Neurobiology of Aging, Vol. 18, No. S4, S99–S105, 1997

0197-4580/97 $17.00 + .00

ELSEVIER

PII:S0197-4580(97)00063-8

# Comparison of Neuropathologic Criteria for the Diagnosis of Alzheimer's Disease

J. W. GEDDES,*†[1] T. L. TEKIRIAN,*† N. S. SOULTANIAN,* J. W. ASHFORD,*‡ D. G. DAVIS*§ AND W. R. MARKESBERY*§¶

*Sanders-Brown Center on Aging, and the †Departments of Anatomy and Neurobiology, ‡Psychiatry, §Pathology, and ¶Neurology, College of Medicine, University of Kentucky, Lexington, Kentucky 40536, USA*

GEDDES, J. W., T. L. TEKIRIAN, N. S. SOULTANIAN, J. W. ASHFORD, D. G. DAVIS AND W. R. MARKESBERY. *Comparison of neuropathologic criteria for the diagnosis of Alzheimer's disease.* NEUROBIOL AGING **18**(S4) S99–S105, 1997.—The National Institute on Aging and Reagan Institute (NIA-RI) criteria, and other neuropathologic criteria for Alzheimer's disease (AD), were compared with the clinical diagnosis of dementia in a well defined population of Catholic sisters. The 47-participant subset examined in this study were college educated and lacked complicating conditions such as brain infarcts or diffuse Lewy body disease. Sixteen participants had a clinical diagnosis of dementia. The NIA-RI criteria imply a perfect correlation between neuritic plaque (NP) density and neurofibrillary tangle distribution. However, NP density often did not coincide with tangle distribution. As a result, it was not possible to categorize many of the participants using the NIA-RI guidelines. The 'high likelihood' category of the NIA-RI criteria for AD research settings (neocortical Braak stage and frequent neocortical NP) had relatively high specificity (90% of nondemented participants did not meet this criteria). However, only half of the demented participants were in this category. Neuropathologic criteria requiring the presence of neocortical tangles (rather than neocortical Braak stage) had relatively high sensitivity, accounting for 87–94% of participants with dementia, but also included 32–35% of nondemented participants. Criteria based on neocortical NP or senile plaques had 100% sensitivity, but a majority of nondemented participants also met these criteria. The results support consideration of both tangles and NP for the neuropathologic diagnosis of AD, but indicate that refinement of the NIA-RI criteria is necessary. A possible refinement is suggested for further consideration. © 1997 Elsevier Science Inc.

Alzheimer's disease  Cerebral cortex  Cognition disorders  Epidemiology  Neurofibrillary tangles
Prospective studies  Psychological tests  Diagnosis  Pathology

A definitive diagnosis of Alzheimer's disease (AD) requires both a clinical history of dementia and neuropathologic confirmation at autopsy. However, there are no universally accepted neuropathologic criteria for AD, and the use of different criteria can alter the clinicopathologic correlation (31). Previous recommended neuropathologic criteria for AD include those established by the Neuropathology Panel of a workshop sponsored by the National Institute on Aging and other agencies and organizations, as reported by Khachaturian (18), referred to hereafter as the Khachaturian criteria and the Consortium to Establish a Registry for Alzheimer's Disease (CERAD), as described by Mirra et al. (22). The Khachaturian criteria (18) are based on the number of senile plaques (SP) [SP, includes both neuritic and diffuse plaques] and neurofibrillary tangles (NFT) in neocortex (frontal, temporal, and parietal lobes). Because AD neuropathology becomes increasingly prevalent in nondemented individuals with advancing age, the criteria consider the age of the individual (Table 1). The Khachaturian criteria require SP for all ages and NFT for individuals younger than age 75 (18). CERAD criteria (22) are based on the semiquantitative assessment of neocortical neuritic plaques (NP), and it also uses specific criteria for different age groups.

The reliance on SP density for the neuropathologic diagnosis of AD is consistent with studies in which the number of SP in various neocortical regions correlated with the severity of dementia (6,28). However, the density of neocortical SP in nondemented individuals can be sufficient to meet the Khachaturian criteria (10,16). Other reports indicate that the presence and severity of neocortical NFT may provide a better agreement with a clinical diagnosis of probable AD (2,5,11,12,17,21,24). Tierney and colleagues (31) examined three sets of neuropathologic criteria for AD based on the presence of one or more NFT and NP in hippocampus (A1 criteria), neocortex (A3 criteria), or both hippocampus and neocortex (A2 criteria). Their results indicated that the A1 criteria provided the greatest agreement with a clinical diagnosis of probable AD (31). In a previous report from the Nun Study (29), neuropathologic diagnosis of AD was based on the Khachaturian criteria for individuals age 75 or older and the presence of neocortical NFT.

Braak and Braak (8) described six stages of AD neuropathology based on the pattern of neurofibrillary changes (NFT, neuropil threads, and plaque dystrophic neurites), but they did not compare neuropathology with clinical symptoms. However, they did speculate that the entorhinal stages (I/II) represent clinically silent periods of the disease, the limbic stages (III/IV) correspond to

[1] To whom requests for reprints should be addressed. James W. Geddes, Ph.D., Sanders-Brown Center on Aging, Room 209, University of Kentucky, 800 S. Limestone, Lexington, KY 40536-0230.

TABLE 1
NEUROPATHOLOGIC CRITERIA OF ALZHEIMER'S DISEASE

| Neuropathologic Criteria | Age/Stage | SP | NP | NFT | | |
|---|---|---|---|---|---|---|
| | | | | NeoCtx | Hippo | EntCtx |
| Khachaturian (18) | < 50 | > 2–5/mm$^2$ | | > 2–5/mm$^2$ | | |
| | 50–65 | > 8 | | some | | |
| | 66–75 | > 10 | | some | | |
| | > 75 | > 15 | | | | |
| CERAD for age > 75 (22) | Uncertain | | Sparse | | | |
| | Suggest | | Moderate | | | |
| | Indicate | | Frequent | | | |
| Braak (8) | I-II (Entorhinal) | | | Sparse | Sparse | Present |
| | III-IV (Limbic) | | | Sparse | Occasional | Abundant |
| | V-VI (Neocortical) | | | Abundant | Abundant | Abundant |
| NIA-RI AD Research (25) | Low Likelihood | | Sparse | Sparse | Sparse | Present |
| | Intermediate | | Moderate | Sparse | Occasional | Abundant |
| | High Likelihood | | Frequent | Abundant | Abundant | Abundant |
| NIA-RI Routine* (25) | Low Likelihood | | Sparse | Absent | Absent | Present |
| | Intermediate | | Moderate | Absent | Present | Present |
| | High Likelihood | | Frequent | Present | Present | Present |
| Tierney A1 (31) | | | Present in Hippo | | Present | |
| Tierney A2 | | | Hippo and NeoCtx | Present | Present | |
| Tierney A3 | | | Present in NeoCtx | Present | | |
| Nun Study (29) | | > 15/mm$^2$ | | Present | | |
| Neocortical NFT | | | | Present | | |

Parietal cortex is shown as a representative neocortical region.
* Our interpretation of the NIA-RI routine criteria, see text for details.
Abbreviations: CERAD, Consortium to Establish a Registry for Alzheimer's Disease; EntCtx, Entorhinal Cortex; Hippo, Hippocampus (including subiculum); NIA-RI, National Institute on Aging and Reagan Institute; NeoCtx, Neocortex; NFT, neurofibrillary tangle; NP, neuritic plaque; SP, senile plaque.

clinically incipient AD, and the neocortical stages (V/VI) represent fully developed AD. Subsequent comparisons of Braak stage with cognitive function have revealed an increasing prevalence of dementia with advancing Braak stage, but these comparisons have been made in studies involving relatively few participants (3,20).

Recently, a panel of neuropathologists under the aegis of the National Institute on Aging and the Ronald and Nancy Reagan Research Institute of the Alzheimer's Association proposed neuropathologic criteria for AD based on both NFT and NP (25). For the routine neuropathologic diagnosis of AD, these criteria (referred to here as the NIA-RI criteria) recommend the CERAD semiquantitative assessment of NP and the examination of the hippocampal formation and neocortex for the presence of NFT. In AD research settings, the NIA-RI criteria recommend use of Braak staging to establish the extent of neurofibrillary lesions. The Braak stage and CERAD score are used to provide an estimate of the likelihood that the AD neuropathological changes underlie dementia (Table 1).

The purpose of this study was to compare neuropathologic criteria for AD with the clinical diagnosis of dementia in a well defined population of Catholic sisters (the Nun Study), which includes demented and nondemented participants who have been followed prospectively using a standard neuropsychological test battery (29,30). The 678 participants in the Nun Study had the same reproductive and marital histories, had similar social activities and support, did not smoke, did not drink excessive amounts of alcoholic beverages, had similar income and socioeconomic status, and had similar access to medical care services. The 47-participant subset of the Nun Study, examined in the present study, had similar education (all had a bachelor's degree or higher) and occupations (all were teachers). Consequently, factors that confound many epidemiological studies are minimized or eliminated in this group of participants.

## METHODS

### *Study Population*

Participants in the Nun Study have been described in detail previously (29,30). Participants agreed to annual assessments, and all agreed to brain donation at death. At the first assessment in 1991–93, the 678 participants in the Nun Study were 75–102 years old (mean = 83, SD = 5). Of the first 130 who died, neuropathologic evaluations were done on 118. Brain tissue was not available for 12 participants primarily due to difficulties in retrieving the brains from distant locations. To minimize confounding by attained education, the subset for this study was restricted to 90 participants who had earned at least a bachelor's degree. Because brain infarcts were associated with diminished cognitive performance in the presence of AD pathology in the participants (29), we also excluded individuals with lacunar or larger brain infarcts. This subset of 54 participants was further reduced to 47 by excluding participants with other conditions including diffuse Lewy body disease, meningioma, glioblastoma multiforme, hydrocephalus, a history of psychiatric illness, and one individual with cerebral contusions. Subsequent investigation revealed five participants with cortical argyrophilic grains. These five participants were not demented, and they were not excluded from analysis.

*Cognitive function and dementia*

Cognitive function was assessed annually by a battery of seven tests compiled by CERAD (23). The CERAD test battery includes the Mini-Mental State Exam (MMSE), Modified Boston Naming Test, Verbal Fluency, Word List Memory, Word List Recall, Word List Recognition, and Constructional Praxis. Cognitive data used in the analysis were derived from each participant's last examination. Cut scores used to identify impairments in cognition represented scores that were less than the fifth percentile for the normative data described by the CERAD group (32). Social and daily functioning was assessed by performance-based tests (19,26).

Individuals considered to be demented in this study had each of the following conditions: impairment in a memory test, i.e., $<4$ on the Delayed Word Recall; impairment in at least one other area of cognition, i.e., $<11$ on the Verbal Fluency, $<13$ on the Boston Naming, or $<8$ on the Constructional Praxis; impairment in social or daily function, i.e., inability to use a telephone, handle money or dress oneself, and a decline in function from a previous level. The criteria are similar to the DSM-IV criteria for dementia (1).

*Neuropathologic evaluation*

At autopsy, the brain was placed in 10% buffered formalin and fixed for a minimum of two weeks. Gross examination of the brain was performed by a neuropathologist who was blinded to the participant's cognitive test scores. Tissue blocks from multiple neocortical regions, hippocampus, entorhinal cortex, amygdala, basal ganglia, midbrain, pons, medulla, and cerebellum were paraffin embedded. To assess Braak stage, sections (10 μm) containing entorhinal cortex and anterior hippocampus were stained with the Gallyas method for neurofibrillary pathology (9,14). In some cases, sections were also stained with the PHF-1 antibody against phosphorylated tau protein (15) that produced similar results (not shown). For staining of NP and NFT, the modified Bielschowsky method was used on sections from neocortical regions, including the middle frontal gyrus (Brodmann area 9), inferior parietal lobule (areas 39/40), middle temporal gyrus (area 21), and occipital cortex (area 18).

NP scores were determined by semiquantitative estimation of neocortical NP density using the CERAD criteria for age $>75$ (22). In addition to the semiquantitative estimates, the number of NFT were counted in the 10 most severely involved microscopic fields; SP and NP were counted in the 10 most severely involved SP fields. NP were counted per 2.35 $mm^2$ and NFT counted per 0.586 $mm^2$ field. Braak stage and NP scores were assessed without knowledge of diagnosis or cognitive test scores.

## RESULTS

The 47 participants in this study ranged in age from 76–98 (mean 85 $\pm$ 5 SD) and were assessed between 7 and 760 days before death (351 $\pm$ 191). Sixteen (34%) of the individuals were classified as demented. The neuropathologic criteria examined are outlined in Table 1.

*CERAD NP Score and Braak stage*

Thirty participants had a frequent CERAD NP score, 6 were moderate, and the remaining 11 participants did not have neocortical NP. None of the participants had a sparse CERAD NP score (Table 2). Eight of the 47 participants were classified as Braak Stage I, 11 as Stage II, 8 as Stage III, 8 as Stage IV, 10 as Stage V, and 2 as Stage VI. The six Braak Stages were pooled into three groups [Entorhinal (I/II), Limbic (III/IV) and Neocortical Stages (V/VI)] corresponding to the classification used by the NIA-RI criteria (Table 2).

*NIA-RI Research Setting Criteria*

The NIA-RI criteria for AD research settings (25) utilize both CERAD NP scores and Braak stages. The recommendations imply that CERAD NP scores have a perfect correlation with Braak stages (e.g., individuals with frequent CERAD NP scores are also neocortical Braak stage, see Table 1). Using these recommendations, 11 participants were in the 'high likelihood', 2 in the 'intermediate likelihood', and 9 in the 'low likelihood' categories. The remaining 25 participants could not be classified using the NIA-RI criteria. For example, of the 30 participants with a CERAD frequent NP score, 19 were not neocortical Braak stage (10 were limbic and 9 were entorhinal Braak stage), therefore, they did not meet the high, intermediate, or low likelihood criteria (Table 1).

As an initial attempt to accommodate participants excluded from the NIA-RI neuropathologic criteria for AD research settings, we modified the intermediate likelihood category to include participants excluded by the high and low likelihood categories (Table 1). This modified intermediate category included participants with entorhinal ($n = 10$), limbic ($n = 16$), and neocortical ($n = 1$) Braak stages, and moderate ($n = 8$) and frequent ($n = 19$) CERAD NP scores. Of these 27 participants, 14 had some neocortical NFT. CA1 NFT counts for the 13 participants without neocortical NFT were 2.9 $\pm$ 4.9 (mean $\pm$ SD). For the 14 participants with neocortical NFT, CA1 NFT counts were 13.5 $\pm$ 13.2 and parietal cortex NFT counts were 1.7 $\pm$ 3.0. Eight of these 27 participants were demented, and seven of these demented participants had neocortical NFT.

*NIA-RI Routine Settings Criteria*

The routine NIA-RI criteria (for pathologists and neuropathologists outside of AD research settings) involve CERAD NP scores and examination of the hippocampal formation and neocortex for the presence of NFT (25). Our interpretation of the NIA-RI report (25) indicates that the high likelihood category requires a frequent CERAD NP score and the presence of NFT in neocortex (Table 1). The intermediate likelihood category requires limbic (hippocampal) NFT and a moderate CERAD NP score. The low likelihood category requires a sparse CERAD NP score and NFT in a more limited distribution or severity. We interpret this as requiring NFT in entorhinal cortex, although the guidelines for the routine criteria do not include examination of entorhinal cortex for NFT (Table 1).

As with the NIA-RI criteria for AD research settings, the NIA-RI routine criteria imply a strong correlation between NFT distribution and CERAD NP score (Table 1). Using our interpretation of the guidelines, 13 participants were in the high likelihood category, 2 in the intermediate likelihood and 1 participant was in the low likelihood category (Table 2). The remaining 21 participants did not meet criteria recommended for the 3 categories.

Additional criteria examined are outlined in Table 1. The number of participants that meet each of the neuropathologic criteria for AD are indicated in Table 2, along with mean age, MMSE score, and CA1 and parietal cortex NFT counts. CA1 NFT counts were not available for one participant with dementia. Participants which met the Tierney A3 criteria also met the Nun Study criteria (Table 2).

*Specificity, Sensitivity, and Overall Accuracy of Selected Neuropathologic Criteria*

Sensitivity (percentage of demented participants correctly identified by neuropathologic criteria) ranged from 50–100% (Table

TABLE 2

AGE, MMSE, AND TANGLE COUNTS FOR PARTICIPANTS IN EACH NEUROPATHOLOGIC CRITERIA

| Criteria | Stage | N | Age | MMSE | CA1 NFT* | Parietal NFT |
|---|---|---|---|---|---|---|
| Khachaturian | Do Not Meet Criteria | 11 | 81.9 ± 5.4 | 27.1 ± 2.9 | 2.7 ± 5.4 | 0.0 ± 0.0 |
| | Meet Criteria | 36 | 85.9 ± 4.3 | 19.7 ± 10.4 | 11.3 ± 13.0 | 2.6 ± 4.3 |
| CERAD | Uncertain | 11 | 79.6 ± 2.6 | 27.9 ± 1.7 | 2.2 ± 5.0 | 0.0 ± 0.0 |
| | Suggest | 6 | 88.9 ± 2.8 | 20.2 ± 10.6 | 13.8 ± 18.5 | 1.9 ± 4.1 |
| | Indicate | 30 | 86.2 ± 4.2 | 19.3 ± 10.3 | 11.0 ± 11.9 | 2.8 ± 4.4 |
| Braak | Entorhinal | 19 | 82.9 ± 4.3 | 26.9 ± 4.1 | 0.8 ± 1.2 | 0.0 ± 0.0 |
| | Limbic | 16 | 87.1 ± 5.1 | 21.7 ± 8.2 | 10.8 ± 9.5 | 0.8 ± 1.7 |
| | Neocortical | 12 | 85.5 ± 4.0 | 12.4 ± 11.4 | 21.6 ± 14.9 | 6.7 ± 5.1 |
| NIA-RI AD Research | Low Likelihood | 9 | 79.4 ± 1.9 | 28.3 ± 1.6 | 0.7 ± 0.9 | 0.0 ± 0.0 |
| | Intermediate | 2 | 91.8 ± 3.1 | 22.0 ± 2.8 | 5.0 ± 5.7 | 0.0 ± 0.0 |
| | High Likelihood | 11 | 85.2 ± 4.1 | 13.5 ± 11.3 | 19.3 ± 13.5 | 6.4 ± 5.3 |
| | Unclassified | 25 | 86.4 ± 4.3 | 22.4 ± 8.8 | 8.6 ± 11.7 | 0.9 ± 2.4 |
| NIA-RI Routine | Low Likelihood | 1 | 79.6 | 28.0 | 0.0 | 0.0 |
| | Intermediate | 2 | 91.8 ± 3.1 | 22.0 ± 2.8 | 5.0 ± 5.7 | 0.0 ± 0.0 |
| | High Likelihood | 23 | 86.5 ± 4.3 | 17.4 ± 10.6 | 14.1 ± 12.1 | 3.6 ± 4.7 |
| | Unclassified | 21 | 82.9 ± 4.4 | 25.5 ± 7.2 | 5.1 ± 11.3 | 0.5 ± 2.2 |
| Tierney A1 | Do Not Meet Criteria | 19 | 82.4 ± 4.7 | 27.2 ± 2.7 | 1.9 ± 4.2 | 0.0 ± 0.0 |
| | Meet Criteria* | 27 | 86.9 ± 4.1 | 17.9 ± 10.8 | 14.5 ± 13.3 | 3.4 ± 4.7 |
| Tierney A2 | Do Not Meet Criteria | 22 | 83.1 ± 4.9 | 26.7 ± 4.1 | 2.0 ± 3.9 | 0.0 ± 0.0 |
| | Meet Criteria* | 24 | 86.8 ± 4.1 | 17.1 ± 10.9 | 15.9 ± 13.4 | 3.9 ± 4.8 |
| Tierney A3 | Do Not Meet Criteria | 22 | 83.1 ± 4.9 | 26.7 ± 4.1 | 2.0 ± 3.9 | 0.0 ± 0.0 |
| | Meet Criteria | 25 | 86.7 ± 4.1 | 16.8 ± 10.8 | 15.9 ± 13.4 | 3.8 ± 4.7 |
| Nun Study | Do Not Meet Criteria | 22 | 83.1 ± 4.9 | 26.7 ± 4.1 | 2.0 ± 3.9 | 0.0 ± 0.0 |
| | Meet Criteria | 25 | 86.7 ± 4.1 | 16.8 ± 10.8 | 15.9 ± 13.4 | 3.8 ± 4.7 |
| Neocortical NFT | Absent | 21 | 83.3 ± 4.9 | 26.6 ± 4.1 | 2.1 ± 4.0 | 0.0 ± 0.0 |
| | Present | 26 | 86.3 ± 4.4 | 17.3 ± 10.9 | 15.3 ± 13.5 | 3.6 ± 4.7 |

Values for Mini-Mental Status Exam (MMSE), age and neurofibrillary tangle (NFT) counts are the mean ± SD. Age represents the age of participant at the last cognitive assessment. NFT means are for the 10 most severely involved microscopic (0.586 $mm^2$) fields.

* CA1 NFT counts were not available for one participant with dementia; this participant was excluded from analysis using the Tierney A1 and A2 criteria.

3). Nine of 16 (56%) of the demented participants had neocortical (V/VI) Braak stage pathology, including 8 (50%) that were in the high likelihood category of the NIA-RI criteria for AD research settings. The high likelihood category of the routine NIA-RI criteria included 13 (81%) of the 16 demented participants. Each of the 16 participants with dementia had neocortical NP, and they also had SP sufficient to meet the Khachaturian criteria. All but one participant with dementia had neocortical NFT. Sensitivity for the Tierney A3 criteria (NP and NFT in neocortex) was similar to the Tierney A1 criteria (NP and NFT in hippocampus), however, the pattern of AD pathology in the two regions was distinct. Of 22 participants who did not meet the Tierney A3 criteria, 11 had NP but not NFT in neocortex, 1 had NFT but not NP, and the remaining 10 had neither NFT nor NP. Of 19 participants who did not meet the Tierney A1 criteria, 13 had NFT but not NP in hippocampus; 1 had NP but not NFT.

Specificity (percentage of nondemented participants correctly identified by the neuropathologic criteria) ranged from 35–90% (Table 3). Twenty-eight of 31 (90%) nondemented participants did not have neocortical Braak stage pathology. Specificity was also 90% for the high-likelihood category of the NIA-RI criteria for AD research settings, which requires neocortical Braak stage. Specificity of criteria that required the presence of neocortical NFT (e.g., Tierney A2, Tierney A3, and the Nun Study), rather than neocortical Braak stage, was 65–68% (Table 3). Specificity was lower (58%) for the Tierney A1 criteria, which requires hippocampal NP and NFT (Table 3). For criteria that required neocortical NP (CERAD) or SP (Khachaturian) but not NFT, specificity was 35–48%.

Criteria based on eleven (35%) of the nondemented participants lacked neocortical NP (CERAD uncertain category). Each of these participants also lacked neocortical NFT. Neocortical SP density was below that required by the Khachaturian criteria in 35% of nondemented participants.

The three nondemented participants with neocortical Braak pathology also had frequent CERAD NP scores thereby meeting the high likelihood criteria of the NIA-RI criteria. It was surprising that individuals with this magnitude of pathology were not demented. However, two of the three were at or below the cut scores on at least five of the seven cognitive tests in the CERAD battery. They were not classified as demented because they were not below cut scores on tests for memory impairment, or they did not display impaired social and daily functioning. The other nondemented participant classified with neocortical Braak stage pathology scored well on each test of cognitive function (i.e., MMSE = 28). This was an individual with numerous NFT in occipital cortex, but relatively mild pathology in hippocampus and entorhinal cortex.

In this study, overall accuracy (percentage of participants correctly identified as demented or nondemented by the neuropathologic criteria) was similar (72–79%) for Braak staging, NIA-RI, Tierney (A1, A2 and A3), Nun Study, and neocortical NFT criteria (Table 3). As shown in Table 4, the overall ability of neuropathologic criteria to correctly classify demented and non-

TABLE 3

SENSITIVITY, SPECIFICITY, AND OVERALL ACCURACY OF THE VARIOUS NEUROPATHOLOGIC CRITERIA

| Criteria | Stage | Dementia | | Percent of Participants | | Sensitivity (95% CI) | Specificity (95% CI) | Overall Accuracy (95% CI) |
|---|---|---|---|---|---|---|---|---|
| | | Yes | No | With Dementia | Without Dementia | | | |
| Khachaturian | Do Not Meet Criteria | 0 | 11 | 0 | 35 | | | |
| | Meet Criteria | 16 | 20 | 100 | 65 | 100 (79–100) | 35 (19–55) | 57 (42–72) |
| CERAD | Uncertain | 0 | 11 | 0 | 35 | | | |
| | Suggest | 2 | 4 | 13 | 13 | | | |
| | Indicate | 14 | 16 | 88 | 52 | 88 (62–98) | 48 (30–67) | 62 (46–75) |
| Braak | Entorhinal | 2 | 17 | 13 | 55 | | | |
| | Limbic | 5 | 11 | 31 | 35 | | | |
| | Neocortical | 9 | 3 | 56 | 10 | 56 (30–80) | 90 (74–98) | 79 (64–89) |
| NIA-RI AD Research | Low Likelihood | 0 | 9 | 0 | 29 | | | |
| | Intermediate | 0 | 2 | 50 | 6 | | | |
| | *High Likelihood | 8 | 3 | 50 | 10 | 50 (25–75) | 90 (74–98) | 77 (62–88) |
| | Unclassified | 8 | 17 | 50 | 55 | | | |
| NIA-RI Routine | Low Likelihood | 0 | 1 | 0 | 3 | | | |
| | Intermediate | 0 | 2 | 0 | 6 | | | |
| | *High Likelihood | 13 | 10 | 81 | 32 | 81 (54–96) | 68 (49–83) | 72 (57–84) |
| | Unclassified | 3 | 18 | 10 | 58 | | | |
| Tierney A1 | Do Not Meet Criteria | 1 | 18 | 7 | 58 | | | |
| | Meet Criteria | 14 | 13 | 93 | 42 | 93 (68–100) | 58 (39–75) | 70 (54–82) |
| Tierney A2 | Do Not Meet Criteria | 1 | 21 | 7 | 68 | | | |
| | Meet Criteria | 14 | 10 | 93 | 32 | 93 (68–100) | 68 (49–83) | 76 (60–87) |
| Tierney A3 | Do Not Meet Criteria | 1 | 21 | 6 | 68 | | | |
| | Meet Criteria | 15 | 10 | 94 | 32 | 94 (70–100) | 68 (49–83) | 77 (62–88) |
| Nun Study | Do Not Meet Criteria | 1 | 21 | 6 | 68 | | | |
| | Meet Criteria | 15 | 10 | 94 | 32 | 94 (70–100) | 68 (49–83) | 77 (62–88) |
| Neocortical NFT | Do Not Meet Criteria | 1 | 20 | 6 | 65 | | | |
| | Meet Criteria | 15 | 11 | 94 | 35 | 94 (70–100) | 65 (45–81) | 74 (60–86) |

* Calculation of sensitivity and specificity for the NIA-RI 'high-likelihood' criteria assumes that unclassified participants are in the 'low-likelihood' or 'intermediate likelihood' groups.

Abbreviations: CI, exact confidence intervals. Other abbreviations are as indicated in Table 1.

demented participants depends on the prevalence of dementia in the sample. For example, when one-half of a hypothetical sample of 100 participants is demented, the overall accuracy of the NIA-RI criteria is 70%. That is, 90% of the 50 nondemented individuals in the hypothetical sample were accurately classified as nondemented, and 50% of the 50 demented participants in the hypothetical sample were accurately classified as demented. If the sample population consists primarily of individuals with dementia, plaque-based criteria display the highest accuracy. As the percentage of nondemented individuals in the sample population increases, criteria that require abundant neocortical NFT exhibit the highest accuracy.

TABLE 4

OVERALL ACCURACY OF VARIOUS NEUROPATHOLOGIC CRITERIA FOR ALZHEIMER'S DISEASE BY DIFFERENT PREVALENCE OF DEMENTIA

| Neuropathologic Criteria | Category | Overall Accuracy by Prevalence of Dementia | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
| Khachaturian | Meet Criteria | 35 | 42 | 48 | 55 | 61 | 68 | 74 | 81 | 87 | 94 | 100 |
| CERAD | Indicate | 48 | 52 | 56 | 60 | 64 | 68 | 72 | 76 | 80 | 84 | 88 |
| Braak | Neocortical | 90 | 87 | 83 | 80 | 76 | 73 | 70 | 66 | 63 | 59 | 56 |
| NIA-RI AD Research | High Likelihood | 90 | 86 | 82 | 78 | 74 | 70 | 66 | 62 | 58 | 54 | 50 |
| NIA-RI Routine | High Likelihood | 68 | 69 | 71 | 72 | 73 | 75 | 76 | 77 | 78 | 80 | 81 |
| Tierney A1 | Meet Criteria | 58 | 62 | 65 | 69 | 72 | 76 | 79 | 83 | 86 | 90 | 93 |
| Tierney A2 | Meet Criteria | 68 | 71 | 73 | 76 | 78 | 81 | 83 | 86 | 88 | 91 | 93 |
| Tierney A3 | Meet Criteria | 68 | 70 | 73 | 76 | 78 | 81 | 83 | 86 | 89 | 91 | 94 |
| Nun Study | Meet Criteria | 68 | 70 | 73 | 76 | 78 | 81 | 83 | 86 | 89 | 91 | 94 |
| Neocortical NFT | Meet Criteria | 65 | 68 | 71 | 74 | 77 | 80 | 82 | 85 | 88 | 91 | 94 |

## DISCUSSION

The NIA-RI report (25) recommends that the proposed criteria be investigated using well characterized populations, and anticipates future refinement of the criteria. The results of the present study indicate difficulties in applying the NIA-RI criteria and suggest one possible way to refine the criteria.

The proposed NIA-RI neuropathologic criteria for AD (25) includes two different sets of recommendations, one for AD research settings and the other for diagnosis of AD by pathologists and neuropathologists outside of AD research settings. The recommendations for AD research settings utilize CERAD NP scores and Braak staging of neurofibrillary pathology. Outside of AD research settings, the criteria utilize CERAD NP scores and the presence of NFT in hippocampus and/or neocortex. These criteria are clearly different because NFT may be present in neocortex in entorhinal (I/II) and limbic (III/IV) Braak stages, but they are abundant in neocortical (V/VI) Braak stage (8). Both sets of NIA-RI recommendations imply a perfect correlation between neocortical NP density and NFT distribution. However, there is often a mismatch between NP and NFT pathology (8,13). This was evident in the present study; NP density often did not coincide with NFT distribution. As a result, many participants did not meet the requirements for the low, intermediate, and high likelihood categories of the NIA-RI criteria. Although the NIA-RI report (25) indicates that AD may occur in individuals who have combinations of NP and NFT other than those specified, it does not indicate how to categorize such individuals.

Of the examined neuropathologic criteria, the greatest overall accuracy was achieved by criteria that required neocortical NFT. The accuracy of the NIA-RI high likelihood group and neocortical Braak stage resulted primarily from relatively high specificity, or an ability to identify correctly most nondemented participants. In contrast, sensitivity of these criteria was relatively low. The overall accuracy of neuropathologic criteria requiring the presence (rather than the abundance) of neocortical NFT was similar to that of the neocortical Braak stage and high likelihood NIA-RI categories, but in this case, it resulted from relatively high sensitivity (correct identification of demented participants). All but one of the demented participants in this study had neocortical NFT. The preferred neuropathologic criteria may depend on the prevalence of dementia within the study population. If the sample has a high prevalence of dementia, sensitivity is usually most important. If the sample has a low prevalence of dementia, specificity is usually most important.

Taken together, the results demonstrate an association between neocortical NFT and dementia. The results further indicate that the prevalence of dementia increases as neocortical NFT become more abundant (e.g., neocortical Braak stage). This is consistent with results of several previous studies (2,4,5,7,10,12,17,21,24). Although the presence of neocortical NFT is associated with dementia, it is unlikely that occasional neocortical NFT are sufficient to trigger dementia. The appearance of NFT in neocortex typically reflects the spread of neurofibrillary pathology and the worsening of NFT density in hippocampus and entorhinal cortex (27). This is consistent with results observed in the present study.

The results do not exclude a role for NP in determining if the AD pathology may contribute to dementia. All but one participant with neocortical NFT had moderate-to-severe CERAD NP scores. In this exception, a single NFT was observed in frontal cortex and the participant was not demented. At the opposite end of the neuropathologic spectrum, participants without neocortical NP also lacked neocortical NFT and were not demented. Moreover, the presence of NP helps to rule out other conditions such as progressive supranuclear palsy.

The participants in this study were female, elderly ($>$ 75 years of age), and highly educated. Although it may be difficult to generalize from this unique population of Catholic sisters, the results indicate that refinement of the NIA-RI criteria is necessary. We suggest a possible refinement of the NIA-RI criteria for further consideration. The suggested refinement incorporates the guiding principles of the NIA-RI criteria, relies on both NP and NFT for the diagnosis of AD, and includes aspects of several neuropathologic criteria examined in the present study. It could be used within AD research settings and by pathologists and neuropathologists outside of AD research settings. The suggested refinement is outlined below.

It is "highly likely" that dementia results from the AD pathology if NFT are abundant in neocortex and hippocampus (e.g., neocortical Braak stage) and NP are present in neocortex. In the present study, all the participants with neocortical Braak stage had neocortical NP. It is "likely" that dementia results from the AD pathology if NFT and NP present in neocortex, and if NFT are also present in hippocampus. This is similar to the Tierney A3 criteria (31) with the additional requirement of hippocampal NFT. It is "less likely" that dementia results from AD pathology in individuals with NP or NFT in neocortex. Finally, it is "unlikely" that dementia results from AD pathology if both NP and NFT are absent from neocortex. This is illustrated by results for the CERAD uncertain category in the present study, because all participants without neocortical NP also lacked neocortical NFT.

## ACKNOWLEDGEMENTS

We thank Ela Patel for excellent technical assistance and Dr. Huaichen Liu for quantitative assessment of AD pathology. This study would not have been possible without the support of the members, leaders, and health care providers of the School Sisters of Notre Dame. This study was funded by grants R35 AG08974 (J. W. G.), P50AG05144 (W. R. M.) and F31MH11650 (T. L. T.), and a grant from the Abercrombie Foundation.

## REFERENCES

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed. Washington, DC: American Psychiatric Association; 1994.
2. Arriagada, P. V.; Growdon, J. H.; Hedley-Whyte, E. T.; Hyman, B. T. Neurofibrillary tangles but not senile plaques parallel duration and severity of Alzheimer's disease. Neurology 42:631–639; 1992.
3. Bancher, C.; Braak, H.; Fischer, P.; Jellinger, K. A. Neuropathological staging of Alzheimer lesions and intellectual status in Alzheimer's and Parkinson's disease patients. Neurosci. Lett. 162:179–182; 1993.
4. Bancher, C.; Jellinger, K.; Lassmann, H.; Fischer, P.; Leblhuber, F. Correlations between mental state and quantitative neuropathology in the Vienna Longitudinal Study on Dementia. Eur. Arch. Psychiatry Clin. Neurosci. 246:137–146; 1996.
5. Bierer, L. M.; Hof, P. R.; Purohit, D. P.; Carlin, L.; Schmeidler, J.; Davis, K. L.; Perl, D. P. Neocortical neurofibrillary tangles correlate with dementia severity in Alzheimer's disease. Arch. Neurol. 52:81–88; 1995.
6. Blessed, G.; Tomlinson, B. E.; Roth, M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. Br. J. Psychiatry 114:797–811; 1968.
7. Bouras, C.; Hof, P. R.; Giannakopoulos, P.; Michel, J. P.; Morrison, J. H. Regional distribution of neurofibrillary tangles and senile plaques in the cerebral cortex of elderly patients: A quantitative evaluation of a one-year autopsy population from a geriatric hospital. Cereb. Cortex 4:138–150; 1994.

8. Braak, H.; Braak, E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. (Berl.) 82:239–259; 1991.
9. Braak, H.; Braak, E.; Ohm, T.; Bohl, J. Silver impregnation of Alzheimer's neurofibrillary changes counterstained for basophilic material and lipofuscin pigment. Stain Technol. 63:197–200; 1988.
10. Crystal, H.; Dickson, D.; Fuld, P.; Masur, D.; Scott, R.; Mehler, M.; Masdeu, J.; Kawas, C.; Aronson, M.; Wolfson, L. Clinico-pathologic studies in dementia: Nondemented subjects with pathologically confirmed Alzheimer's disease. Neurology 38:1682–1687; 1988.
11. Dayan, A. D. Quantitative histological studies on the aged human brain. I. Senile plaques and neurofibrillary tangles in "normal" patients. Acta Neuropathol. (Berl.) 16:85–94; 1970.
12. Dayan, A. D. Quantitative histological studies on the aged human brain. II. Senile plaques and neurofibrillary tangles in senile dementia (with an appendix on their occurrence in cases of carcinoma). Acta Neuropathol. (Berl.) 16:95–102; 1970.
13. Dickson, D. W. Mismatch between plaques and tangles in staging Alzheimer pathology. Neurobiol. Aging 16:283–284; 1995.
14. Gallyas, F. Silver staining of Alzheimer's neurofibrillary changes by means of physical development. Acta Morphol. Acad. Sci. Hung. 19:1–8; 1971.
15. Greenberg, S. G.; Davies, P. A preparation of Alzheimer paired helical filaments that displays distinct tau proteins by polyacrylamide gel electrophoresis. Proc. Natl. Acad. Sci. U.S.A. 87:5827–5831; 1990.
16. Katzman, R.; Terry, R.; DeTeresa, R.; Brown, T.; Davies, P.; Fuld, P.; Renbing, X.; Peck, A. Clinical, pathological, and neurochemical changes in dementia: A subgroup with preserved mental status and numerous neocortical plaques. Ann. Neurol. 23:138–44; 1988.
17. Kazee, A. M.; Eskin, T. A.; Lapham, L. W.; Gabriel, K. R.; McDaniel, K. D.; Hamill, R. W. Clinicopathologic correlates in Alzheimer disease: Assessment of clinical and pathologic diagnostic criteria. Alzheimer Dis. Assoc. Disord. 7:152–164; 1993.
18. Khachaturian, Z. S. Diagnosis of Alzheimer's disease. Arch. Neurol. 42:1097–1105; 1985.
19. Kuriansky, J.; Gurland, B. The performance test of activities of daily living. Int. J. Aging Human Dev. 7:343–352; 1976.
20. Masliah, E. Natural evolution of the neurodegenerative alterations in Alzheimer's disease. Neurobiol. Aging 16:280–282; 1995.
21. McKee, A. C.; Kosik, K. S.; Kowall, N. W. Neuritic pathology and dementia in Alzheimer's disease. Ann. Neurol. 30:156–165; 1991.
22. Mirra, S. S.; Heyman, A.; McKeel, D.; Sumi, S. M.; Crain, B. J.; Brownlee, L. M.; Vogel, F. S.; Hughes, J. P.; van Belle, G.; Berg, L. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. Neurology 41:479–486; 1991.
23. Morris, J. C.; Heyman, A.; Mohs, R. C.; Hughes, J. P.; van Belle, G.; Fillenbaum, G.; Mellits, E. D.; Clark, C. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 39:1159–1165; 1989.
24. Nagy, Z.; Esiri, M. M.; Jobst, K. A.; Morris, J. H.; King, E. M.; McDonald, B.; Litchfield, S.; Smith, A.; Barnetson, L.; Smith, A. D. Relative roles of plaques and tangles in the dementia of Alzheimer's disease: Correlations using three sets of neuropathological criteria. Dementia 6:21–31; 1995.
25. 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 18(Suppl 1)S1–S3.
26. Potvin, A. R.; Tourtellotte, W. W.; Dailey, J. S.; Alberts, J. W.; Walker, J. E.; Pew, R. W.; Henderson, W. G.; Snyder, D. N. Simulated activities of daily living examination. Arch. Phys. Med. Rehabil. 53:476–486; 1972.
27. Price, J. L.; Davis, P. B.; Morris, J. C.; White, D. L. The distribution of tangles, plaques and related immunohistochemical markers in healthy aging and Alzheimer's disease. Neurobiol. Aging 12:295–312; 1991.
28. Roth, M.; Tomlinson, B. E.; Blessed, G. Correlation between scores for dementia and counts of 'senile plaques' in cerebral grey matter of elderly subjects. Nature 209:109–110; 1966.
29. Snowdon, D. A.; Greiner, L. H.; Mortimer, J. A.; Riley, K. P.; Greiner, P. A.; Markesbery, W. R. Brain infarction and the clinical expression of Alzheimer disease. The Nun Study. JAMA 227:813–817; 1997.
30. Snowdon, D. A.; Kemper, S. J.; Mortimer, J. A.; Greiner, L. H.; Wekstein, D. R.; Markesbery, W. R. Linguistic ability in early life and cognitive function and Alzheimer's disease in late life. Findings from the Nun Study. JAMA 275:528–532; 1996.
31. Tierney, M. C.; Fisher, R. H.; Lewis, A. J.; Zorzitto, M. L.; Snow, W. G.; Reid, D. W.; Nieuwstraten, P. The NINCDS-ADRDA Work Group criteria for the clinical diagnosis of probable Alzheimer's disease: A clinicopathologic study of 57 cases. Neurology 38:359–364; 1988.
32. Welsh, K. A.; Butters, N.; Mohs, R. C.; Beekly, D.; Edland, S.; Fillenbaum, G.; Heyman, A. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. Neurology 44:609–614; 1994.