

ORIGINAL ARTICLE

Accuracy of the Clinical Diagnosis of Alzheimer Disease at National Institute on Aging Alzheimer Disease Centers, 2005–2010

Thomas G. Beach, MD, PhD, Sarah E. Monsell, PhD, Leslie E. Phillips, PhD, and Walter Kukull, PhD

Abstract

The neuropathologic examination is considered to provide the gold standard for Alzheimer disease (AD). To determine the accuracy of currently used clinical diagnostic methods, clinical and neuropathologic data from the National Alzheimer's Coordinating Center, which gathers information from the network of National Institute on Aging (NIA)-sponsored Alzheimer Disease Centers (ADCs), were collected as part of the National Alzheimer's Coordinating Center Uniform Data Set (UDS) between 2005 and 2010. A database search initially included all 1198 subjects with at least one UDS clinical assessment and who had died and been autopsied; 279 were excluded as being not demented or because critical data fields were missing. The final subject number was 919. Sensitivity and specificity were determined based on "probable" and "possible" AD levels of clinical confidence and 4 levels of neuropathologic confidence based on varying neuritic plaque densities and Braak neurofibrillary stages. Sensitivity ranged from 70.9% to 87.3%; specificity ranged from 44.3% to 70.8%. Sensitivity was generally increased with more permissive clinical criteria and specificity was increased with more restrictive criteria, whereas the opposite was true for neuropathologic criteria. When a clinical diagnosis was not confirmed by minimum levels of AD histopathology, the most frequent primary neuropathologic diagnoses were tangle-only dementia or argyrophilic grain disease, frontotemporal lobar degeneration, cerebrovascular disease, Lewy body disease and hippocampal sclerosis. When dementia was not clinically diagnosed as AD, 39% of these cases met or exceeded minimum threshold levels of AD histopathology. Neurologists of the NIA-ADCs had higher predictive accuracy when they diagnosed AD in subjects with dementia than when they diagnosed dementing diseases other than AD. The mis-

diagnosis rate should be considered when estimating subject numbers for AD studies, including clinical trials and epidemiologic studies.

Key Words: Alzheimer disease, Autopsy, Clinical trials, Diagnosis, Histopathology, Neuropathology, Non-Alzheimer dementia.

INTRODUCTION

For Alzheimer disease (AD), as for any disease, it is extremely important to know, with precision and confidence, the accuracy of currently used clinical diagnostic methods. This information is critical for AD research, including epidemiologic studies, economic impact studies, and in particular, treatment and prevention trials. With the exception of those diseases caused by single gene defects, the most accurate diagnosis is obtained from the histologic examination of tissue samples from affected sites. For many disorders, this is possible during life through biopsy, but biopsy has long been contraindicated for AD because of a high risk/benefit ratio. However, it is generally accepted that histologic examination is the best indicator of AD diagnosis. The emergence of new biomarkers may have a major impact on clinical diagnostic practice (1); however, because measures of all biomarkers studied to date have considerable overlap with those found in other types of dementia (as well the cognitively normal elderly population), it remains difficult to determine whether cognitive impairment is caused by AD or another more dominant and concurrent process. Therefore, for the foreseeable future, autopsy will still serve the role of "gold standard" for the determination of clinical diagnostic accuracy rates.

The existing body of data on the accuracy of current clinical AD diagnostic methods shows much variability among studies (2–25). Sensitivity estimates have ranged between 41% and 100% (median, 87%), whereas specificity has ranged between 37% and 100% (median, 58%). Because many studies have reported only sensitivity or positive predictive value (PPV), a general impression has arisen that the clinical diagnosis of AD is extremely accurate. The interpretation of these data, however, is problematic because of differences in time and setting. In particular, whereas the clinical methods for diagnosing AD have not changed substantially since the introduction of the National Institute of Neurological Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (NINDS-ADRDA) criteria in 1984 (26), there have been several changes in neuropathologic diagnostic criteria,

From the Banner Sun Health Research Institute, Sun City, Arizona (TGB); and National Alzheimer's Coordinating Center, University of Washington, Seattle, Washington (SEM, LEP, WK).

Send correspondence and reprint requests to: Thomas G. Beach, MD, PhD, Banner Sun Health Research Institute, Civil Laboratory for Neuropathology, 10515 West Santa Fe Drive, Sun City, AZ 85351; E-mail: thomas.beach@bannerhealth.com

This study was supported by National Institute on Aging grants to the National Alzheimer's Coordinating Center (U01 AG016976) and the Arizona Alzheimer's Disease Core Center (P30 AG19610). Thomas Beach was supported by grants from the Arizona Department of Health Services, the Arizona Biomedical Research Commission, and the Michael J. Fox Foundation for Parkinson's Research.

including the Khachaturian criteria of 1985 (27), the Tierney criteria of 1988 (28), the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) criteria in 1991 (29), and the National Institute on Aging (NIA)–Reagan criteria in 1997 (30). The major goal in all of these efforts has been to determine the degree of AD histopathologic abnormalities necessary to cause dementia. Indeed, until the present time, the consensus of expert opinion has required that the clinically documented presence of dementia be part of the definition of AD.

This study used clinical and neuropathologic data collected by the National Alzheimer's Coordinating Center (NACC), which gathers information from the network of Alzheimer Disease Centers (ADCs) sponsored by the US NIA. The data analyzed in this study have been collected as part of the NACC Uniform Data Set (UDS) since 2005 (31) and thus represent the most current research practices. Because the last diagnostic accuracy studies using a complete ADC data set were performed in 1998 and 1999 (4, 23), it is important, for current research and clinical use, to obtain more appropriate figures.

MATERIALS AND METHODS

Subjects

Subject data were obtained with the assistance of NACC personnel through an NACC UDS database search. The NACC UDS has been collected since September 2005 from more than 30 ADCs located throughout the United States (31–33). Most ADCs are at university medical centers in urban areas. Research subjects are generally recruited from the practices of participating neurologists, with some additional community-based recruitment. The initial data pull included all 1198 subjects who had at least one UDS clinical assessment, had died, and were autopsied. From this group, 279 subjects were excluded as being recorded as “not demented” or because critical data fields were either not filled out or marked “missing” or “not done.” Subjects without dementia were excluded because it has been the common practice for both clinical and neuropathologic diagnostic definitions of AD to require the presence of dementia. Therefore, subjects with no dementia are not generally classified as AD. Excluded critical data fields were those used to

enter a response for the presence or absence of clinically probable AD and for the CERAD neuritic plaque density and Braak stage. After exclusion of these subjects, the final subject number for the study was 919.

Analysis Strategy

The study goal was to estimate the sensitivity and specificity of the clinical diagnosis of AD using the neuropathologic diagnosis as the gold standard. Clinical diagnosis was that given at the last assessment during life. Both the clinical and neuropathologic diagnoses were stratified by level of confidence. For the clinical diagnosis of AD, NINDS-ADRDA criteria (26) were stratified by considering “probable AD” alone as well as “probable” plus “possible AD.”

For the neuropathologic diagnosis of AD, the NIA-Reagan criteria (30) are the most current guidelines; these stratify the neuropathologic confidence level as “high,” “intermediate,” and “low.” Because of the idiosyncratic assignment of these categories from case to case (32), we instead stratified all relevant combinations of the deterministic histopathologic scores, consisting of the CERAD-defined neuritic plaque density score (29) and the Braak neurofibrillary tangle stage (33). These scoring methods have been shown to have a reasonably high level of reproducibility between observers and among research centers (34–37).

After the determination of specificities and sensitivities associated with the levels of clinical and pathologic diagnostic confidence, further analysis was directed at determining the final neuropathologic diagnoses for cases having mismatched clinical and neuropathologic diagnoses.

Statistical methods included calculation of sensitivity and specificity, with no adjustments made for age, gender, or other subject characteristics. Groups were compared using *t*-tests, analysis of variance, and Kruskal-Wallis analysis of variance. For all tests, the significance level was set at $p < 0.05$.

RESULTS

Subject Characteristics

Some subject characteristics are given in Table 1. Subjects were classified based on their clinical categorization as

TABLE 1. Clinical and Neuropathologic Data

Clinical Diagnosis	Age (Mean [SD]), yr	Gender	Interval Between Last Assessment and Death (Mean [SD]), mo	NP Density (Median; Mean [SD])	Braak Stage (Median; Mean [SD])
Included subjects (n = 919)	79.0 (11.4)	368 F/551 M	10.8 (9.1)	3; 2.1 (1.2)	5; 4.1 (1.9)
Excluded subjects (n = 279)	85.4 (11.0)*	142 F/137 M*	10.9 (7.8)	1; 1.2 (1.1)*	3; 2.6 (1.5)*
Probable AD (n = 526)	81.2 (10.3)	220 F/306 M	11.5 (9.2)	3; 2.5 (0.9)	5; 4.8 (1.5)
Possible AD (n = 122)	83.2 (11.6)	53 F/69 M	10.4 (8.4)	2; 1.8 (1.2)	4; 4.2 (1.6)
Not AD (n = 271)	72.8 (10.9)†	95 F/176 M	9.8 (9.0)	1; 1.3 (1.3)†	2; 2.7 (2.1)†

NP density is expressed on a 0 to 3 scale as follows: 0 = none; 1 = sparse; 2 = moderate; and 3 = frequent. The excluded subjects had incomplete data available for NP density and Braak stage assessment.

Probable AD, possible AD, and not AD refer to the clinical diagnosis.

*Significantly different from included subjects.

†Clinically probable, possible, and not AD groups are significantly different from each other.

AD, Alzheimer disease; F, female; M, male; NP, neuritic plaque.

TABLE 2. Sensitivity and Specificity of the Clinical Diagnosis of AD Relative to Stratified Clinical Confidence Levels and Minimum Threshold Levels for Histopathologic Severity

Neuropathologic AD Definition	Clinically Probable AD, n = 526	Clinically Probable or Possible AD, n = 648
CERAD NP Freq	n = 327	n = 373
Braak Stage V or VI n = 427	Sensitivity 76.6% Specificity 59.5%	Sensitivity 87.3% Specificity 44.3%
CERAD NP Mod or Freq	n = 366	n = 418
Braak Stage V or VI n = 486	Sensitivity = 75.3% Specificity = 63.0%	Sensitivity = 85.9% Specificity = 47.0%
CERAD NP Freq	n = 370	n = 421
Braak Stage III–VI n = 490	Sensitivity = 75.5% Specificity = 63.6%	Sensitivity = 85.9% Specificity = 47.1%
CERAD NP Mod or Freq	n = 438	n = 511
Braak Stage III–VI n = 618	Sensitivity = 70.9% Specificity = 70.8%	Sensitivity = 82.7% Specificity = 54.5%

Histopathologic severity is determined by specific combinations of CERAD NP and Braak neurofibrillary stage.

AD, Alzheimer disease; CERAD NP, Consortium to Establish a Registry for Alzheimer's Disease neuritic plaque density score; Mod, moderate; Freq, frequent; Braak Stage, Braak neurofibrillary tangle stage.

“probable AD,” “possible AD,” or “not AD.” The “not AD” group are those not clinically diagnosed as either probable or possible AD; this group included only non-AD dementias because subjects with no dementia were excluded from the study. “Probable AD” and “possible AD” were defined according to the NINDS-ADRDA guidelines (26). The mean age of the “not AD” group was significantly lower than that of the probable or possible AD groups (72.8, 81.2, and 83.2 years, respectively). The gender distribution was generally skewed toward more men but did not significantly differ among groups. The median scores for neuritic plaque

density and Braak stage were significantly different, with progressively lower scores moving from probable AD through possible AD and not AD groups. The 3 groups did not significantly differ in the mean interval between last clinical assessment and death.

The excluded subjects significantly differed from the 919 included subjects in terms of age, gender distribution, and AD-related histopathologic scores, but not in the mean interval between last clinical assessment and death.

Diagnostic Classification Comparison

Measures of agreement between stratified levels for the clinical and neuropathologic diagnosis of AD are shown in Table 2. Sensitivity ranged from 70.9% to 87.3%, and specificity ranged from 44.3% to 70.8%. In general, sensitivity was increased with more permissive clinical criteria, and specificity was increased with more restrictive clinical criteria, whereas the opposite was true for neuropathologic criteria. When optimizing for sensitivity and specificity, the best result was 70.9% sensitivity and 70.8% specificity. This was achieved when the clinical diagnosis was defined as probable AD and the neuropathologic diagnosis as moderate or frequent neuritic plaques with Braak stage III to VI.

Table 3 shows the PPV for the clinical diagnosis of AD, stratified by levels of clinical and neuropathologic confidence. The results were compared with the PPV that would result if all subjects with dementia were diagnosed as AD. The PPV for the clinical diagnosis of AD ranged from 46.0% to 83.3%, with the best result achieved when the clinical diagnosis was defined by probable AD and the neuropathologic diagnosis as moderate or frequent neuritic plaques with Braak stage III to VI (the prevalence of subjects at or above this histopathologic threshold was 67.2%). The PPV for clinically probable AD was consistently about 4.5% to 5% higher than that for probable/possible AD. The PPV for clinically probable

TABLE 3. Positive Predictive Value of the Clinical Diagnosis of AD Stratified by Clinical Confidence Levels and Minimum Threshold Levels for Histopathologic Severity

Neuropathologic AD Definition	Clinically Probable AD, n = 526	Clinically Probable or Possible AD, n = 648	All Dementia Diagnosed as AD, n = 919
CERAD NP Freq	n = 327	n = 373	n = 427
Braak Stage V or VI n = 427	PPV = 62.2%	PPV = 57.6%	PPV = 46.0%
CERAD NP Mod or Freq	n = 366	n = 418	n = 486
Braak Stage V or VI n = 486	PPV = 69.6%	PPV = 64.7%	PPV = 52.9%
CERAD NP Freq	n = 368	n = 421	n = 490
Braak Stage III–VI n = 490	PPV = 70.0%	PPV = 65.0%	PPV = 53.3%
CERAD NP Mod or Freq	n = 438	n = 511	n = 618
Braak Stage III–VI n = 618	PPV = 83.3%	PPV = 78.8%	PPV = 67.2%

Results are compared with the PPV achieved if all subjects with dementia were diagnosed as AD. Histopathologic severity was determined by specific combinations of CERAD neuritic plaque density (CERAD NP) and Braak neurofibrillary stage.

AD, Alzheimer disease; CERAD, the Consortium to Establish a Registry for Alzheimer's Disease; PPV, positive predictive value; Mod, moderate; Freq, frequent; Braak Stage, Braak neurofibrillary tangle stage.

AD was consistently approximately 16% higher than that resulting if all subjects with dementia were considered to have clinical AD.

Analysis of Mismatched Clinical and Neuropathologic Diagnoses

Of the 526 subjects diagnosed as clinically probable AD, 438 were confirmed as neuropathologic AD, as defined above, and 88 did not meet neuropathologic criteria. For this analysis, a relatively permissive neuropathologic definition was used, that is, CERAD neuritic plaque density of moderate or frequent in combination with any Braak neurofibrillary tangle stage between III and VI, inclusive.

The primary neuropathologic findings for the cases not meeting the defined lower threshold for histopathologic severity are summarized in Table 4. The most frequent primary neuropathologic diagnosis among these subjects was AD, primary (NACC database code NPPAD), assigned by the neuropathologist to 17 cases, despite the relatively low levels of AD histopathology. Other relatively frequent primary neuropathologic findings were tangle-only dementia or argyrophilic grain disease (15 cases; NACC database code NPTAU), frontotemporal lobar degeneration (15 cases; NACC database code NPPFTLD), cerebrovascular disease (10 cases; NACC database code NPPVASC), Lewy body disease, with or without AD (9 cases; NACC database code NPPLEWY), and hippocampal sclerosis (9 cases; NACC database code NPPHIP). A small number of cases received primary neuropathologic diagnoses of progressive supranuclear palsy (3 cases), corticobasal degeneration (2 cases), and neuroaxonal dystrophy/Hallervorden-Spatz-like disease (2 cases); there were several

TABLE 4. Primary Neuropathologic Diagnosis for the 88 Subjects Clinically Diagnosed as Probable AD But Not Meeting a Defined Minimum Threshold Level of Histopathologic Severity

Primary Neuropathologic Findings	No. Cases
Primary neuropathologic diagnosis of AD despite low level of AD histopathology	17
Tangle-only dementia or argyrophilic grain disease	15
Frontotemporal lobar degeneration*	15
Cerebrovascular disease	10
Lewy body disease, with or without AD†	9
Hippocampal sclerosis, with or without AD‡	9
Progressive supranuclear palsy	3
Corticobasal degeneration	2
Neuroaxonal dystrophy/Hallervorden-Spatz-like condition	2
Miscellaneous (1 case each of amyloid angiopathy, "small vessel disease," "TDP-43 proteinopathy," limbic encephalitis, Rosenthal fiber encephalopathy, "clinical dementia, no neuropathological substrate")	6

Criteria for Alzheimer disease (AD) diagnosis were based on the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) neuritic plaque density of moderate or frequent in combination with any Braak neurofibrillary tangle stage between III and VI, inclusive.

*Of cases with frontotemporal lobar degeneration, 7 had ubiquitin or TDP-43-positive inclusions; 3 had tauopathies.

†Of cases with primary Lewy body disease, 2 also had a contributory diagnosis of AD.

‡Of cases with primary hippocampal sclerosis, 1 also had a contributory diagnosis of AD.

TABLE 5. Primary Neuropathologic Findings in the 271 Subjects Clinically Diagnosed as Not Being Either Probable or Possible AD

Primary Neuropathologic Diagnosis*	No. Cases
Histopathologically defined AD	107†
Frontotemporal lobar degeneration	60‡
Lewy body disease, with or without AD	31§
Creutzfeldt-Jakob disease and other prion encephalopathies	23
Progressive supranuclear palsy	18
Tangle-only dementia or argyrophilic grain disease	9
Corticobasal degeneration	8
Pick's disease	6
Cerebrovascular disease	6
Hippocampal sclerosis, with or without AD‡	2
Amyotrophic lateral sclerosis	2
Miscellaneous (1 case each of neuronal intermediate filament disease, "leukodystrophy" and cerebellar atrophy)	3

*One hundred fifty-four subjects had primary neuropathologic findings other than AD. Some subjects had more than 1 diagnosis that could be considered primary.

†The largest category is the 107 cases that, despite the clinical diagnosis, did not meet a defined minimum threshold level of histopathologic severity, that is, the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) neuritic plaque density of moderate or frequent in combination with any Braak neurofibrillary tangle stage between III and VI, inclusive.

‡Of cases with frontotemporal lobar dementia, 19 had ubiquitin- or TDP-43-positive inclusions; 9 had tauopathies.

§Of cases with primary Lewy body disease, 3 also had a contributory diagnosis of AD.

||Of cases with progressive supranuclear palsy, 3 also had a contributory diagnosis of AD.

other miscellaneous neuropathologic diagnoses (1 case each) (Table 4).

There were 271 subjects who were clinically diagnosed as not having either probable or possible AD (Table 5). Of these, 107 cases met a minimum histopathologic threshold for AD, that is, neuritic plaque density of moderate or frequent in combination with any Braak neurofibrillary tangle stage between III and VI, inclusive, despite their negative clinical diagnoses. There were 164 cases that had clinical diagnoses other than AD and that were confirmed neuropathologically, in that minimum AD histopathology thresholds were not present. The most frequent primary neuropathologic diagnoses in these cases were frontotemporal lobar degeneration (60 cases), followed by Lewy body disease (39 cases), Creutzfeldt-Jakob disease (20 cases), and progressive supranuclear palsy (18 cases). Smaller numbers of cases were diagnosed as tangle-only dementia, argyrophilic grain disease, Pick disease, corticobasal degeneration, cerebrovascular disease, hippocampal sclerosis, and amyotrophic lateral sclerosis. Many cases had contributing neuropathologic diagnoses in addition to their primary diagnosis; the most frequent contributing diagnoses were similar to the most frequent primary diagnoses.

DISCUSSION

Assessment of diagnostic accuracy rates necessarily requires a gold standard for accuracy that is ideally presumed to be 100% correct. For AD, the neuropathologic diagnosis

has always been considered the gold standard, however, because the neuropathologic criteria for AD have changed several times during the past 30 years (27–30), the question arises “how good is the present neuropathologic gold standard?” The neuropathologic criteria for AD used by all NIA ADCs since 1997 have been the NIA-Reagan criteria (30). These, however, do not give entirely specific guidelines for making the diagnosis, but rather set probability levels, given threshold levels of histopathologic severity (CERAD neuritic plaque density and Braak neurofibrillary stage) for when dementia may be caused by AD. Thus there are “low,” “intermediate,” and “high” probability designations for classifying autopsy subjects, although definitive recommendations on which of these 3 levels should be attained to justify the conclusion that dementia is primarily caused by AD have not been defined. There are 25 possible combinations of CERAD neuritic plaque density and Braak neurofibrillary stage, but NIA-Reagan provides instructions for classification of only 3 (32, 38). Thus, a large fraction of subjects are not classifiable by NIA-Reagan criteria, and the neuropathologist is left to make an arbitrary assignment.

In a recent examination of NACC data, Nelson et al estimated that 18% of subjects with dementia fell outside of the NIA-Reagan guidelines (32); here, we found that when controls with no dementia were not included, 33% of subjects were not classifiable. The shortcomings of the NIA-Reagan criteria have been recognized, and a revised set of criteria are currently in press (39, 40). In a recent multivariate analysis, approximately two thirds of the variation in cognitive ability of elderly subjects was accounted for by variation in CERAD neuritic plaque density and Braak stage (41). We feel that our approach here has been cautious as the available body of knowledge does not allow neuropathologists to be dogmatic about exactly how much plaque and tangle pathology is necessary to cause dementia. It is likely that, just as some subjects with 90% coronary artery stenosis have myocardial infarctions and some do not, individual subjects will vary in their ability to withstand a given lesion density (perhaps because of cognitive reserve), differences in genetic background or differences in environmental exposures. Therefore, we thought it appropriate to offer a range of possible gold standards.

We found that sensitivity for AD diagnosis ranged from 70.9% to 87.3%, whereas specificity ranged from 44.3% to 70.8%. Sensitivity was increased with more permissive clinical criteria, that is, by allowing either probable or possible AD to serve as the clinical diagnosis, whereas specificity was increased with more restrictive criteria, that is, by only accepting “probable” as the clinical diagnosis. Changes in neuropathologic criteria had the opposite effect. More liberal criteria decreased sensitivity but increased specificity. For the set of minimum neuropathologic criteria that are perhaps most commonly used to define AD (i.e. CERAD neuritic plaque density of moderate or frequent and Braak stage III–VI), the sensitivity and specificity of the clinical diagnosis of probable AD were each approximately 71%, whereas for combined probable and possible AD, sensitivity was about 83%, with specificity about 55%. Comparisons with previous studies are difficult, but these data agree with the median sensitivity (87%) and specificity (58%) calculated from a large set of

previous publications (2–25). There is also relatively close agreement with the last 2 previous assessments of diagnostic accuracy using the total NIA-ADC data set: Mayeux et al (4) estimated a sensitivity of 93% and a specificity of 55%, and Tsuang et al (23) reported a sensitivity of 84% and a specificity of 50%.

In most studies, sensitivity is relatively high whereas specificity is low, and many studies have reported only sensitivity or PPV, which has led to a false impression that the clinical diagnosis of AD is extremely accurate. Whereas the clinical diagnosis of AD often is validated by neuropathologic examination, a clinical diagnosis of a non-AD dementia is often not verified by neuropathology. For selecting subjects for clinical trials, the relatively high sensitivity is somewhat reassuring because it means that a relatively small percentage of non-AD subjects will be included in clinical trials. The PPV, although not generally regarded as a good measure of diagnostic accuracy, is perhaps most relevant to clinical trial selection. The present results indicate that when the minimum neuropathologic threshold for diagnosis is defined as moderate or frequent neuritic plaques together with Braak stage III to VI, the PPV of the clinical diagnosis of probable AD is 83%, that is, 83% of subjects with that clinical diagnosis were confirmed neuropathologically to have AD lesions sufficient to cause dementia. In addition, when neurologists of the NIA ADCs diagnose clinical AD, the PPV is approximately 16% more accurate than if they diagnosed all dementia subjects with AD. However, even a modest level of diagnostic misclassification could have a significant effect on clinical trial calculations for minimum subject number and thus should be considered in clinical trial planning. For example, with an estimated response rate to a trial medication of 50%, a 20% misdiagnosis rate would lower the actual response rate to 40% (if non-AD dementia subjects do not respond to the medication), which would require an approximate doubling of subject recruitment to maintain statistical power.

For calculating diagnostic mismatches, we selected a relatively permissive combination, that is, moderate or frequent neuritic plaques (NPs) with Braak stage III to VI. Lowering the lesion density threshold further (e.g. allowing sparse neuritic plaques and/or Braak stage 0–II) seemed counterintuitive because a large fraction of cognitively unimpaired elderly possesses such attributes. The new NIA criteria use this same combination of NPs and Braak stage as the minimum lesion density necessary for attributing the cause of dementia to AD (39, 40). Some might argue that the decision of whether to make the final diagnosis of AD should be left to the judgment of the individual neuropathologist rather than to rigid consensus criteria. We agree that individual neuropathologists should have some flexibility, but deviations from consensus criteria should be only an occasional occurrence, otherwise, the definition of what constitutes AD becomes so variable as to hinder effective research.

Analysis of the subjects mismatched in terms of clinical and neuropathologic diagnosis reveals some interesting trends. Neuropathologists not infrequently diagnosed AD in subjects clinically diagnosed as AD but not meeting a minimum histopathologic threshold. This was despite the application of a relatively low minimum histopathologic threshold

(moderate or frequent neuritic plaques with Braak stage III–VI). This is most likely caused by the lack of a specified minimum histopathologic threshold for the primary neuropathologic diagnosis of AD in the 1997 NIA-Reagan criteria. Other relatively frequent neuropathologic findings are shown in Table 4. Of subjects clinically diagnosed as not having either probable or possible AD, the most frequent primary neuropathologic diagnosis was AD, accounting for 39% of the total. Although not formally analyzed here, we did note a high rate of multiple neuropathologic diagnoses in subjects with dementia, whereas clinicians rarely assign more than one diagnosis.

The reasons for the observed mismatches between clinical and neuropathologic diagnoses cannot be understood completely, but we offer the following tentative explanations. First, as extensively documented, the clinical diagnosis of dementia is relatively nonspecific (2–25). There are many alternate causes of dementia in the elderly, and AD has generally been attributed to be the neuropathologic cause in 65% to 75% of dementia cases (42, 43). Correlation studies between molecular genetics and neuropathology also indicate that, for several autosomal dominant dementias, the presenting clinical phenotype may vary widely, from various dementia subtypes to parkinsonism to motor neuron disease–like, even among individuals with the same causative mutation (44–47). These findings suggest that phenotype is not rigidly linked to etiology and may be moldable to a large extent by an individual's genetic and epigenetic background and/or total environmental exposure.

Because ADCs are tertiary referral centers, diagnostic accuracy rates may differ from those that may be achieved in secondary or primary care settings. For example, patients with a more complex clinical syndrome are more likely to be referred to a tertiary care center. Thus, there might be relatively more complicated cases that are more difficult to diagnose. On the other hand, the greater level of neurologic expertise at these centers would presumably contribute to more accurate diagnoses. Because large reported autopsy series are only done in tertiary care settings, it is not possible to know for certain whether such differences exist. However, one study has reported no evident selection bias between autopsied and non-autopsied subjects among an incident community-diagnosed dementia series, suggesting that dementia case composition may not be substantially different in different settings (48).

The 1997 NIA-Reagan criteria for the neuropathologic assessment of AD acknowledged that there are many uncertainties with defining the neuropathologic gold standard, and indeed the criteria were intended to be only a tentative starting point that would invite an organized assessment once sufficient data had accumulated. In the revision of the 1997 criteria, combinations of plaque and tangle pathologies considered sufficient to cause dementia and qualify for the neuropathologic diagnosis of AD are explicitly stated (40, 41). Topographic whole-brain amyloid staging has been incorporated as an adjunct to the preceding approach that was limited to lobar cortical plaque density estimates. Amyloid staging may be of immediate clinical diagnostic utility when used to guide amyloid imaging as the progression of amyloid plaques beyond the cortex appears to be a useful marker of higher Braak tangle

stage and of the presence of dementia (49, 50). In addition, more detailed advice has been given on how to assess common comorbid neuropathology and AD subtypes such as AD with Lewy bodies, vascular brain injury, hippocampal sclerosis, and TDP-43–immunoreactive tissue elements. This is a useful addition, although other clinically and neuropathologically defined AD subtypes have not been addressed (51). A major turn of direction occurred with the tacit acceptance that AD biologic changes begin long before the clinical presentation of dementia or even cognitive impairment, whereas the 1997 criteria defined AD as requiring the presence of dementia. The present study also provides frequency estimates for various levels of plaque and tangle pathology in subjects with no dementia, giving an overview of this important disease stage. Because there is a need for determining the biologic cause of dementia, the probabilistic approach, which offers certain combinations of plaque and tangle pathology as being of “low,” “intermediate,” and “high” likelihood of causing dementia, has been maintained. The approach to concurrent brain diseases of various types is a useful beginning, but there is still a great need for large multivariable clinicopathologic studies that will ground probabilistic impact estimates on hard data rather than on expert opinion alone. This is especially true for the contribution of vascular brain injury, the documentation of which needs to become much more detailed before useful analyses can begin. It has become more apparent in recent years that dementia in the elderly is most often heterogeneous in origin and, therefore, cannot be understood solely by isolated studies of the “pure” conditions.

Although neuropathologic consensus criteria for AD have changed several times over the last 3 decades, the clinical criteria have remained virtually unchanged since the NINDS-ADRDA-sponsored criteria published in 1984 (6). New clinical diagnostic consensus criteria have now been finalized under the auspices of the NIA and Alzheimer's Association (52), with novel features including recognition of a preclinical stage of AD and the incorporation of imaging and laboratory-based cerebrospinal fluid biomarkers. It is expected that the latter will increase clinical diagnostic accuracy, but verification will require the accumulation of a sufficient number of autopsied subjects to compare with the neuropathologic diagnosis.

In conclusion, between 2005 and 2010, the accuracy rate of the clinical diagnosis of AD at NIA ADCs varied depending on the exact clinical and neuropathologic criteria used. Among subjects with dementia, ADC neurologists were more accurate when they diagnosed AD than when they diagnosed subjects with another dementing disease. Those conducting clinical trials, epidemiologic studies, and governmental healthcare analyses should take diagnostic misclassification into consideration when determining experimental design and data analysis strategies. Whenever possible, efforts should be made to obtain data linked to neuropathologic confirmation of diagnoses. With the maturation of many longitudinal clinicopathologic programs, including those represented by the NIA ADCs, the sample sizes of such autopsy-confirmed cases are now becoming adequate to undertake such analyses. Although new diagnostic biomarkers hold promise for increasing the clinical diagnostic accuracy for AD, it is expected that there will continue to be

overlap in these measures between AD subjects, subjects with non-AD dementias and the elderly with no dementia. The present study may be useful as a baseline against which to assess future improvements.

REFERENCES

- Jack CR Jr, Knopman DS, Jagust WJ, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol* 2010;9:119–28
- Haneuse S, Larson E, Walker R, et al. Neuropathology-based risk scoring for dementia diagnosis in the elderly. *J Alzheimer Dis* 2009;17:875–85
- Knopman DS, DeKosky ST, Cummings JL, et al. Practice parameter: Diagnosis of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* 2001;56:1143–53
- Mayeux R, Saunders AM, Shea S, et al. Utility of the apolipoprotein E genotype in the diagnosis of Alzheimer's disease. Alzheimer's Disease Centers Consortium on Apolipoprotein E and Alzheimer's Disease. *N Engl J Med* 1998;338:506–11
- Gay BE, Taylor KI, Hohl U, et al. The validity of clinical diagnoses of dementia in a group of consecutively autopsied memory clinic patients. *J Nutr Health Aging* 2008;12:132–37
- McKeel DW Jr, Price JL, Miller JP, et al. Neuropathologic criteria for diagnosing Alzheimer disease in persons with pure dementia of Alzheimer type. *J Neuropathol Exp Neurol* 2004;63:1028–37
- Rasmusson DX, Brandt J, Steele C, et al. Accuracy of clinical diagnosis of Alzheimer disease and clinical features of patients with non-Alzheimer disease neuropathology. *Alzheimer Dis Assoc Disord* 1996;10:180–88
- Petrovitch H, White LR, Ross GW, et al. Accuracy of clinical criteria for AD in the Honolulu-Asia Aging Study: A population-based study. *Neurology* 2001;57:226–34
- Gearing M, Mirra SS, Hedreen JC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part X. Neuropathology confirmation of the clinical diagnosis of Alzheimer's disease. *Neurology* 1995;45:461–66
- Galasko D, Hansen LA, Katzman R, et al. Clinical-neuropathological correlations in Alzheimer's disease and related dementias. *Arch Neurol* 1994;51:888–95
- Bowler JV, Munoz DG, Merskey H, et al. Fallacies in the pathological confirmation of the diagnosis of Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 1998;64:18–24
- Kosunen O, Soininen H, Paljarvi L, et al. Diagnostic accuracy of Alzheimer's disease: A neuropathological study. *Acta Neuropathol (Berl)* 1996;91:185–93
- Saunders AM, Hulette O, Welsh-Bohmer KA, et al. Specificity, sensitivity, and predictive value of apolipoprotein-E genotyping for sporadic Alzheimer's disease. *Lancet* 1996;348:90–93
- Villareal DT, Morris JC. The diagnosis of Alzheimer's disease. *J Alzheimer Dis* 1999;1:249–63
- Jellinger KA. Accuracy of clinical criteria for AD in the Honolulu-Asia Aging Study: A population-based study. *Neurology* 2002;58:989–90
- Salmon DP, Thomas RG, Pay MM, et al. Alzheimer's disease can be accurately diagnosed in very mildly impaired individuals. *Neurology* 2002;59:1022–28
- Lopez OL, Becker JT, Klunk W, et al. Research evaluation and diagnosis of probable Alzheimer's disease over the last two decades: I. *Neurology* 2000;55:1854–62
- Victoroff J, Mack WJ, Lyness SA, et al. Multicenter clinicopathological correlation in dementia. *Am J Psychiatry* 1995;152:1476–84
- Klatka LA, Schiffer RB, Powers JM, et al. Incorrect diagnosis of Alzheimer's disease. A clinicopathologic study. *Arch Neurol* 1996;53:35–42
- Lim A, Kukull W, Nochlin D, et al. Cliniconeuropathological correlation of Alzheimer's disease in a community-based case series. *J Am Geriatr Soc* 1999;47:564–69
- Jobst KA, Barnetson LP, Shepstone BJ. Accurate prediction of histologically confirmed Alzheimer's disease and the differential diagnosis of dementia: The use of NINCDS-ADRDA and *DSM-III-R* criteria, SPECT, X-ray CT, and Apo E4 in medial temporal lobe dementias. Oxford Project to Investigate Memory and Aging. *Int Psychogeriatr* 1998;10:271–302
- Nagy Z, Esiri MM, Hindley NJ, et al. Accuracy of clinical operational diagnostic criteria for Alzheimer's disease in relation to different pathological diagnostic protocols. *Dement Geriatr Cogn Disord* 1998;9:219–26
- Tsuang D, Larson EB, Bowen J, et al. The utility of apolipoprotein E genotyping in the diagnosis of Alzheimer disease in a community-based case series. *Arch Neurol* 1999;56:1489–95
- Joachim CL, Morris JH, Selkoe DJ. Clinically diagnosed Alzheimer's disease: Autopsy results in 150 cases. *Ann Neurol* 1988;24:50–56
- Hogervorst E, Barnetson L, Jobst KA, et al. Diagnosing dementia: Interrater reliability assessment and accuracy of the NINCDS/ADRDA criteria versus CERAD histopathological criteria for Alzheimer's disease. *Dement Geriatr Cogn Disord* 2000;11:107–13
- McKhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–44
- Khachaturian ZS. Diagnosis of Alzheimer's disease. *Arch Neurol* 1985;42:1097–105
- Tierney MC, Fisher RH, Lewis AJ, et al. The NINCDS-ADRDA Work Group criteria for the clinical diagnosis of probable Alzheimer's disease: A clinicopathologic study of 57 cases. *Neurology* 1988;38:359–64
- Mirra SS, Heyman A, McKeel D, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. *Neurology* 1991;41:479–86
- The National Institute on Aging, and Reagan Institute Working Group on Diagnostic Criteria for the Neuropathological Assessment of Alzheimer's Disease. Consensus recommendations for the postmortem diagnosis of Alzheimer's disease. *Neurobiol Aging* 1997;18:S1–S2
- Beekly DL, Ramos EM, Lee WW, et al. The National Alzheimer's Coordinating Center (NACC) database: The uniform data set. *Alzheimer Dis Assoc Disord* 2007;21:249–58
- Nelson PT, Frosch MP. Thinking outside the box: Alzheimer-type neuropathology that does not map directly onto current consensus recommendations. *J Neuropathol Exp Neurol* 2010;69:449–54
- Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol (Berl)* 1991;82:239–59
- Nagy Z, Vatter-Bittner B, Braak H, et al. Staging of Alzheimer-type pathology: An interrater-intrarater study. *Dement Geriatr Cogn Disord* 1997;8:248–51
- Mirra SS, Gearing M, McKeel DW Jr, et al. Interlaboratory comparison of neuropathology assessments in Alzheimer's disease: A study of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD). *J Neuropathol Exp Neurol* 1994;53:303–15
- Nagy Z, Yilmazer-Hanke DM, Braak H, et al. Assessment of the pathological stages of Alzheimer's disease in thin paraffin sections: A comparative study. *Dement Geriatr Cogn Disord* 1998;9:140–44
- Halliday G, Ng T, Rodriguez M, et al. Consensus neuropathological diagnosis of common dementia syndromes: Testing and standardizing the use of multiple diagnostic criteria. *Acta Neuropathol (Berl)* 2002;104:72–78
- Newell KL, Hyman BT, Growdon JH, et al. Application of the National Institute on Aging (NIA)-Reagan Institute criteria for the neuropathological diagnosis of Alzheimer disease. *J Neuropathol Exp Neurol* 1999;58:1147–55
- Hyman BT, Phelps CH, Beach TG, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimers Dement* 2012;8:1–13.
- Montine TJ, Phelps CH, Beach TG, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: A practical approach. *Acta Neuropathol* 2012;123:1–11
- Nelson PT, Abner EL, Schmitt FA, et al. Modeling the association between 43 different clinical and pathological variables and the severity of cognitive impairment in a large autopsy cohort of elderly persons. *Brain Pathol* 2010;20:66–79
- Neuropathology Group of the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). Pathological correlates of late-onset dementia in a multicenter, community-based population in England and Wales. *Lancet* 2001;357:169–75

43. Brayne C, Richardson K, Matthews FE, et al. Neuropathological correlates of dementia in over-80-year-old brain donors from the population-based Cambridge city over-75s cohort (CC75C) study. *J Alzheimer Dis* 2009;18:645–58
44. Chen-Plotkin AS, Martinez-Lage M, Sleiman PM, et al. Genetic and clinical features of progranulin-associated frontotemporal lobar degeneration. *Arch Neurol* 2011;68:488–97
45. Larner AJ, Doran M. Genotype-phenotype relationships of presenilin-1 mutations in Alzheimer's disease: An update. *J Alzheimer Dis* 2009;17:259–65
46. Pearson JP, Williams NM, Majounie E, et al. Familial frontotemporal dementia with amyotrophic lateral sclerosis and a shared haplotype on chromosome 9p. *J Neurol* 2011;258:647–55
47. Wszolek ZK, Tsuboi Y, Ghetti B, et al. Frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17). *Orphanet J Rare Dis* 2009;1:30
48. Tsuang D, Simpson KL, Li G, et al. Evaluation of selection bias in an incident-based dementia autopsy case series. *Alzheimer Dis Assoc Disord* 2005;19:67–73
49. Thal DR, Rub U, Orantes M, et al. Phases of A beta-deposition in the human brain and its relevance for the development of AD. *Neurology* 2002;58:1791–800
50. Beach TG, Sue LI, Walker DG, et al. Striatal amyloid plaque density predicts Braak neurofibrillary stage and clinicopathological Alzheimer's disease: Implications for amyloid imaging. *J Alzheimer Dis* 2012;28:869–76
51. Murray ME, Graff-Radford NR, Ross OA, et al. Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: A retrospective study. *Lancet Neurol* 2011;10:785–96
52. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the pre-clinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer Dement* 2011;7:280–92