Research criteria for the diagnosis of Alzheimer’s disease: revising the NINCDS–ADRDA criteria


The NINCDS–ADRDA and the DSM-IV-TR criteria for Alzheimer’s disease (AD) are the prevailing diagnostic standards in research; however, they have now fallen behind the unprecedented growth of scientific knowledge. Distinctive and reliable biomarkers of AD are now available through structural MRI, molecular neuroimaging with PET, and cerebrospinal fluid analyses. This progress provides the impetus for our proposal of revised diagnostic criteria for AD. Our framework was developed to capture both the earliest stages, before full-blown dementia, as well as the full spectrum of the illness. These new criteria are centred on a clinical core of early and significant episodic memory impairment. They stipulate that there must also be at least one or more abnormal biomarkers among structural neuroimaging with MRI, molecular neuroimaging with PET, and cerebrospinal fluid analysis of amyloid β or tau proteins. The timelines of these criteria is highlighted by the many drugs in development that are directed at changing pathogenesis, particularly at the production and clearance of amyloid β as well as at the hyperphosphorylation state of tau. Validation studies in existing and prospective cohorts are needed to advance these criteria and optimise their sensitivity, specificity, and accuracy.

Background

For research purposes, the diagnosis of Alzheimer’s disease (AD) is based on the criteria of the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV-TR) and the National Institute of Neurological Disorders and Stroke–Alzheimer Disease and Related Disorders (NINCDS–ADRDA) working group. These accepted criteria are fulfilled in a stepwise process where there is initial identification of a dementia syndrome and then the application of criteria based on the clinical features of the AD phenotype. The DSM-IV-TR criteria require the presence of both a memory disorder and impairment in at least one additional cognitive domain, both of which interfere with social function or activities of daily living (ADL). ADL impairment has come to define the threshold for the diagnosis of dementia beyond the identification of a cognitive abnormality. The NINCDS–ADRDA clinical criteria of probable AD do not require evidence of interference with social or occupational functioning but they include the specification that the onset of AD is insidious and that there is a lack of other systemic or brain diseases that may account for the progressive memory and other cognitive deficits. The currently accepted criteria support a probabilistic diagnosis of AD within a clinical context where there is no definitive diagnostic biomarker. A definite diagnosis of AD is only made according to the NINCDS–ADRDA criteria when there is histopathological confirmation of the clinical diagnosis.

Since the publication of the NINCDS–ADRDA criteria in 1984, the elucidation of the biological basis of AD has advanced greatly, allowing an unprecedented understanding of the disease process. The clinical phenotype of AD is no longer described in exclusionary terms, but can be characterised more definitively on a phenotypic basis. Distinctive markers of the disease are now recognised including structural brain changes visible on MRI with early and extensive involvement of the medial temporal lobe (MTL), molecular neuroimaging changes seen with PET with hypometabolism or hypoperfusion in temporoparietal areas, and changes in cerebrospinal fluid biomarkers. A driving force behind this emerging identity of AD has been the intense research interest in characterising the earliest stages of AD that predate the crossing of the dementia threshold, defined by functional disability. Prodromal AD (see glossary, panel 1) must be distinguished within the broad and heterogeneous state of cognitive functioning that falls outside normal ageing. This state has been described by a wide range of nosological terms including age-associated memory impairment, age-related cognitive decline, age-associated cognitive decline, mild cognitive disorder, mild neurocognitive disorder, cognitively impaired not demented, and mild cognitive impairment. Mild cognitive impairment (panel 1) is the most widely used diagnostic term for the disorder in individuals who have subjective memory or cognitive symptoms, objective memory or cognitive impairment, and whose activities of daily living are generally normal. Progression to clinically diagnosable dementia occurs at a higher rate from mild cognitive impairment than from an unimpaired state, but is clearly not the invariable clinical outcome at follow-up. A more refined definition of AD is still needed to reliably identify the disease at its earliest stages.

The case for revising the research criteria for AD diagnosis

There are several factors that highlight the need to update the current research criteria for AD.
Panel 1: Glossary of terms

Mild cognitive impairment
Varially defined but includes subjective memory or cognitive symptoms or both, objective memory or cognitive impairment or both, and generally unaffected activities of daily living; affected people do not meet currently accepted dementia or AD diagnostic criteria

Amnestic mild cognitive impairment
A more specified term describing a subtype of mild cognitive impairment, in which there are subjective memory symptoms and objective memory impairment; other cognitive domains and activities of daily living are generally unaffected; affected people do not meet currently accepted dementia or AD diagnostic criteria

Preclinical AD
The long asymptomatic period between the first brain lesions and the first appearance of symptoms and which concerns normal individuals that later fulfil AD diagnostic criteria

Prodromal AD
The symptomatic predementia phase of AD, generally included in the mild cognitive impairment category; this phase is characterised by symptoms not severe enough to meet currently accepted diagnostic criteria for AD

AD dementia
The phase of AD where symptoms are sufficiently severe to meet currently accepted dementia and AD diagnostic criteria

Insufficient diagnostic specificity
The DSM-IV-TR and NINCDS–ADRDA criteria have been validated against neuropathological gold standards with diagnostic accuracy ranging from 65–96%.11–14 However, the specificity of these diagnostic criteria against other dementia is only 21–88%.13,15 The accuracy of these estimates is difficult to assess, given that the neuropathological standard is not the same in all studies. Nevertheless, the low specificity must be addressed through both revised AD and accurate non-AD dementia diagnostic criteria

Improved recognition of non-AD dementia
Since the publication of the NINCDS–ADRDA criteria, operational definition and characterisation of non-AD dementias has improved. Entities for which there are diagnostic criteria include the frontotemporal lobar degenerations (frontotemporal dementia frontal variant, semantic dementia, progressive non-fluent aphasia),16–20 corticobasal degeneration,21 posterior cortical atrophy,22 dementia with Lewy bodies,23 and vascular dementia.24,25 Varma and colleagues26 showed that many of these disorders can fulfil the NINCDS–ADRDA criteria and it is likely that they have been included in AD research studies. Meanwhile, for each of these disorders, criteria have been developed that aim for high specificity.

The development of disease-specific criteria that are applicable in some cases before dementia is fully manifested has enabled the criteria to be used without going through the two-step process of dementia recognition (the syndrome) followed by the specific disease (the aetiology). For example, the identification of a dementia syndrome is not required for the diagnosis of primary progressive aphasia, corticobasal degeneration, or posterior cortical atrophy even though a dementia as currently defined will occur during or at the end of the course of these diseases.27 The histopathological diagnosis of the non-AD dementias has also advanced. In the example of frontotemporal lobar degeneration, the identification of ubiquitin-immunoreactive cytoplasmic and intranuclear inclusions as an important pathology in patients has reduced the neuropathological diagnostic prevalence of dementia lacking distinctive histopathology from 40% to 10% in autopsy series.28–30 There is no doubt that progress in the clinical definition of non-AD dementia improves the sensitivity of the currently accepted diagnostic criteria for AD by reducing the level of uncertainty.

Improved identification of AD phenotype
When the NINCDS–ADRDA criteria were first published, the authors noted that they were not yet fully operational because of insufficient knowledge about the disease.21 Since then, the clinical phenotype of AD has been much more clearly elucidated. In most patients with AD (86–94%), there is a progressive amnestic core that appears as an impairment of episodic memory.25,26 The pathological pathway of Alzheimer’s related changes has been fully described31,32 and involves the medial temporal structures (eg, entorhinal cortex, hippocampal formation, parahippocampal gyrus) early in the course of the disease. Moreover, the episodic memory disorder of AD correlates well with the distribution of neurofibrillary tangles within the MTL3 and with MRI volumetric loss of the hippocampus,34 structures known to be critical for episodic memory. The availability of neuroimaging techniques that can reliably measure the MTL have further supported this vital cliniconeuroanatomic correlation.

Need to test early intervention
The rapid growth of knowledge about the potential pathogenic mechanisms of AD including the amyloidopathy and tauopathy has spawned numerous experimental therapeutic approaches to enter into clinical trials. There is accruing evidence that, years before the onset of clinical symptoms, there is an AD process evolving along a predictable pattern of progression in the brain.30,35 The neurobiological advantage of earlier intervention within this cascade is clear. Earlier intervention with disease-modifying therapies36 is likely to be more effective when there is a lower burden of amyloid and hyperphosphorylated tau
and may truncate the ill effects of secondary events due to inflammatory, oxidation, excitotoxicity, and apoptosis. Early intervention may also be directly targeted against these events because they may play an important part in early phases of AD. By the time there is clear functional disability, the disease process is significantly advanced and even definitive interventions are likely to be suboptimal. Revised research criteria would allow diagnosis when symptoms first appear, before full-blown dementia, thus supporting earlier intervention at the prodromal stage (figure).

Problems with definition of mild cognitive impairment
One of the proposed advantages of mild cognitive impairment has been its potential usefulness for clinical trials directed at delaying the time to onset of AD. A series of large randomised controlled trials with both non-steroidal anti-inflammatory drugs and acetylcholinesterase inhibitors have sought to establish the usefulness of these drugs in delaying the conversion of mild cognitive impairment to AD. However, the lessons learned have highlighted the problems of mild cognitive impairment within this type of randomised controlled trial. With only small variations in the inclusion criteria for mild cognitive impairment, four trials (ADCS-MIS, InDDEx, Gal-Int II, and Rofecoxib) have had a very wide range of annual rates of progression to AD dementia (panel 1, 4.8–16%). The intention in these trials on mild cognitive impairment was to include many individuals with prodromal AD (ie, individuals with symptoms not sufficiently severe to meet currently accepted diagnostic criteria) that later progress to meet these criteria. When the mild cognitive impairment inclusion criteria of these trials were applied to a cohort of memory clinic patients in an observational study, they had diagnostic sensitivities of 46–88% and specificities of 37–90% in identifying prodromal AD. Given these numbers, these trials have clearly treated many patients who do not have AD or are not going to progress to AD for a long time. This has diluted the potential for a significant treatment effect and may have contributed to the negative outcomes where none of these drugs were successful at delaying the time to diagnosis of AD. These trials have also incurred significant costs where sample sizes of 750 to 1000 have been called for with durations of 3–4 years. Increasing the severity of mild cognitive impairment needed for inclusion in trials might improve sensitivity, specificity, and predictive values. However, participants would then be much closer to the current dementia threshold and would have a greater pathological burden, making the clinical gain marginal and disease modification difficult.

Neuropathological findings in mild cognitive impairment have also reinforced the heterogeneity of the clinical disorders subsumed under the definition mild cognitive impairment. To address the recognised clinical and pathological heterogeneity, it has been proposed that subtyping of mild cognitive impairment might be useful. The term amnestic mild cognitive impairment (panel 1) has been proposed to include individuals with subjective memory symptoms, objective memory impairment, and with other cognitive domains and activities of daily living generally assessed as being normal. However, only 70% of a selected cohort of people with amnestic mild cognitive impairment clinically identified to have progressed to dementia actually met neuropathological criteria for AD. This finding indicates that applying the criteria for this subtype of mild cognitive impairment clinically, without other objective evidence such as neuroimaging or results of cerebrospinal fluid analyses, will lack specificity for predicting the future development of AD since at least 30% of these will have non-AD pathology. If 30% of cases enrolled in a study that assesses drugs targeting amyloid or neurofibrillary tangles were to have non-AD pathology, there would be a substantial loss of power and possibly a conclusion that a medication was not effective. If all cases could be ascertained as having AD, a positive outcome might result. Thus, the most accurate determination that an individual has prodromal AD is critical.

In the planning of trials of disease-modifying treatment, special care will be needed to limit not only the exposure of potentially toxic therapies to those with prodromal AD but also to reliably exclude those who are destined to develop non-AD dementia. Our proposal for multidimensionally established identification of AD would place greater potential superiority to the intrinsically heterogeneous state of mild cognitive impairment and would advance the concept of mild cognitive impairment to its natural next level of more desirably identifying prodromal AD.

Unclear distinction between mild cognitive impairment and AD
The transition from mild cognitive impairment to AD has been an a priori primary endpoint in several randomised controlled trials. There is an inherent arbitrariness in...
determining a binary outcome, that is, conversion or no conversion, when the underlying disease is a continuous process. Individual clinicians’ experience in dementia diagnosis and the quality of the information they receive on the cognitive and functional status of patients will affect the threshold of detection of the transition to AD. Our revised research criteria will eliminate the mild cognitive impairment construct, thus bypassing the binary outcome in the clinical categorisation process associated with it as well as problems with reliability.

New biomarkers for AD

Over the past two decades since the NINCDS–ADRDA criteria were published, great progress has been made in identifying the AD-associated structural and molecular changes in the brain and their biochemical footprints. MRI enables detailed visualisation of MTL structures implicated in the core diagnostic feature of AD. PET with fluorodeoxyglucose (FDG) has been approved in the USA for diagnostic purposes and is sensitive and specific in detecting AD in its early stages. Cerebrospinal fluid biomarkers for detecting the key molecular pathological changes in the brain and their biochemical footprints (primary progressive aphasia, visuospatial dysfunction) where an antemortem diagnosis would at best receive the designation of possible AD from the framework. This may change in the future as work on diagnostic biomarkers advances and reliance on a well characterised clinical phenotype is lessened. In the absence of completely specific biomarkers, the clinical diagnosis of AD can still be only probabilistic, even in the case of typical AD. To meet criteria for probable AD, an affected individual must fulfil criterion A (the core clinical criterion) and at least one or more of the supportive biomarker criteria: B, C, D, or E (panel 2).

Core diagnostic criterion: early episodic memory impairment (A)

1. Gradual and progressive change in memory function at disease onset reported by patients or informants for a period greater than 6 months

The reporting of subjective memory complaints is a common symptom in an ageing population, at a prevalence that far exceeds the risk of being classified as having AD. Subjective memory complaints in elderly people may result from normal ageing or various medical disorders, and they are commonly associated with depression. However, such self-reported symptoms are associated with a high risk of future development of AD and, therefore, should be carefully taken into account. The perceptions of patients’ symptoms from an informant or proxy are perhaps more significant as they are more strongly related to objective memory performance and are predictive of progression to AD. To satisfy criterion A, memory symptoms must start gradually and show progressive decline over at least 6 months. Particular attention should be paid to intraindividual decline, which improves the identification of those individuals with prodromal AD.

2. Objective evidence of significantly impaired episodic memory on testing

A diagnosis of AD requires an objective deficit on memory testing (recall deficit with intrusions). Tests of
AD is considered definite if the following are present:

**Criteria for definite AD**

- Both clinical and histopathological (brain biopsy or autopsy) evidence of the disease, as required by the NIA-Reagan criteria for the post-mortem diagnosis of AD; criteria must both be present
- Both clinical and genetic evidence (mutation on chromosome 1, 14, or 21) of AD; criteria must both be present

**Probable AD: A plus one or more supportive features B, C, D, or E**

**Core diagnostic criteria**

A. Presence of an early and significant episodic memory impairment that includes the following features:
   - Gradual and progressive change in memory function reported by patients or informants over more than 6 months
   - Objective evidence of significantly impaired episodic memory on testing: this generally consists of recall deficit that does not improve significantly or does not normalise with cueing or recognition testing and after effective encoding of information has been previously controlled
   - The episodic memory impairment can be isolated or associated with other cognitive changes at the onset of AD or as AD advances

**Supportive features**

B. Presence of medial temporal lobe atrophy
   - Volume loss of hippocampi, entorhinal cortex, amygdala evidenced on MRI with qualitative ratings using visual scoring (referenced to well characterised population with age norms) or quantitative volumetry of regions of interest (referenced to well characterised population with age norms)
   - Abnormal cerebrospinal fluid biomarker
      - Low amyloid $\beta_{1-42}$ concentrations, increased total tau concentrations, or increased phospho-tau concentrations, or combinations of the three
      - Other well validated markers to be discovered in the future
   - Specific pattern on functional neuroimaging with PET
      - Reduced glucose metabolism in bilateral temporal parietal regions
      - Other well validated ligands, including those that foreseeably will emerge such as Pittsburgh compound B or FDDNP
   - Proven AD autosomal dominant mutation within the immediate family

**Exclusion criteria**

- Sudden onset
- Early occurrence of the following symptoms: gait disturbances, seizures, behavioural changes

**Clinical features**

- Focal neurological features including hemiparesis, sensory loss, visual field deficits
- Early extrapyramidal signs

Other medical disorders severe enough to account for memory and related symptoms
- Non-AD dementia
- Major depression
- Cerebrovascular disease
- Toxic and metabolic abnormalities, all of which may require specific investigations
- MRI FLAIR or T2 signal abnormalities in the medial temporal lobe that are consistent with infectious or vascular insults

**Criteria for definite AD**

AD is considered definite if the following are present:

- Both clinical and histopathological (brain biopsy or autopsy) evidence of the disease, as required by the NIA-Reagan criteria for the post-mortem diagnosis of AD; criteria must both be present
- Both clinical and genetic evidence (mutation on chromosome 1, 14, or 21) of AD; criteria must both be present

Delayed recall discriminate very mild AD from normal healthy controls with high accuracy (>90%). Such tests also predict prodromal AD better than other memory or non-memory measures, with accuracy greater than 80%. Delayed recall is a reliable predictor of AD in individuals with mild cognitive impairment. A meta-analysis of 47 studies calculated pooled effect sizes between incident AD and control groups. Episodic memory yielded the largest pooled effect sizes (>1 SD), along with executive functioning and perceptual speed. Within episodic memory, delayed recall testing showed a larger effect size than immediate recall testing.

Impaired delayed recall is not itself evidence of an AD-related memory disorder. Genuine deficits in encoding and storage processes that are characteristic for AD must be distinguished from non-AD deficits that can also affect delayed recall, including attentional difficulties that may be present in depression, or inefficient retrieval strategies associated with normal ageing, frontotemporal dementia, or subcortical-frontal dementias. The accurate diagnosis of the episodic memory deficit of AD can be improved by use of test paradigms that provide encoding specificity. Within such paradigms, test materials are encoded along with specific cues—for example, semantic cues, which are used to control for an effective encoding and are subsequently presented to maximise retrieval. Coordinated encoding and retrieval paradigms of this type include the free and cued recall test or similar cued recall paradigms. Within these neuropsychological test paradigms, measures of sensitivity to semantic cueing can successfully differentiate patients with AD from healthy controls, even when patients are equated to controls on mini-mental state examination scores or when disease severity is very mild. Buschke and co-workers derived sensitivity and specificity estimates of 93% and 99%, respectively, for the discrimination of patients with mild AD from healthy people with such a strategy. Patients with non-AD disorders including progressive supranuclear palsy and Parkinson’s and Huntington’s diseases do almost as well as control individuals under encoding specificity conditions whereas patients with AD do not normalise their recall deficit. Patients with very mild AD also have a measurable reduction in sensitivity to cueing. Reduced benefit from cueing at recall reliably identifies prodromal AD.
comprehension); praxis (impaired imitation, production, or recognition of gestures); and complex visual processing and gnosis (impaired recognition of objects or faces). The emergence of neuropsychiatric symptoms, including apathy or delusions, also constitutes a clinical marker of disease.9,77 Neuropsychiatric symptoms cannot be a core diagnostic feature because they are less specific and generally occur at a high prevalence later in the course of the disease. When there is evidence of impairment in multiple cognitive domains, functional disability, and neuropsychiatric symptoms, a more widespread diffusion of neuronal lesions in cortical and subcortical structures can be established.8 However, even in these more advanced cases, there should be evidence of an early and previous episodic memory deficit as a mandatory requirement for the diagnosis of AD.

Supportive features

Atrophy of medial temporal structures on MRI (B)

Atrophy of the MTL on MRI seems to be common in AD (71–96%, depending on disease severity), frequent in mild cognitive impairment (59–78%), but less frequent in normal ageing (29%).6,7 MTL atrophy is related to the presence of AD neuropathological and its severity, both in terms of fulfillment of AD neuropathological criteria and Braak stages.44,88 MRI measurements of MTL structures include qualitative ratings of the atrophy in the hippocampal formation46 or quantitative techniques with tissue segmentation and digital computation of volume.42 Both techniques can reliably separate AD group data from normal age-matched control group data, with sensitivities and specificities greater than 85%.6,14–15 Hippocampal volumes can distinguish AD equally at younger (<70 years) and old ages (>70 years).46 Qualitative measures have been useful in distinguishing patients with AD from those with non-AD dementia including vascular, frontotemporal, and dementia of unspecified cause, with combined mini-mental state examination and MTL atrophy ratings yielding sensitivity and specificity greater than 85%.47

In studies of mild cognitive impairment, the accuracy of MTL atrophy measures in identifying prodromal AD has been generally lower, possibly because individuals who did not meet currently accepted AD diagnostic criteria at study completion included some cases that would have done so at a later time. Qualitative MTL ratings can identify prodromal AD; however, the sensitivities and specificities, respectively, of 51–70% and 68–69%46 at present limit their usefulness. The predictive usefulness of quantitative measures of hippocampal volume in identifying prodromal AD is inconsistent.6,48,53–56 Measures of hippocampal subfields might be more useful than measures of the entire structure.55,56 Other structures or combinations of structures within and beyond the MTL may prove to be more sensitive to early AD pathology. For example, entorhinal cortex volume identifies prodromal AD more accurately than hippocampal volume, with a sensitivity of 83% and specificity of 73%.59 There are, however, technical difficulties in measuring this region that must be resolved.60,61 Combinations of MTL volumes and lateral temporal lobe or anterior cingulate volumes also detect prodromal AD with variable success (sensitivity 68–93%, specificity 48–96%).100,101

There is a strong correlation between MTL volumes and episodic memory performance.11,12 In turn, there is a potential incremental value of MTL measurement beyond the episodic memory impairment in the identification of prodromal AD. In several studies, MTL measures (quantitative and qualitative) contributed independently of memory scores to the identification of prodromal AD.60,61 The reported accuracy of identifying prodromal AD increased from 74% to 81%106 and from 88% to 96%46 when MTL measures were added to age and memory scores, respectively.

Inclusion of MTL atrophy as a diagnostic criterion of AD, irrespective of the age at onset,9 mandates exclusion of other causes of MTL structural abnormality including bilateral ischaemia, bilateral hippocampal sclerosis, herpes simplex encephalitis, and temporal lobe epilepsy. T2 weighted MRI, coupled to history and examination, and potentially adjunctive directed tests such as cerebrospinal fluid analysis and EEG should facilitate this discrimination.103,104

Abnormal cerebrospinal fluid biomarkers (C)

In the NINCDS–ADRDA guidelines, cerebrospinal fluid examination was recommended as an exclusion procedure for non-AD dementia, due to inflammatory disease, vasculitis, or demyelination.7 Since then, there has been a lot of research into the usefulness of AD-specific biomarkers that are reflective of the central pathogenic processes of amyloid β aggregation and hyperphosphorylation of tau protein. These markers have included amyloid β1–42 (Aß42), total tau (t-tau), and phospho-tau (p-tau).105–107 In AD, the concentration of Aß42 in cerebrospinal fluid is low and that of t-tau is high compared with those in healthy controls.105,106 Concentrations of different phosphorylated tau epitopes may also be high.107 Aß42 concentration in the cerebrospinal fluid is normal in patients with depression and decreased in dementia with Lewy bodies, frontotemporal lobar degeneration, and vascular dementia.10 This lack of specificity is not fully explained, but may relate to the lack of histopathological verification or the presence of comorbid AD. T-tau concentration is normal in depression, may be slightly raised in dementia with Lewy bodies and frontotemporal lobar degeneration, and is very high in Creutzfeldt-Jakob disease.10,11 Measurement of the concentration of p-tau, notably p-tau 231, increases the specificity for AD, especially in contrast to frontotemporal lobar degeneration.109 The pooled sensitivity and specificity for Aß42 in AD versus controls from 13 studies involving 600 patients and 450 controls were
86% and 90%, respectively. For t-tau, the sensitivity was 81% and the specificity 90%, pooled from 36 studies with 2500 patients and 1400 controls. Across 11 studies with a total of 800 patients and 370 controls, p-tau had a mean sensitivity of 80% when specificity was set at 92%, but sensitivities varied widely among studies using different methods. By use of a combination of concentrations of Aβ42 and t-tau for AD versus controls, high sensitivities (85–94%) and specificities (83–100%) can be reached.

Several recent studies have specifically addressed the value of cerebrospinal fluid biomarkers in identifying prodromal AD. Combinations of abnormal markers (low Aβ42, high t-tau, high p-tau) reached a hazard ratio of 17 to 20 for predicting AD in a follow-up of 4–6 years. Sensitivities and specificities in this study were >90% and >85%, respectively, which agreed with those in a similar one with much shorter follow-up (1 year). This high diagnostic usefulness of cerebrospinal fluid markers in the mild cognitive impairment stage supports their incremental value over memory impairment in the diagnostic scheme and justifies their inclusion as a diagnostic criterion.

Using an adapted spinal needle (Sprotte 24 g), lumbar puncture can be done with a very low rate of clinically significant adverse events and with a good acceptability in cognitively impaired people and healthy adults of all ages.

Specific metabolic pattern evidenced with molecular neuroimaging methods (D)

PET and single photon emission computed tomography (SPECT) are in vivo nuclear radioisotopic scans that can measure blood flow (99mTc-HMPAO or 133Xe), glucose metabolism (18F-FDG PET), and, more recently, protein aggregates of amyloid and tau. Within an AD diagnostic framework their ideal role is to increase the specificity of clinical criteria.

A reduction of glucose metabolism as seen on PET in bilateral temporal parietal regions and in the posterior cingulate is the most commonly described diagnostic criterion for AD. A recent meta-analysis of nine studies reported that for the discrimination of patients with AD from healthy controls, pooled sensitivities and specificities were 86% for temporoparietal hypometabolism. There is a wide range for both sensitivities and specificities, without clear explanation of this variability. When histopathological examination has been used as the gold standard, sensitivity is 88–95% and specificity is 62–74%. Because of cognitive reserve, the reduction in temporoparietal glucose metabolism diagnostic for AD might vary with educational attainment or IQ at any level of clinical severity.

PET has been successful in distinguishing dementia with Lewy bodies from AD, with sensitivity of 86–92% and specificity of 80–81% when visual-association cortex was considered. Discrimination from frontotemporal dementia has also been achieved, with sensitivity and specificity of 78% and 71%. There has been limited accuracy in differentiating AD from vascular dementia (sensitivity 75–88%, specificity 18–53%).

The usefulness of FDG-PET in the detection of prodromal AD has only just begun to be addressed in studies with small samples of patients with mild cognitive impairment and limited follow-up (≤3 years). Metabolic reductions in the anterior cingulate, posterior cingulate, and temporal, parietal, and medial temporal cortices detected prodromal AD, with accuracy estimates ranging from 75% to 84%. The potential incremental value of PET over other diagnostic markers in identifying prodromal AD is poorly defined. PET may be more accurate when delayed recall scores are severely impaired (sensitivity and specificity >90%).

There are promising PET techniques that provide in-vivo visualisation of amyloid and potentially neurofibrillary tangles. Studies using PiB (N-methyl-[11C]2-(4-[(1RS,2RS)-6-hydroxybenzothiazole]-2-propionamido)ethyl-2-naphthylamino-2-naphthyl-ethylidene)malononitrile) have shown a pattern of increased radioligand retention in patients with AD compared with control individuals that is consistent with AD pathology. Furthermore, positive cortical PiB binding has been associated with low cerebrospinal fluid Aβ42 concentrations in AD. These protein visualisation techniques clearly have the potential of increasing the usefulness of PET in AD within the diagnostic framework, but their diagnostic accuracy, in particular their specificity for AD, requires further investigation as there is evidence of high AD-like PiB retention in some healthy people and some people with mild cognitive impairment. AD-like PiB retention in healthy people might signal a preclinical AD state in asymptomatic individuals who later meet currently accepted dementia and AD diagnostic criteria, whereas in mild cognitive impairment it might reveal prodromal AD. Longitudinal follow-up is essential for the verification of the presumption that these are indeed preclinical and prodromal AD cases and not false positives.

Because SPECT is more widely available and cheaper than PET, it has received much attention as an alternative to PET. However, at present, the technique is not included in these proposed criteria as the diagnostic accuracy estimates for this modality generally fall below the requisite 80% levels specified by the Reagan Biomarker Working Group. 99m Tc-HMPAO SPECT identifies diagnosed AD with moderate sensitivity (77–80%) and specificity (65–93%). A pattern of bilateral temporal parietal hypoperfusion increases diagnostic certainty over clinical diagnosis alone.

According to a recent meta-analysis SPECT distinguished AD from non-AD in studies including healthy controls with pooled weighted sensitivities ranging from 65% to 71%, with a specificity of 79%. There are few SPECT studies that have adequately addressed the comparison between AD and non-AD...
dementias. The few that did provided a pooled weighted sensitivity and specificity for AD versus frontotemporal dementia of 71% and 78%, respectively, and for AD versus vascular dementia of 71% and 75%, respectively. More specific ligand methods, for example dopamine SPECT scanning with fluoropropyl-CIT, may have particular utility in distinguishing dementia with Lewy bodies and Parkinson’s disease dementia from AD (sensitivity 88%, specificity 85%).

Two small retrospective SPECT studies of mild cognitive impairment suggest that hypoperfusion in parietal and temporal lobe regions, and in the precuneus, may be brain functional patterns occurring very early in AD. In both studies, patterns of regional blood flow on SPECT distinguished prodromal AD with accuracy greater than 80%. These studies require replication with larger samples and prospective methodology before the technique can become a recommended criterion. The thioflavin derivative IMPY (6-iodo-2-(4-L-dimethylamino-)phenyl-imidazo[1,2-a]pyridine), which targets amyloid plaques for in vivo imaging in SPECT has not yet been investigated in living patients with either AD or mild cognitive impairment, but this ligand might help measure amyloid plaque burden in the future, thus providing an additional new approach and usefulness for functional imaging.

Familial genetic mutations (E)
Three autosomal dominant mutations that cause AD have been identified on chromosomes 21 (amyloid precursor protein), 14 (presenilin 1), and 1 (presenilin 2). The presence of a proband with genetic-testing evidence of one of these mutations can be considered as strong supportive for the diagnosis of AD for affected individuals within the immediate family who did not themselves have a genetic test for this mutation. If individuals with a positive mutation history of the described type present with the core amnestic criterion A, they will be considered to meet criteria for AD within the revised diagnostic framework. The gold standard for definite diagnosis of AD in this setting would of course be genetic testing and verification of a genetic mutation in these individuals.

Exclusion criteria
Probable AD diagnosis cannot be established if the illness begins with a sudden onset, has focal neurological findings including hemiparesis, sensory loss, visual field deficits, or where there are seizures, gait disturbances, or extrapyramidal signs at the onset or very early in the course of the illness. Other medical, neurological, or psychiatric disorders that could otherwise account for the deterioration in memory and related symptoms must be excluded. The diagnosis should be questioned in case of the following red flags: early behavioural disturbances (particularly disinhibition, euphoria, or psychosis), early extrapyramidal symptoms, early visual hallucinations, early visuospatial impairment, marked fluctuations in cognition and REM sleep behavioural disorders. The presence of cerebrovascular lesions, particularly white-matter lesions and lacunar infarctions both symptomatic and asymptomatic, are common with ageing. To establish probable AD, cerebrovascular disease that is sufficiently severe to account for the cognitive and functional deficits must be excluded. Probable AD cannot be diagnosed if the symptom profile suggestive of dementia with Lewy bodies (pronounced fluctuations in attention and cognition, recurrent prominent visual hallucinations, and motor parkinsonism) or if any other non-AD dementia is present. The presence of a delirium or toxic metabolic cause for the cognitive disorder precludes a diagnosis of probable AD (at least until the delirium has cleared) as does an unexplained altered state of consciousness.

Discussion
This working group has identified, by consensus, that new research criteria are timely, realistic, and feasible. Our proposed AD diagnostic framework (panel 2) is anchored around a core clinical phenotype supported by brain-structure abnormalities, molecular imaging impairment, biochemical changes, or genetic mutations associated with AD. The timeliness of these criteria is underscored by the many drugs in development that are directed at changing the disease pathogenesis through amyloid immunotherapy, gamma or beta secretase inhibitors and modulators, alpha secretase activators, tau kinase inhibitors, and nerve growth factors. Further new approaches directed at tau and tangles are foreseen. There is a neurobiological imperative to identify AD before the point of disease where irreversible pathological injury would prevent effective intervention. The proposed criteria should allow an earlier and more specific AD diagnosis than their predecessor, the NINCDS–ADRDA criteria.

These proposed criteria move away from the traditional two-step approach of first identifying dementia according to degree of functional disability, and then specifying its cause. Rather, they aim to define the clinical, biochemical, structural, and metabolic presence of AD. The cornerstone clinical criterion A specifies that there is an episodic memory deficit within test conditions of encoding specificity. This criterion should allow 86–94% of cases to be included. Beyond this core criterion, the presence of at least one biological footprint of the disease, either by criterion B (structural imaging), criterion C (cerebrospinal fluid), criterion D (molecular imaging), or criterion E (dominant mutation within the immediate family) is also needed to establish a positive diagnosis. The requirement, for diagnosis, of a clinical phenotype in combination with any one of the supportive features currently represents the most balanced approach because the clinical phenotype of AD is better known than its biological phenotype. We have no empirical basis at this
time for assigning different weightings to the supportive features or recommending combinations of features or, alternatively, requiring the presence of all. However, as new evidence accrues on biological markers for AD, especially those detecting AD-pathology specific markers such as amyloid imaging, the weighting of the supportive features may change. Other combinations may prove to have greater diagnostic accuracy or new features may be introduced. This will evolve as data sets gathered with all modalities are assessed.

We recognise that these criteria represent a cultural shift requiring more biologically focused work-up than previous approaches; however, this seems to be the best way to integrate the profound advances into the clinical arena. When effective disease-modifying medications are available, the argument for such biologically based studies will be even more compelling. Some research needs will be better addressed with a more stringent approach requiring that each diagnostic criterion be met. For example, proof-of-principle studies may benefit from the most highly selected AD study samples where the presence of all supportive features might be specified. This could maximise specificity for AD, but impose a substantial loss of sensitivity that would need to be re-addressed in later stages of development.

There are important differential diagnostic challenges that can be anticipated with the application of the proposed criteria. Of primary concern are non-AD amnestic disorders that can be associated with MTL damage including bilateral ischaemic injury, hippocampal sclerosis, limbic encephalitis, and temporal lobe epilepsy. Non-AD neurodegenerative disorders including tangle-only pathological changes and argyrophilic grain disease may also involve the MTL and limbic system. Depression can present with episodic memory impairment and MRI changes in hippocampal volume. These disorders could potentially satisfy criteria A and B, and in turn, where possible, must be ruled out in each instance. Careful clinical assessment with the use of specific memory tests, careful attention to T2 signal abnormalities within the MTL, and other investigations as clinically indicated will be called for to establish the AD diagnosis. We also cannot exclude that the non-AD dementias including frontaltemporal lobar degeneration, vascular dementia, and dementia with Lewy bodies may in some cases have the core clinical amnestic presentation specified in this framework. These non-AD dementias may also have positive molecular imaging or cerebrospinal fluid findings as has been shown in the reviewed evidence. If a non-AD cause is suspected, it must be ruled out carefully on a case-by-case basis by applying in parallel the diagnostic criteria for the other disorders.

We have specifically not addressed the issue of mixed disease for two reasons. First, in many instances mixed stands for comorbid disease being present but not being the principal cause of the dementia syndrome. As such, nothing has changed over current practice. Second, by narrowing the definition to a strict memory presentation with additional evidence for underlying AD pathological change, the chance of a patient fulfilling more than one set of criteria, as was the case with the NINCDS–ADRDA criteria, has actually been reduced. Even when a patient has abundant white-matter changes on MRI, thought to be of vascular origin, the presence of criterion A and the absence of overt dementia in the sense of the NINDS–AIREN criteria (ie, involvement of other cognitive domains), renders the patient more likely to have AD (with concomitant vascular disease) than vascular dementia. The MRI changes may still be used to guide therapy towards secondary prevention.

Moreover, the proposed criteria depict typical AD. There are also atypical forms of neuropathologically confirmed AD. These forms clearly deviate from the described amnestic presentation and include focal cortical syndromes, particularly posterior cortical degeneration where there is visual or visuospatial impairment, or frontal forms with prominent behavioural symptoms. The non-memory clinical phenotype might be influenced by the apolipoprotein genotype. Estimates of the relative prevalence of these atypical presentations have ranged from 6% to 14%. These presentations will still clearly elude diagnosis according to the revised criteria as they did in the NINCDS–ADRDA criteria. Their inclusion in research protocols remains too uncertain to be made with sufficient reliability and for this reason we have excluded them from the present diagnostic framework.

The strength of these proposed research criteria is the introduction of neurobiological measures on to the clinically based criteria. There are, however, many limitations and steps still needed. In their current formulation, these proposed diagnostic criteria still require decisions around how they are to be put into practice. For example, for the core criterion of significant episodic memory impairment, we have identified the memory test paradigms that can distinguish AD-associated deficits from other memory difficulties, but we have not defined a magnitude of deficit or the comparative norms that should be used. In structural imaging, we have not presented a specific best test or method for MTL atrophy. There remains uncertainty as to the most effective method of assessment, qualitative or quantitative, and for the latter, the specific region within the MTL for measurement. There is no specification of the amount of atrophy that is optimally diagnostic of AD. Within molecular neuroimaging, there are similar open questions with regard to which regions are optimally diagnostic, whether a qualitative versus quantitative approach should be taken, and what degree of hypometabolism is diagnostic. Finally, we have not specified which cerebrospinal fluid marker or combination of markers should be used to support a positive diagnosis. Concentrations of cerebrospinal fluid...
markers vary substantially with different assays but also with the same assay done in different centres, raising important questions about measurements and sources of error. Although most of these questions will receive empirical answers in the future, this will not entirely resolve an issue that is also philosophical, around whether an approach primarily based on applied clinical judgment or one based on fully operational definitions will work better for the research diagnosis of AD. At this point we favour the former approach of clinical judgment being applied to the determination of each criterion.

Validation studies of the proposed diagnostic criteria will clearly be needed because, inherent in this new definition of AD is the assumption that they indicate the presence of the neurodegenerative process of AD, including those cases presenting very early in the course of the disease. Their validity will need to be established for different types of discrimination and at different time points in the disease course, including discrimination of early AD from normal ageing, and AD from non-AD dementias. Two major strategies can be used for the validation of the proposed criteria. First, the criteria can be applied retrospectively to existing cohorts that have detailed investigations including neuropsychology, MRI, cerebrospinal fluid analysis, and PET. The current large cohort studies of the European Alzheimer Disease Consortium (EADC), the Alzheimer Disease Cooperative Study (ADCS), the Alzheimer Disease Neuroimaging Initiative (ADNI), and other ongoing studies will provide ideal sources within which to validate the diagnostic criteria of probable AD. Because these are all multicentre studies, they will also allow the assessment of measurement reliability across centres as a step towards standardising measurements and putting our proposed criteria into practice. Second, new prospective cohorts should be acquired that are followed to post-mortem. This prospective validation approach will need to focus on non-demented individuals with and without cognitive complaints. At their initial study visit, individuals should be assessed in parallel with the newly proposed criteria for AD and with the criteria for mild cognitive impairment. In subsequent visits, the stability of the diagnosis under these proposed criteria should be determined. In addition, the standard NINCDS–ADRDA criteria for AD should be applied. The first validation measure will be the sensitivity and specificity of our proposed criteria, obtained at baseline, for predicting cognitive decline as well as eventually meeting the existing AD criteria. In addition, the present criteria should be compared with the NINCDS–ADRDA criteria to verify whether we achieve the goal of increasing the specificity for diagnosis. Ultimately, the sensitivity and specificity for the pathological diagnosis of AD and other dementias will need to be assessed. In all future studies using these criteria, the supportive features used and the number of patients that have a positive MRI, or cerebrospinal fluid test, or PET or SPECT should be specified. Validation studies will necessarily begin with selected samples—for example those accrued in the ADCS, EADC, and ADNI—but validation will eventually need to be extended to unselected heterogeneous community samples. The most informative studies are those that will use the four criteria on patients at different stages of the disease, including the prodromal stage, with a long-term follow-up including post-mortem examination.

We recognise that the proposed research criteria require significant expertise, technical skills, and financial resources to allow the comprehensive assessment of MRI, PET, and cerebrospinal fluid. MRI will be contraindicated in some patients or not easily available in some countries, as may be the case for cerebrospinal fluid biomarkers or molecular neuroimaging with PET or SPECT. The multidisciplinary approach that is required for our diagnostic framework may not yet be feasible in all memory clinics and certainly not in most epidemiological studies. The validation studies being proposed will need to take place within highly specialised AD centres, and if successful, the research criteria will need to be adapted for use in standard clinical settings. We foresee that technically less demanding criteria for clinical settings might develop from the more technically challenging research criteria once these are validated.

Finally, these proposed criteria acknowledge the progress that has been made in the past two decades in refining our understanding of the neurobiology and clinical phenomenology of AD. Their usefulness will be determined in the future as investigators apply the criteria in a variety of research studies and as key issues in their application are resolved.

Contributors
Each of the authors has contributed to the writing and the revision of the paper and has approved the final version. In addition, most of the authors (16 out of 19) participated in the Florence meeting (June 2005) and in the related discussions about the revised criteria.

Conflicts of interest
We have no conflicts of interest.

Acknowledgments
The conference for the international working group held on June 9, 2005, was sponsored through a grant-in-aid received from Eisai, Janssen-Cilag, Novartis, and Servier. BD obtained financial support from several pharmaceutical companies (Servier, Eisai, Novartis, Lundbeck, and Janssen-Cilag) for the extra costs needed to organise this 1 day satellite symposium, which included airplane tickets and accommodation for the following experts: B Dubois, H H Feldman, P Barberge-Gateau, A Delacourte, S Gauthier, G Jelic, K Meguro, P Robers, M Rosson, S Salloway, Y Stern, and P Visser. We thank Marsel Mesulam for helpful comments on the article.

References
56 Storandt M, Grant EA, Miller JP, Morris JC. Longitudinal course and neuropsychologic outcomes in original vs revised MCI and in pre-MCI. Neurology 2006; 67: 467–73.


