithm or ROI approach.

Research is needed to further substantiate the role of PET in distinguishing among dementing disorders and to characterize the full range of metabolic findings. Multicenter trials of PET should study consecutive patients and consider how PET may contribute further certainty to the standard clinical diagnosis. It
will be important to define the determinants of typical patterns of hypometabolism in AD. Some evidence suggests that typical AD patterns are less often seen in older individuals or in those with superimposed extensive cerebrovascular disease\(^{(126, 127)}\). FDG PET is helpful in distinguishing frontotemporal dementia (FTD) from AD. One recent study\(^{(129)}\) found that the agreement between FDG PET and neuropathological diagnosis was between 75% and 90% among six raters and was better than that found between the clinical examination and neuropathological diagnosis (75% to 80%). Consensus criteria for FTD consider PET findings supportive evidence for clinical diagnosis of FTD\(^{(129)}\). Other studies are needed to confirm whether PET aids in recognition and predicting prognosis of mixed dementias. PET may be cost-effective in case of suspected dementia\(^{(129, 130)}\). Prospective multicenter trials that compare MRI and PET with postmortem confirmation are needed to establish the reliability, sensitivity, and specificity of distinguishing different causes of dementia and to further demonstrate whether such imaging adds benefit to the accuracy of clinical diagnoses rendered by both dementia specialists and by nonspecialists.

**D. Improving the Accurate Recognition of Progressive Dementia**

PET is a useful biomarker for many causes of dementia and helps to determine whether an individual with cognitive complaints will suffer further decline\(^{(15, 114)}\). Patients with dementia due to neurological disease almost uniformly have abnormal PET scans, while those with cognitive complaints from other sources often have normal scans\(^{(114, 131, 132)}\). PET scans are abnormal even when symptoms of AD are mild\(^{(112, 114, 121)}\). Thus PET may be useful in differentiating neurological disease from psychiatric and drug-induced causes of behavioral and cognitive dysfunction.

An abnormal PET scan may provide supportive evidence for the presence of a neurodegenerative dementing disease, rather than a reversible psychiatric illness or symptoms due to normal aging. PET should be considered as an option to assist in the recognition of dementia when coexisting illness or baseline abilities and cooperation are uncertain.

**E. Assessing the Prognosis of Individuals at Increased Risk for Dementia**

PET may provide additional information about risk in individuals who already are at increased risk because they have memory impairment or carry the apolipoprotein E4 allele (APOE-4). Several studies have demonstrated CMRgl reductions in patients who have various forms of MCI as well as the value of these reductions in predicting subsequent cognitive decline\(^{(112, 133)}\). Moreover, one study showed that FDG PET measures predict decline of normal individuals to MCI\(^{(15)}\). Further work is needed to determine the predictive value of PET in patients with memory impairments.

The risk of AD is increased in those with one copy of APOE-4, an AD susceptibility gene found in almost 25% of the population, and is even greater when two copies of APOE-4 are inherited\(^{(139)}\). PET has shown that carriers of the APOE-4 allele have metabolic reductions compared with noncarriers\(^{(135-138)}\). These metabolic changes follow a pattern typical for AD and show progressive regional metabolic decline indicating that PET might serve as a surrogate marker in clinical trials designed to prevent future cognitive decline\(^{(136, 138)}\). The available data do not support the routine use of PET for assessing asymptomatic individuals at genetic risk.

**F. Role of SPECT**

Single Photon Emission Computed Tomography (SPECT) is another molecular imaging technique that can provide information similar to that obtained with PET. Because of differences in the physics of single photon emissions as compared with positron emissions, SPECT has lower resolution than PET. In addition, the availability of single photon emitting radionuclides has generally limited the number of SPECT radiotracer ligands that have been developed. Nevertheless, SPECT imaging has received widespread clinical application because of its simplicity, use of long-lived radionuclides, and lack of need for a local cyclotron. For many years, SPECT was much more available than PET, although this trend may be changing as the clinical application of PET to cancer management and other conditions becomes more widespread. Although the theoretical limitation of SPECT resolution is significant, clinically useful SPECT images of cerebral blood flow still can be produced when multidetector SPECT cameras are employed.

Many studies have evaluated SPECT perfusion imaging in the diagnosis of dementia. These studies are often difficult to compare with one another and with PET studies, since methods for data analysis differ substantially and because SPECT utilizes measures of perfusion, while most PET studies involve measures of cerebral metabolism. The majority of SPECT studies find that the pattern of hypoperfusion in temporal and parietal cortex have reasonable sensitivity and specificity for AD. Autopsy studies have indicated sensitivity and specificity in the range of .7 to .8 for the diagnosis of AD in comparison with control subjects or other dementias\(^{(139, 140)}\). One study that employed the clinically useful approach of investigating the added value of a SPECT scan in addition to the clinical evaluation found that when a diagnosis of probable AD was made clinically, a positive SPECT scan increased the likelihood of autopsy confirmed AD from 84% to 92%\(^{(141)}\). SPECT has also detected abnormalities in those who are presymptomatic for AD, similar to those described with FDG PET. Baseline SPECT scans differ in individuals with questionable A\(_\beta\)\(^{(142)}\) or MCI\(^{(143)}\) who subsequently convert to dementia in comparison with normal controls. These SPECT perfusion differences occur in brain regions associated with early AD, such as the cingulate...
cortex. Like FDG PET, SPECT measures of cerebral blood flow provide independent complementary information to structural imaging studies\(^{(146)}\).

In general, therefore, the results of SPECT studies are similar to those obtained with PET. The few studies that have directly compared SPECT with PET in AD find the performance of PET superior to SPECT in both detecting abnormalities and in differentiating AD patients from controls, presumably because of the greater resolution and sensitivity of the technique\(^{(145, 146)}\). The inherent superiority of PET probably makes this technique preferable, although SPECT can provide clinically useful information that may be comparable.

G. Assisting in the Discovery of New Treatments and Preventions

PET is a promising surrogate marker in the assessment of a putative AD treatment’s ability to improve clinical outcome and modify disease progression. FDG PET may provide evidence of disease progression, and it provides complementary information to the best-established structural brain imaging measurements of disease progression (i.e., MRI measurements of hippocampal, entorhinal cortex, and whole brain volume). A longitudinal study comparing MRI and PET modalities is in order. PET may have greater statistical power in the assessment of putative treatments than traditional outcome measures\(^{(109)}\).

FDG PET is reasonably likely to provide information about the disease-modifying effects of putative treatments, particularly when used in conjunction with volumetric MRI and when the findings are supported by studies that include a randomized start or withdrawal design. FDG PET studies of eventually established disease-modifying treatments are needed to help validate this surrogate marker for the discovery of putative AD treatments and prevention therapies. For these reasons, the use of FDG PET is encouraged as an ancillary outcome measure in Phase III clinical trials of putative AD treatments. Reductions in regional CMRgl are progressive in patients with probable AD\(^{(108, 109)}\). They are correlated with dementia severity in both their magnitude and spatial extent\(^{(123)}\), and they are also correlated with the severity of the symptoms caused by damage in that brain region\(^{(147, 148)}\). These reductions could reflect a decrease in the activity or density of terminal neuronal fields or perisynaptic glial cells, an abnormality in glucose metabolism itself, the combined effects of atrophy and partial-volume averaging, or a combination of these factors; they do not appear to be solely attributable to the effects of atrophy\(^{(113)}\). In a one-year follow-up study, mildly to moderately affected patients with probable AD had an annual rate decline in regional and whole brain glucose metabolism between 4% and 11%, with no normalization for the variation in absolute measurements\(^{(109)}\). Based on this decline, PET was estimated to have almost 10 times the power of dementia rating scales for detecting a treatment effect in a randomized, placebo-controlled clinical trial of patients with probable AD, and about the same as that reported using MRI measurements of whole brain atrophy\(^{(106)}\).

Multisite trials using SPECT and PET have been effectively utilized in clinical trials in Parkinson’s disease\(^{(149, 150)}\), demonstrating the ability to perform multisite trials of this type. Although any surrogate end point can be misleading in clinical trials\(^{(151)}\), the accumulation of evidence argues for the utility of PET as a surrogate marker in clinical trials of age-related memory loss and dementia. Use of more than one marker (for example, PET CMRgl measurements, PET measurements of amyloid binding, and MRI measurements of hippocampal and whole brain volume) may provide converging evidence that is stronger than any individual measure. If validated as a predictor of clinical outcome, PET could reduce the number of patients needed to evaluate drug effectiveness.

H. Additional Research Opportunities

There are many other potential applications for PET that utilize the diverse radioligands developed to measure neurotransmitter activity and neuroreceptor distribution. Particularly exciting are recent developments in small molecule probes for the putative in vivo assessment of neurofibrillary tangles (NFTs) and neuritic plaques (NPs)\(^{(152, 153)}\). Several groups have been developing PET probes of nicotinic and muscarinic\(^{(154, 155)}\), but availability and previous experience with these approaches are limited. Recent work using autoradiography suggests that it may also be possible to assess drug effects on transgenic animals using micro-PET before extending observations to humans\(^{(156)}\); if abnormalities in regional CMRgl can be detected despite limitations in the spatial resolution of this imaging technique\(^{(157)}\). Additional studies are recommended to extend FDG and histopathological PET methods to the study of transgenic mice and other relevant animal models in order to refine screening of candidate treatments.

References


58. Chui H, Young L. Clinical criteria for various dementia subtypes: Alzheimer’s disease (AD), vascular dementia (VaD) dementia with Lewy bodies (DLB), and frontotemporal dementia (FTD). In: *Evidence-Based Dementia*. Oxford: Blackwell Science, LTD; 2002.


