

Validity and Reliability of the Preliminary NINDS Neuropathologic Criteria for Progressive Supranuclear Palsy and Related Disorders

I. LITVAN, MD, J.J. HAUW, MD, J.J. BARTKO, PhD,
P.L. LANTOS, MD, S.E. DANIEL, MD, D.S. HOROUPIAN, MD,
A. MCKEE, MD, D. DICKSON, MD, C. BANCHER, MD,
M. TABATON, MD, K. JELLINGER, MD, AND D.W. ANDERSON, PhD

Abstract. We investigated the validity and reliability of diagnoses made by eight neuropathologists who used the preliminary NINDS neuropathologic diagnostic criteria for progressive supranuclear palsy (PSP) and related disorders. The specific disorders were typical, atypical, and combined PSP, postencephalitic parkinsonism, corticobasal ganglionic degeneration, and Pick's disease. These disorders were chosen because of the difficulties in their neuropathologic differentiation. We assessed validity by measuring sensitivity and positive predictive value. Reliability was evaluated by measuring pairwise and group agreement. From a total of 62 histologic cases, each neuropathologist independently classified 16 to 19 cases for the pairwise analysis and 5 to 6 cases for the group analysis. The neuropathologists were unaware of the study design, unfamiliar with the assigned cases, and initially had no clinical information about the cases. Our results showed that with routine sampling and staining methods, neuropathologic examination alone was not fully adequate for differentiating the disorders. The main difficulties were discriminating the subtypes of PSP and separating postencephalitic parkinsonism from PSP. Corticobasal ganglionic degeneration and Pick's disease were less difficult to distinguish from PSP. The addition of minimal clinical information contributed to the accuracy of the diagnosis. On the basis of results obtained, we propose clinicopathologic diagnostic criteria to improve on the NINDS criteria.

Key Words: Neuropathology; Pick's disease; Postencephalitic parkinsonism; Progressive supranuclear palsy; Reliability; Validity.

INTRODUCTION

There are no absolute clinical markers for the accurate diagnosis of neurodegenerative disorders with extrapyramidal features. Neuropathologic examination remains the "gold standard" for their diagnosis. Criteria for the neuropathologic examination are important, as they pro-

vide operational definitions of diseases and improve communication between scientists. They also help to validate clinical criteria that are useful in case selection for drug intervention trials, pathogenetic investigations, and epidemiologic studies. Yet, despite an extensive literature identifying neuropathologic characteristics of neurodegenerative disorders, neuropathologic criteria intended for wide use have been developed only recently (1-4).

Many atypical parkinsonian disorders, including progressive supranuclear palsy (PSP) (also known as Steele-Richardson-Olszewski syndrome), corticobasal ganglionic degeneration, and postencephalitic parkinsonism, share not only clinical symptoms but also histopathologic features such as neurofibrillary tangles, gliosis, and neuronal loss (5-8). Thus, the differential diagnosis of these disorders can be challenging. For example, it is often impossible to distinguish postencephalitic parkinsonism from PSP solely on the basis of neuropathologic evidence (8). Furthermore, the basophilic inclusions in corticobasal ganglionic degeneration may appear identical to the neurofibrillary tangles in PSP (9), suggesting a relationship between these two disorders (10, 11). There are additional conundrums between corticobasal ganglionic degeneration and Pick's disease. Achromatic neurons in corticobasal ganglionic degeneration share many morphologic and immunocytochemical characteristics with swollen cells in Pick's disease, and cases of corticobasal ganglionic degeneration and Pick's disease may overlap (9, 12-14). Overlap between PSP and Pick's disease is also reported (15, 16).

From the Neuroepidemiology Branch (IL) and the Biometry and Field Studies Branch (DWA), National Institute of Neurological Disorders and Stroke, Bethesda, Maryland, USA; the Raymond Escourolle Neuropathology Laboratory (JJH), INSERM U 360, Association Claude Bernard, Hôpital de la Salpêtrière, Paris, France; the Division of Epidemiology and Research Studies (JJB), National Institute of Mental Health, Bethesda, Maryland, USA; the Department of Neuropathology (PLL), Institute of Psychiatry, London, UK; the Parkinson's Disease Society Brain Tissue Bank and Department of Neuropathology (SED), Institute of Neurology, London, UK; the Department of Pathology (Neuropathology) (DSH), Stanford School of Medicine, Stanford, California, USA; the Department of Neuropathology (AM), Massachusetts General Hospital, Boston, Massachusetts, USA; the Department of Neuropathology (DD), Albert Einstein College of Medicine, Bronx, New York, USA; the Ludwig Boltzmann Institute of Clinical Neurobiology (CB, KJ), Vienna, Austria; the Division of Neuropathology (MT), Case Western Reserve University, Cleveland, Ohio, USA.

Correspondence to: Irene Litvan, MD, Federal Building, Room 714, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892.

Dr Daniel was supported by a grant from the Parkinson's Disease Society of the United Kingdom. Prof. Lantos received support from the Medical Research Council, United Kingdom.

Presented in part at the 119th Annual Meeting of the American Neurological Association, San Francisco, California, October 11, 1994, and at the Annual Meeting of the World Federation of Neurology, Bethesda, Maryland, May 6, 1994.

TABLE 1
Characteristics of the 62 cases studied

Disorder*	No. of cases	Sex	Patient's age at disease onset† (yr)	Patient's age at death† (yr)	Duration of disease† (yr)
PSP, typical	18	6M/12F	64 ± 2	71 ± 2	6 ± 2
PSP, atypical	9	4M/5F	63 ± 3	73 ± 3	10 ± 3
PSP, combined	4	1M/3F	59 ± 5	68 ± 4	9 ± 4
PEP	8	1M/7F	35 ± 4‡	71 ± 3	36 ± 3‡
CBGD	12	4M/8F	62 ± 3	70 ± 2	8 ± 2
PKD	11	7M/4F	50 ± 3	67 ± 3	7 ± 2

* PSP, progressive supranuclear palsy; PEP, postencephalitic parkinsonism; CBGD, corticobasal ganglionic degeneration; PKD, Pick's disease.

† Values are mean ± SEM.

‡ $p < 0.0001$ (ANOVA) for the comparison with other disorders. Cases were from the departments of Neuropathology of the following hospitals: Hôpital de la Salpêtrière, Paris; Institute of Psychiatry, London; the Parkinson's Disease Society Brain Tissue Bank, London; Stanford School of Medicine, Stanford; Massachusetts General Hospital, Boston; Albert Einstein College of Medicine, Bronx, NY; the Ludwig Boltzmann Institute of Clinical Neurobiology, Vienna; Case Western Reserve University, Cleveland.

There are only a few published studies on the validity and reliability of the neuropathologic diagnosis of neurodegenerative disorders, and these pertain to Alzheimer's disease (17–20). Agreement of neuropathologists using Alzheimer's disease classifications is variable when multiple centers and diverse techniques are involved (17). By contrast, agreement is substantial when specimens originating from the investigator's own laboratory are evaluated (18).

In the present study, we investigated the accuracy of eight neuropathologists (three senior, five junior) in using the preliminary NINDS neuropathologic diagnostic criteria for PSP and related disorders (4). Specifically, our objective was to estimate the validity and reliability of the neuropathologists in classifying PSP (three subtypes), postencephalitic parkinsonism, corticobasal ganglionic degeneration, and Pick's disease. These disorders were specifically chosen because they present major diagnostic difficulties for neuropathologists when pathologic markers are similar and coexistent diseases are present. Since the information provided to neuropathologists varies in practice, the importance of clinical information in their decision making was also studied. Each participating neuropathologist was familiar with the diagnosis of neurodegenerative disorders, and seven of the eight had collaborated on the development of the diagnostic criteria (4). These criteria were based on the experience of the neuropathologists and also on an extensive review of the literature.

MATERIALS AND METHODS

Neuropathologic Criteria

The preliminary NINDS neuropathologic criteria for PSP and related disorders (4) distinguish three histologic subtypes of PSP, as well as postencephalitic parkinsonism, corticobasal gan-

glionic degeneration, and Pick's disease. The histologic subtypes of PSP consist of: (a) typical PSP, which conforms to the original description (21); (b) atypical PSP, which consists of histologic variants where the severity or distribution of abnormalities, or both, deviate from the typical pattern; and (c) combined PSP, in which typical PSP is accompanied by lesions characteristic of another disease (4, 22).

The present study followed the recommendations set forth in the criteria (4) for the minimum number of brain areas that should be sampled, the stains or histologic techniques that should be used, the main histologic features that should be considered, and the semiquantitative scoring of lesions. Cases were included only if tissue specimens were available from the globus pallidus, putamen, caudate nucleus, subthalamic nucleus, midbrain, pons, medulla, dentate nucleus of the cerebellum, hippocampus, parahippocampal gyrus, and motor, frontal, and parietal cortices (only one brain area could be omitted from the sampling), and only if the specimens were prepared by hematoxylin-eosin staining and silver impregnation (modified Bielschowsky, Bodian, Gallyas and its modifications) or for tau and ubiquitin immunohistochemistry.

Selection and Allocation of Cases

Although the participating neuropathologists had no knowledge of the study design, they did know what disorders were to be studied, and, indeed, they selected the cases from the research and clinical files of the medical centers in which they worked. Each neuropathologist had an assigned number of cases to contribute for each disorder (Table 1). The cases were coded so that both the place of origin and the "correct" diagnosis were masked. For purposes of this study, the correct diagnosis was the one made, on the basis of the NINDS criteria (4), by the neuropathologist who provided the tissue specimen.

For the study itself, there were some basic constraints. The neuropathologists would meet together for three days, during which they were to finish their assessment of cases. Assuming that 60 to 90 minutes were required to evaluate a case, each neuropathologist would have time for about 25 cases. No neu-

rotopathologist would evaluate cases that he or she had contributed.

With these constraints and the objectives of the study, it was reasonable to test the neuropathologists pairwise and as a group. For the pairwise testing, there were four randomly chosen pairs of neuropathologists, and within pairs each member independently evaluated the same cases. The allocation of cases to pairs was as follows: pair 1, 16 cases—postencephalitic parkinsonism (7 cases), typical PSP (5 cases), atypical PSP (4 cases); pair 2, 18 cases—typical PSP (5 cases), combined PSP (4 cases), corticobasal ganglionic degeneration (4 cases), Pick's disease (5 cases); pair 3, 19 cases—typical PSP (5 cases), atypical PSP (3 cases), postencephalitic parkinsonism (4 cases), corticobasal ganglionic degeneration (4 cases), Pick's disease (3 cases); pair 4, 18 cases—typical PSP (6 cases), atypical PSP (3 cases), corticobasal ganglionic degeneration (5 cases), Pick's disease (4 cases). Some cases were evaluated by more than one pair of neuropathologists. For the group testing, each of the eight neuropathologists evaluated the same 6 cases, which were chosen to represent all possible diagnoses (i.e. typical PSP, atypical PSP, combined PSP, postencephalitic parkinsonism, corticobasal ganglionic degeneration, and Pick's disease).

Data Collection and Diagnosis

The neuropathologists noted on a standardized form (available from the corresponding author) the main histologic features of each specimen, including neurofibrillary tangles, swollen cells (achromatic cells), Pick bodies, neuropil threads, Lewy bodies, basophilic inclusions, and neuritic plaques. They assessed the severity of neuronal loss, gliosis, and neurofibrillary tangles in 23 sampled areas, including the cortical and subcortical regions. Lesions were scored on a scale of 0 to 2, 0 meaning absent, 2 meaning severe. From their observations, the neuropathologists made a neuropathologic diagnosis and indicated the degree of diagnostic certainty. Also, for each case, neuropathologists commented on the quality of the specimens and noted if more samples or stains were preferred to make a better diagnosis.

Initially, only gross neuropathologic information (including a detailed written description and photographs when available) and the patient's age at death were supplied. Subsequently, clinical information was provided, and the neuropathologists were allowed to reconsider the case and to offer a revised, *clinico-pathologic* diagnosis.

Epidemiologic and Statistical Methods

We investigated validity by considering sensitivity and positive predictive value. For a given disorder and neuropathologist, *sensitivity* refers to the proportion of genuine cases correctly diagnosed by the neuropathologist, among all genuine cases of the disorder presented. *Positive predictive value* refers to the proportion of genuine cases of the disorder, among all cases so diagnosed by the neuropathologist.

We investigated the reliability of diagnoses by measuring pairwise and group agreement among the neuropathologists. The kappa (κ) statistic was used for pairwise agreement and the generalized κ statistic for group agreement. Both statistics take chance agreement into account. Like a correlation coefficient, κ varies from -1.0 (complete disagreement) to 0.0

(chance agreement) to $+1.0$ (perfect agreement). Strength of agreement was designated as poor ($\kappa < 0.0$), slight ($0.0 \leq \kappa \leq 0.2$), fair ($0.21 \leq \kappa \leq 0.4$), moderate ($0.41 \leq \kappa \leq 0.6$), substantial ($0.61 \leq \kappa \leq 0.8$), and near-perfect to perfect ($0.81 \leq \kappa \leq 1.0$), as previously suggested (23). The statistical significance between κ values was determined using the pooled κ test (24).

In other aspects of the study, analysis of variance (ANOVA), Fisher's exact test, χ^2 test of association, and logistic regression were used as appropriate. Statistical significance was defined as $p < 0.05$.

RESULTS

Table 1 presents characteristics of the 62 cases included in the study. The mean age at death was similar for each disorder considered. Postencephalitic parkinsonism had an earlier age at onset and a longer duration than the other disorders.

Table 2 gives estimates of the sensitivity and positive predictive value of the neuropathologic and clinicopathologic diagnoses. Individual estimates of sensitivity and positive predictive value were obtained for each neuropathologist, and the median (middle value) and range (minimum to maximum values) are presented in the table. Only for PSP, corticobasal ganglionic degeneration, and Pick's disease did the median neuropathologic sensitivity reach or exceed 80%. On the other hand, for atypical PSP, the median neuropathologic sensitivity was 0%. For each disorder, the median clinicopathologic sensitivity was greater than, or equal to, the median neuropathologic sensitivity. Analogously, the median clinicopathologic positive predictive value exceeded or equaled the median neuropathologic positive predictive value. Also, the median neuropathologic positive predictive value exceeded 75% for every disorder except the PSP subtypes.

Table 3 shows a cross-classification of the neuropathologic and clinicopathologic diagnoses with the correct diagnoses. For example, with respect to typical PSP, there were 18 cases on which 41 observations were made. Of these 41 observations, 29 were correctly diagnosed neuropathologically, and 12 were misdiagnosed neuropathologically (5 as atypical PSP, 5 as combined PSP, and 2 as postencephalitic parkinsonism). Atypical and combined PSP were misdiagnosed more often than they were correctly diagnosed, regardless of whether the diagnosis was neuropathologic or clinicopathologic. The misdiagnosis of postencephalitic parkinsonism was greatly reduced when the diagnosis was made clinicopathologically rather than neuropathologically ($28 - 24 = 4$ vs $28 - 12 = 16$ misdiagnosed cases).

Table 4 displays κ statistics showing pairwise reliability of the neuropathologists. The relevant pairs are presented for each disorder. Only for Pick's disease did all the pairs (i.e. pairs 2, 3, 4) show near-perfect to perfect agreement, for both neuropathologic and clinicopathologic diagnoses.

TABLE 2
Sensitivity and Positive Predictive Value of the Neuropathologic and Clinicopathologic Diagnoses

Disorder	Sensitivity		Positive predictive value		Relative frequency (%)
	Neuropathologic diagnosis (%)	Clinicopathologic diagnosis (%)	Neuropathologic diagnosis (%)	Clinicopathologic diagnosis (%)	
PSP	89 (78–100)	89 (78–100)	82 (67–100)	95 (75–100)	50 (42–56)
Typical PSP	71 (50–100)	75 (50–100)	60 (50–67)	70 (50–80)	27 (25–33)
Atypical PSP	0 (0–50)	17 (0–50)	0 (0–40)	25 (0–100)	25 (16–25)
Combined PSP	25 (25)	25 (25)	50 (50)	50 (50)	22 (22)
PEP	41 (25–57)	88 (75–100)	100 (80–100)	100 (100)	33 (21–44)
CBGD	100 (40–100)	100 (40–100)	100 (67–100)	100 (50–100)	22 (21–28)
PKD	80 (75–100)	80 (75–100)	78 (75–100)	100 (75–100)	22 (16–28)

PSP = progressive supranuclear palsy; PEP = postencephalitic parkinsonism; CBGD = corticobasal ganglionic degeneration; PKD = Pick's disease.

All eight neuropathologists evaluated cases of PSP, including typical PSP; six evaluated cases of atypical PSP, CBGD, and PKD; two evaluated cases of combined PSP. Values computed for each neuropathologist; presented are median and range (range in parentheses). Relative frequency is the proportion of cases of a given disorder among all assigned cases.

For postencephalitic parkinsonism, this level of agreement was reached by all the pairs only for clinicopathologic diagnoses. With regard to typical PSP, all the pairs achieved at least moderate agreement, regardless of whether the diagnoses were neuropathologic or clinicopathologic. For each of atypical PSP, combined PSP, and corticobasal ganglionic degeneration, the pairs did not consistently achieve this level of agreement. Note that the pairs had different case assignments, with some overlap. Thus, a lower or higher κ might, in part, be a reflection of more difficult or less difficult cases to evaluate.

To complement the pairs analysis, generalized κ statistics were used to measure the reliability of diagnoses by the neuropathologists as a group. As stated earlier, each neuropathologist was to examine the same six cases, consisting of one of each from typical PSP, atypical PSP, combined PSP, postencephalitic parkinsonism, corticobasal ganglionic degeneration, and Pick's disease. However, the neuropathologist was not permitted to examine

any case contributed by his or her own medical center. Because of this constraint, each of two neuropathologists examined all six cases, and each of the other six examined five cases (the case not examined for each neuropathologist was treated as missing data). In contrasting PSP vs non-PSP, the neuropathologists achieved moderate agreement for both neuropathologic ($\kappa = 0.58$) and clinicopathologic diagnoses ($\kappa = 0.60$). In contrasting all six disorders, the neuropathologists achieved moderate agreement ($\kappa = 0.52$) for neuropathologic diagnoses and substantial agreement ($\kappa = 0.71$) for clinicopathologic diagnoses. Regardless of the comparison, the group significantly improved ($p \leq 0.001$) in their agreement with the addition of clinical information.

To investigate if suboptimal quality of some cases may have adversely influenced the study results, we related, for each neuropathologist, the accuracy of neuropathologic diagnosis with perceived needs of further sampling areas or additional stains, and with informal assessment

TABLE 3
Neuropathologic and Clinicopathologic Diagnoses Versus the "Correct" Diagnosis

Correct diagnosis	Neuropathologic (clinicopathologic) diagnosis							No. of observations	No. of cases
	Typical PSP	Atypical PSP	Combined PSP	PEP	CBGD	PKD	Other		
Typical PSP	29 (30)	5 (6)	5 (5)	2 (0)	0 (0)	0 (0)	0 (0)	41	18
Atypical PSP	12 (12)	2 (4)	2 (2)	2 (0)	1 (1)	2 (2)	2 (2)	23	9
Combined PSP	5 (3)	1 (2)	7 (5)	0 (0)	0 (1)	0 (1)	2 (3)	15	4
PEP	6 (0)	5 (2)	5 (2)	12 (24)	0 (0)	0 (0)	0 (0)	28	8
CBGD	1 (2)	1 (1)	2 (2)	0 (0)	21 (21)	2 (1)	2 (2)	29	12
PKD	0 (0)	0 (0)	1 (1)	0 (0)	1 (1)	23 (23)	2 (2)	27	11
Total	53 (47)	14 (15)	22 (17)	16 (24)	23 (24)	27 (27)	8 (9)	163	62

PSP = progressive supranuclear palsy; PEP = postencephalitic parkinsonism; CBGD = corticobasal ganglionic degeneration; PKD = Pick's disease; Other = other disorders not specifically studied.

Counts of neuropathologic and clinicopathologic diagnoses are presented side-by-side, with the latter given in parentheses. Note that the same case may be observed by two or more neuropathologists.

TABLE 4
Reliability of the Diagnosis of the Specific Disorders

Disorder	Pair No.	Neuro- pathologic diagnosis κ	Clinico- pathologic diagnosis κ
Typical PSP	1	0.54	0.71
	2	0.72	0.51
	3	0.89	0.76
	4	0.61	0.6
Atypical PSP	1	0.13	0.64
	3	0	0
	4	-0.08	-0.08
Combined PSP	2	0.22	-0.2
	1	0.23	1
PEP	3	-0.07	1
	2	0.2	0.22
CBGD	3	1	1
	4	0.49	0.49
PKD	2	0.85	0.85
	3	1	1
	4	0.82	0.82

κ = kappa (pairwise); PSP = progressive supranuclear palsy; PEP = postencephalitic parkinsonism; CBGD = corticobasal ganglionic degeneration; PKD = Pick's disease.

of quality of specimens provided. An accurate neuropathologic diagnosis was the one that matched the correct diagnosis. Various statistical procedures (Fisher's exact test, χ^2 test of association, and logistic regression) generally showed no significant associations with the factors analyzed.

The mean years of neuropathologic diagnostic experience were: 5 ± 0.5 for junior neuropathologists, and 22 ± 0.7 years for the seniors ($p < 0.001$). To assess whether seniority of the neuropathologists affected the study results, we compared the percentage of accurate diagnoses (neuropathologic and clinicopathologic) for both the senior and junior groups of neuropathologists. There were no clear differences between groups.

DISCUSSION

This study is the first to investigate the validity and reliability of the neuropathologic diagnosis of several neurodegenerative disorders. Some reliability studies of the neuropathologic diagnosis of Alzheimer's disease were conducted (17-19), but they contrasted disease cases with normal cases, and did not consider different disorders or subtypes of the same disorder. Our specific interest was in neurodegenerative disorders with extrapyramidal features, and we focused on disorders likely to present major diagnostic difficulties for neuropathologists.

To increase the relevance of our findings, we used sampling areas and essential stains recommended for routine practice (4). In evaluating cases, the neuropathologists reported some problems with interpreting the stains pre-

sented, and occasionally expressed the desire for more sampling areas and better stains. However, a special methodologic investigation built into the overall study suggested that the accuracy of neuropathologic diagnoses was not related to the perception that more sampling areas or stains were needed. In prospective studies where staining and blocking techniques are standardized, this issue of quality should not be a problem. However, in a multicenter study (25) in which raters from 11 laboratories evaluated 6 cases of Alzheimer's disease by applying their own techniques to serially cut sections from the same specimens, variations in staining methods and techniques evidently contributed greatly to interrater variability.

The neuropathologists who evaluated the cases in the current study were all familiar with the classification of neurodegenerative disorders. This similarity in their experience probably helped to minimize some potential sources of disagreement. Although three of the neuropathologists were senior and five junior, this disparity did not seem important in terms of the results obtained. However, time constraints and fatigue may have influenced the pathologic diagnoses. On the other hand, we should stress that substantial or perfect agreement does not mean that the neuropathologists achieved the correct diagnosis; they can all agree and still be wrong.

Correct diagnoses were needed to investigate validity. It would have been desirable, in some respects, to create an independent panel of expert neuropathologists to review the cases and by consensus arrive at the correct diagnoses. However, for budgetary and operational reasons, we chose to accept as correct the diagnosis of each case given by the neuropathologist from the center which donated the case. This diagnosis, based on the NINDS criteria (4), was considered correct because the donor had the opportunity to sample more brain areas, the option of using additional stains, access to complete clinical records, and the opportunity to discuss the case with clinicians. Recall that, during the study, no neuropathologist examined cases from his or her own center, or was even aware of the study design.

Progressive Supranuclear Palsy

For both neuropathologic and clinicopathologic diagnoses, the PSP subtypes presented problems. Compared with diagnosing atypical and combined PSP, the neuropathologists were better, by far, at recognizing typical PSP, although every neuropathologist misdiagnosed at least 33% of the cases that he or she classified as typical PSP. The misdiagnoses usually pertained to atypical or combined PSP.

The neuropathologists almost always failed to recognize atypical PSP, mistaking it mainly for typical PSP. The neuropathologic criteria proposed for atypical PSP, specifically, the presence of either less brain stem neu-

ro fibrillary involvement or more cortical changes (4, 22), were difficult to implement. Failure to recognize atypical PSP suggests that it is an artificial neuropathologic (or clinicopathologic) entity. Indeed, the nosological importance of typical or atypical clinical presentations of PSP is also open to question, as there appears to be no histologic basis for that distinction (25). Improved neuropathologic or clinicopathologic diagnostic criteria are needed if the atypical PSP subtype is to be maintained.

As a subtype, combined PSP also deserves scrutiny. Too often, the neuropathologists failed to recognize combined PSP. This difficulty was related to the reluctance of some neuropathologists to use the diagnostic criteria pertaining to this disorder, especially when vascular lesions were present. We propose that combined PSP be restricted to cases where there are infarcts in the brain stem or basal ganglia because lesions in those areas could modify the clinical symptoms. When infarcts occur in other areas or when a nonvascular disorder is also present, the diagnosis should be typical PSP and the coexisting disease rather than combined PSP. Although a strong association between PSP and Alzheimer's disease has been reported (26), we did not specifically address this issue.

Postencephalitic Parkinsonism

Postencephalitic parkinsonism shares with PSP the presence of extensive subcortical neurofibrillary tangles. The neuropathologists found it difficult to differentiate postencephalitic parkinsonism from PSP on the basis of neuropathologic information only. However, the sensitivity and reliability of the diagnosis of postencephalitic parkinsonism improved dramatically when clinical information was added. The distinction between the two entities was largely determined by a history of encephalitis lethargica and oculogyric crisis and also disease duration. The duration of symptoms was significantly longer, and the age at onset earlier, in postencephalitic parkinsonism than in PSP.

Our data suggest that more refined neuropathologic criteria or the use of supplementary techniques (e.g. biochemical or genetic studies) are needed to distinguish postencephalitic parkinsonism and PSP. Although immunocytochemical and ultrastructural studies are unable to easily differentiate the neurofibrillary tangles in postencephalitic parkinsonism from those in Alzheimer's disease and PSP (27–29), biochemical studies show differences between the triple tau proteins of Alzheimer's disease and the double tau proteins of PSP (30, 31). The histologic features originally perceived as helpful for separating postencephalitic parkinsonism from PSP, such as the absence of neurofibrillary involvement in the oculomotor complex, trochlear nucleus, or basis pontis, are insufficient for distinguishing these two disorders.

Corticobasal Ganglionic Degeneration

The pairs of neuropathologists varied in their levels of agreement with respect to the diagnosis of corticobasal ganglionic degeneration. The pattern was the same, regardless of whether the diagnoses were neuropathologic or clinicopathologic. Disagreements within pairs might be related to different proficiency levels with the diagnostic criteria, the heterogeneous nature of the disorder, or the absence of characteristic features that define the disorder as a distinct entity. It is unlikely that these disagreements were related to the level of difficulty of cases with corticobasal ganglionic degeneration, as at least one neuropathologist always made the correct diagnosis. Furthermore, the proportion of corticobasal ganglionic degeneration cases evaluated was similar for each neuropathologist, so the relative frequency should not have influenced case detection (32).

The misdiagnoses of corticobasal ganglionic degeneration might be related to the basophilic inclusions characteristic of the condition. These inclusions were confused with the neurofibrillary tangles of PSP, and, less often, with Pick bodies. Recent studies (33, 34) have demonstrated that there are cytochemical, ultrastructural, and biochemical differences between the basophilic inclusions in corticobasal ganglionic degeneration and the neurofibrillary tangles in PSP. In addition, non-amyloid cortical plaques have been suggested as distinctive structures only found in corticobasal ganglionic degeneration (35). On the other hand, these structures were shown to be collections of abnormal tau in the distal processes of astrocytes (36), also found in PSP (37–39). It remains to be determined whether corticobasal ganglionic degeneration should be classified as a distinct neuropathologic or clinicopathologic disorder. The results of the present study suggest that the disorder is distinct.

Pick's Disease

Pick's disease was rather easily diagnosed neuropathologically on the basis of Pick bodies, which happened to be present in every case of this disorder. However, the mean positive predictive value for the diagnosis of Pick's disease improved (from 78 to 100%) when clinical information was provided—clearly supplementing the importance of Pick bodies for the diagnosis of this disorder.

Revised Diagnostic Criteria

On the basis of the results and considerations mentioned above, we have revised the preliminary NINDS criteria (4) and organized them in concise form (Table 5). The following is a summary of the changes: (a) Clinical history compatible with the diagnosis is now required; (b) Atypical PSP is excluded as a PSP subtype (further neuropathological studies of this entity are needed); (c) Combined PSP is the same as typical PSP, but

TABLE 5
NINDS Diagnostic Criteria for PSP and Related Disorders¶

Disorder	Inclusion criteria	Exclusion criteria
Typical PSP	A high density of neurofibrillary tangles and neuropil threads in at least three of the following areas: pallidum, subthalamic nucleus, substantia nigra, or pons, <i>and</i> a low-to-high density of neurofibrillary tangles or neuropil threads in at least three of the following areas: striatum, oculomotor complex, medulla, or dentate nucleus* and clinical history compatible with PSP	Large or numerous infarcts; marked diffuse or focal atrophy; Lewy bodies; changes diagnostic of Alzheimer's disease; oligodendroglial argyrophilic inclusions; Pick bodies; diffuse spongiosis; prion P-positive amyloid plaques
Combined PSP	Same as for typical PSP, with the coexistence of infarcts in the brainstem or basal ganglia, or both**	
PEP	A low density of neurofibrillary tangles and neuropil threads in at least five of the following areas: pallidum, subthalamic nucleus, substantia nigra, pons, medulla, or dentate nucleus, with minimal involvement of oculomotor complex, trochlear nucleus, and inferior olivary nucleus; other lesions include neuronal loss and gliosis in affected areas; clinical history compatible with PEP	Large or numerous infarcts; marked diffuse or focal atrophy; Lewy bodies; changes diagnostic of Alzheimer's disease; oligodendroglial argyrophilic inclusions; Pick bodies; diffuse spongiosis; prion P-positive amyloid plaques
CBGD	Circumscribed or lobar atrophy in parietal or frontoparietal areas; tau-positive neurons in the cortex; swollen and achromatic neurons, basophilic inclusions, numerous neuropil threads, and severe neuronal loss in the substantia nigra, basal ganglia, or dentatorubrothalamic tract; astrogliosis and spongiosis in subcortical white matter and clinical history compatible with corticobasal ganglionic degeneration	Large or numerous infarcts; Lewy bodies; changes diagnostic of Alzheimer's disease; oligodendroglial argyrophilic inclusions; Pick bodies; prion P-positive amyloid plaques
PKD	Lobar atrophy of the frontal and anterior temporal lobes, with Pick argyrophilic inclusions (Pick bodies), and occasionally atrophy of the caudate nucleus, pallidum, basal ganglia, or substantia nigra; other lesions include Pick cells, massive neuronal loss, astrogliosis, spongiosis (often laminar) and clinical history compatible with Pick's disease	Large or numerous infarcts; neurofibrillary tangles; Lewy bodies; changes diagnostic of Alzheimer's disease; oligodendroglial argyrophilic inclusions; prion P-positive amyloid plaques

¶ This is a revision of the preliminary NINDS criteria, from Hauw et al (4).

PSP = progressive supranuclear palsy; PEP = postencephalitic parkinsonism; CBGD = corticobasal ganglionic degeneration; PKD = Pick's disease.

* Tau-positive astrocyte processes or astrocyte cell bodies in the areas of neurofibrillary tangles and neuropil threads confirm the diagnosis; other lesions include various degrees of neuronal loss and gliosis in affected areas. ** The diagnosis should be typical PSP and the coexisting disease when infarcts occur in nonaffected areas or when a nonvascular disorder is also present.

with the further requirement of coexistent infarcts in the brainstem or basal ganglia, or both.

Concluding Remarks

The present study points to the still poor neuropathologic characterization of the disorders studied, with Pick's disease being possibly the only exception, and reaffirms the need for standardizing the neuropathologic criteria for the diagnosis of many neurodegenerative disorders. More easily recognized were disorders with unquestionable markers, such as Pick bodies. Less easily recognized were disorders characterized either by relatively nonspecific, similar markers that are supposed to be diagnostic by virtue of their distribution or density (e.g. neurofibrillary tangles of "normal aging," Alzheimer's disease,

PSP, or postencephalitic parkinsonism, and ballooned cells of Pick's disease or corticobasal ganglionic degeneration) or by markers that are poorly characterized, debated, and generally nondiagnostic (e.g. basophilic inclusions of corticobasal ganglionic degeneration).

Our results suggest that, with present sampling and staining methods, neuropathology alone is not fully adequate as the gold standard for differentiating PSP from non-PSP or other specific neurodegenerative disorders with extrapyramidal features. However, the diagnosis can be better secured by the addition of clinical information. Therefore, as in Alzheimer's disease criteria [NINDS-ADRDA (1) and CERAD (3)], the gold standard should be the clinicopathologic, rather than the neuropathologic, diagnosis. At times, the diagnosis may depend heavily on

clinical information (e.g. when distinguishing between PSP and postencephalitic parkinsonism), but it may also rely mainly on neuropathologic analysis (e.g. as in Pick's disease). In practice, the correct diagnosis will be more readily established when neuropathologists receive accurate clinical information.

The study emphasizes the relevance of ensuring the validity and reliability of neuropathologic data in multicenter studies of neurodegenerative disorders that depend on neuropathologic findings to confirm the clinical diagnosis. Agreement on the neuropathologic diagnosis might be lower in multicenter studies that rely on diverse sampling and staining techniques.

Finally, the present study might serve as a framework for analyzing the accuracy of the neuropathologic diagnosis of other types of neurodegenerative disorders, such as diffuse Lewy body disease, lobar or circumscribed atrophy, and vascular or mixed dementias.

ACKNOWLEDGMENTS

We wish to thank Drs Yves Agid and Gustavo Roman for their contributions to this project, Ms Margaret Eaglin for assistance in developing the histopathologic form, Dr Luis D'Olaberriague and Mr Kurt Lange for technical assistance, and Ms B. J. Hesse for skillful editing. We are grateful for the expert technical help provided by the staffs of the participating laboratories.

REFERENCES

1. Khachaturian ZS. Diagnosis of Alzheimer's disease. *Arch Neurol* 1985;42:1097-1105
2. 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
3. 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
4. Hauw JJ, Daniel SE, Dickson D, et al. Preliminary NINDS neuropathologic criteria for Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy). *Neurology* 1994;44:2015-19
5. Gibb WRG, Luthert PJ, Marsden CD. Corticobasal degeneration. *Brain* 1989;112:1171-92
6. Gibb WRG, Luthert PJ, Marsden CD. Clinical and pathological features of corticobasal degeneration. *Adv Neurol* 1990;53:51-54
7. Duvoisin R, Yahr MD. Encephalitis and parkinsonism. *Arch Neurol* 1965;12:227-39
8. Geddes JF, Hughes AJ, Daniel SE. Pathological overlap in cases of parkinsonism associated with neurofibrillary tangles. A study of recent cases of postencephalitic parkinsonism and comparison with progressive supranuclear palsy and Guamanian parkinsonism-dementia complex. *Brain* 1993;116:281-302
9. Gibb WRG. Neuropathology of Parkinson's disease and related syndromes. *Neurol Clin* 1992;10:361-76
10. Wakabayashi K, Oyanagi K, Makifuchi T, et al. Corticobasal degeneration: Etiopathological significance of the cytoskeletal alterations. *Acta Neuropathol (Berl)* 1994;87:545-53
11. Mori H, Nishimura M, Namba Y, Oda M. Corticobasal degeneration: A disease with widespread appearance of abnormal tau and neurofibrillary tangles, and its relation to progressive supranuclear palsy. *Acta Neuropathol (Berl)* 1994;88:113-21
12. Lennox G, Jackson M, Lowe J. Corticobasal degeneration manifesting as a frontal lobe dementia. *Ann Neurol* 1994;36:273-74
13. Brown J, Lantos PL, Collinge J, Rossor MN. Familial Pick's syndrome with corticobasal inclusion bodies: A clinicopathological study. (Abstract) *Neuropathol Appl Neurobiol* 1993;19:448
14. Jendroska K, Rossor MN, Mathias CJ, Daniel SE. Morphological overlap between corticobasal degeneration and Pick's disease: A clinicopathological report. *Mov Disord* 1995;10:111-14
15. Arima K, Murayama S, Oyanagi S, et al. Presenile dementia with progressive supranuclear palsy tangles and Pick bodies: An unusual degenerative disorder involving the cerebral cortex, cerebral nuclei, and brain stem nuclei. *Acta Neuropathol (Berl)* 1992;84:128-34
16. Seitelberger F, Gross H, Pilz P. Pick's disease: A neuropathological study. In: Hirano A, Miyoshi K, eds. *Neuropsychiatric disorders in the elderly*. Tokyo: Igaku-Shoin, 1983;87-117
17. Duyckaerts C, Delaère P, Hauw JJ, et al. Rating of the lesions in senile dementia of the Alzheimer type: Concordance between laboratories. *J Neurol Sci* 1990;97:295-323
18. Paulus W, Baner C, Jellinger K. Interrater reliability in the neuropathological diagnosis of Alzheimer's disease. *Neurology* 1992;42:329-32
19. 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;3:303-15
20. McKeel DW Jr, Ball MJ, Price JL, Smith DS, Miller JP, Berg L. Interlaboratory histopathologic assessment of Alzheimer's neuropathology: Different methodologies yield comparable diagnostic results. *Alzheimer Dis Assoc Disord* 1993;7:136-53
21. Steele JC, Richardson JC, Olszewski J. Progressive supranuclear palsy. *Arch Neurol* 1964;10:333-59
22. Lantos PL. The neuropathology of progressive supranuclear palsy. *J Neural Transm Suppl* 1994;42:137-52
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74
24. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York: John Wiley & Sons, 1981.
25. Daniel SE, De Bruin VMS, Lees AJ. The clinical and pathological spectrum of Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy). *Brain* 1995;118:759-70
26. Gearing M, Olson DA, Watts RL, Mirra SS. Progressive supranuclear palsy: Neuropathologic and clinical heterogeneity. *Neurology* 1994;44:1015-24
27. Ishii T, Nakamura Y. Distribution and ultrastructure of Alzheimer's neurofibrillary tangles in postencephalitic Parkinsonism of economo type. *Acta Neuropathol (Berl)* 1981;55:59-62
28. Tabaton M, Whitehouse PJ, Perry G, Davies P, Aulilio-Gambetti L, Gambetti L. Alz 50 recognizes abnormal filaments in Alzheimer's disease and progressive supranuclear palsy. *Ann Neurol* 1988;24:407-13
29. Yen SH, Horoupian DS, Terry RD. Immunocytochemical comparison of neurofibrillary tangles in senile dementia of Alzheimer type, progressive supranuclear palsy, and postencephalitic parkinsonism. *Ann Neurol* 1983;13:172-75
30. Flament S, Delacourte A, Verny M, Hauw JJ, Javoy-Agid F. Abnormal tau proteins in progressive supranuclear palsy. Similarities and differences with the neurofibrillary degeneration of the Alzheimer type. *Acta Neuropathol (Berl)* 1991;81:591-96
31. Vermersch P, Robitaille Y, Bernier L, Wattez A, Gauvreau D, Delacourte A. Biochemical mapping of neurofibrillary degeneration in a case of progressive supranuclear palsy: Evidence for general cortical involvement. *Acta Neuropathol (Berl)* 1994;87:572-77
32. Baldessarini RJ, Kinklestein S, Arana GW. The predictive power of diagnostic tests and the effect of prevalence of illness. *Arch Gen Psychiatry* 1983;40:569-73

33. Uchihara T, Mitani K, Mori H, et al. Abnormal cytoskeletal pathology peculiar to corticobasal degeneration is different from that of Alzheimer's disease or progressive supranuclear palsy. *Acta Neuropathol (Berl)* 1994;88:379-83
34. Ksiezak-Reding H, Dickson DW, Tracz E, Wall JS. Biochemical and structural analysis of paired helical filaments (PHF) in corticobasal degeneration. (Abstract) *Neurobiol Aging* 1994;15:S21
35. Mattiace LA, Wu E, Aronson M, Dickson DW. A new type of neuritic plaque without amyloid in corticonigral degeneration with neuronal achromasia. (Abstract) *J Neuropathol Exp Neurol* 1991;50:310
36. Feany MB, Dickson DW. Widespread cytoskeletal pathology characterizes corticobasal degeneration. *Am J Pathol* 1995;146:1388-96
37. Nishimura M, Namba Y, Ikeda K, Oda M. Glial fibrillary tangles with straight tubules in the brains of patients with progressive supranuclear palsy. *Neurosci Lett* 1992;143:35-38
38. Yamada T, McGeer PL, McGeer EG. Appearance of paired nucleated, Tau-positive glia in patients with progressive supranuclear palsy brain tissue. *Neurosci Lett* 1992;135:99-102
39. Inagaki T, Ishino H, Seno H, Yamamori C, Iijima M. An autopsy case of PSP with astrocytic inclusions. *Jpn J Psychiatry Neurol* 1994;48:85-89

Received August 8, 1995

Revision received September 14, 1995

Accepted September 15, 1995