Quantitative application of the primary progressive aphasia consensus criteria

Meredith R. Wicklund,
MD
Joseph R. Duffy, PhD
Edythe A. Strand, PhD
Mary M. Machulda, PhD
Jennifer L. Whitwell,
PhD
Keith A. Josephs, MD,
MST, MSc

Correspondence to
Dr. Josephs:
josephs.keith@mayo.edu

ABSTRACT

Objective: To determine how well the consensus criteria could classify subjects with primary progressive aphasia (PPA) using a quantitative speech and language battery that matches the test descriptions provided by the consensus criteria.

Methods: A total of 105 participants with a neurodegenerative speech and language disorder were prospectively recruited and underwent neurologic, neuropsychological, and speech and language testing and MRI in this case-control study. Twenty-one participants with apraxia of speech without aphasia served as controls. Select tests from the speech and language battery were chosen for application of consensus criteria and cutoffs were employed to determine syndromic classification. Hierarchical cluster analysis was used to examine participants who could not be classified.

Results: Of the 84 participants, 58 (69%) could be classified as agrammatic (27%), semantic (7%), or logopenic (35%) variants of PPA. The remaining 31% of participants could not be classified. Of the unclassifiable participants, 2 clusters were identified. The speech and language profile of the first cluster resembled mild logopenic PPA and the second cluster semantic PPA. Gray matter patterns of loss of these 2 clusters of unclassified participants also resembled mild logopenic and semantic variants.

Conclusions: Quantitative application of consensus PPA criteria yields the 3 syndromic variants but leaves a large proportion unclassified. Therefore, the current consensus criteria need to be modified in order to improve sensitivity. *Neurology*® 2014;82:1119-1126

GLOSSARY

agPPA = agrammatic variant of PPA; **AOS** = apraxia of speech; **BNT** = Boston Naming Test; **DKEFS** = Delis-Kaplan Executive Function System Card Sort; **lvPPA** = logopenic variant of PPA; **MoCA** = Montreal Cognitive Assessment battery; **PPA** = primary progressive aphasia; **PPAOS** = primary progressive apraxia of speech; **PPTT** = Pyramids and Palm Trees Test; **Rey-O** = Rey-Osterrieth Complex Figure Test; **svPPA** = semantic variant of PPA; **WAB** = Western Aphasia Battery-revised; **WAB-AQ** = Western Aphasia Battery-aphasia quotient.

The term primary progressive aphasia (PPA) is synonymous with a neurodegenerative syndrome in which language impairment is the presenting and most salient feature.¹ In 2011, a group of investigators developed consensus criteria recognizing 3 syndromic variants of PPA.² The first, known as the agrammatic variant (agPPA), is characterized by written and verbal language that is grammatically flawed and sometimes associated with apraxia of speech (AOS). The second, known as the semantic variant (svPPA), is characterized by anomia with loss of single-word knowledge. The third, known as the logopenic variant (lvPPA), is characterized by anomia without loss of word knowledge, difficulty with sentence repetition, and, often, phonologic errors. Recently, the consensus criteria have been challenged, particularly in regards to whether subjects with AOS only, or AOS dominant syndromes, should be classified as agPPA³⁻⁵ and whether lvPPA exists as a discrete linguistic syndrome.⁶

METHODS Participants. Between July 2010 and January 2013, all subjects with a suspected degenerative speech and language disorder who presented to the Mayo Clinic in Rochester, Minnesota, were prospectively recruited. Subjects were consecutively recruited to reduce any selection bias. Only subjects over the age of 18, who spoke English as their primary language, and had an independent

Editorial, page 1108

Supplemental data
at Neurology.org

From the Division of Behavioral Neurology (M.R.W., K.A.J.), Speech Pathology, Department of Neurology (J.R.D., E.A.S.), Division of Neurocognitive Disorders, Department of Psychiatry and Psychology (M.M.M.), and Division of Radiology Research, Department of Radiology (J.L.W.), Mayo Clinic, Rochester, MN.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

informant able to provide information about functional capabilities, were included. During this period, 110 subjects were screened and 105 participants were included in the study. Five subjects were excluded due to a stroke, meeting criteria for Alzheimer type dementia,⁷ MRI was contraindicated, speech or language was so impaired that meaningful analysis could not be performed, and consensus agreement between 2 independent speech pathologists (J.R.D. and E.A.S.) did not indicate the presence of a degenerative speech or language disorder.

Standard protocol approvals and patient consents. The study was approved by the Mayo Clinic institutional review board and all subjects consented for enrollment into the study.

Speech and language assessment. All participants underwent a detailed speech and language assessment as described previously.⁴ The Western Aphasia Battery–revised (WAB),⁸ a 22-item version of part V of the Token Test by De Renzi and Vignolo,⁹ the 15-item Boston Naming Test (BNT),¹⁰ and the Pyramids and Palm Trees Test¹¹ (PPTT) were used as part of the language assessment. Motor speech abilities, including assessment for dysarthria and AOS, were assessed according to internally derived rating scales as previously described.⁴ The presence or absence of phonologic errors was assessed on a 0–4 severity scale: 0 = not present; 1 = detectable but not frequent; 2 = frequent but not pervasive; 3 = nearly always evident but not marked in severity; 4 = nearly always evident and marked in severity. Quantitative scores and video recordings of all crucial aspects of the speech and language assessment were reviewed for all study participants by 2 study authors (J.R.D. and E.A.S.) for consensus agreement about the presence or absence of aphasia, AOS, dysarthria, and phonologic errors, and the severity of each disorder.

Study participants who, by consensus agreement, demonstrated solely AOS without features of aphasia, i.e., primary progressive apraxia of speech (PPAOS),⁴ served as control participants for the purpose of this study ($n = 21$). Any participant who, by consensus agreement, demonstrated features of progressive aphasia, regardless of the presence of AOS, was analyzed for the diagnostic features of each PPA variant as described below. All participants presented with a chief complaint of speech or language dysfunction with the speech or language dysfunction being the cause for disrupting normal activities of daily living, e.g., writing a letter, giving a speech, or talking on the phone. None complained of early deficits in other cognitive domains.

Based on the recommended tasks for evaluation of speech and language functions provided by the consensus guidelines for the classification of PPA,² subsets of our battery were chosen accordingly for assessment of the presence or absence of each diagnostic feature of each variant (table e-1 on the *Neurology*[®] Web site at Neurology.org). A threshold of 1.5 SDs worse than the PPAOS control participants was set as the cutoff for impairment. In analyzing AOS, a threshold of 1.5 SDs worse than a control population of 10 participants without AOS by consensus agreement on the AOS rating scale⁴ was set as the cutoff for impairment. In assessing agrammatism, a raw score of ≤ 5 on the fluency rating of the WAB was used given the operational definition of that assigned score. In identifying impaired single-word comprehension, any score < 60 on the Auditory Word Recognition subtest of the WAB was set as the cutoff for impairment. Phonologic errors were scored as present/absent. As an additional sensitivity analysis, classification was also assessed using a 1.0 SD cutoff across tests.

Neurologic and neuropsychological examination. As previously described,⁴ all participants underwent detailed neurologic examination (K.A.J.). Testing included the Montreal Cognitive Assessment battery (MoCA),¹² Frontal Behavioral Inventory,¹³ brief questionnaire form of the Neuropsychiatric Inventory,¹⁴ the Limb Apraxia subscale of the WAB,⁸ and the Movement

Disorders Society–sponsored version of the Unified Parkinson’s Disease Rating Scale part III.¹⁵

Detailed neuropsychological testing was performed by a psychometrist with oversight by a clinical neuropsychologist (M.M.M.). The battery included tests of motor speed (Trail Making Test A¹⁶), executive function (Trail Making Test B¹⁶ and Delis-Kaplan Executive Function System Card Sort [DKEFS]),¹⁷ learning and memory (Logical Memory and Visual Reproduction subsets of the Wechsler Memory Scale–III¹⁸ and Auditory Verbal Learning Test¹⁹) and visuospatial and visuo-perceptual functions (Cube Analysis and Incomplete Letters²⁰ and the Rey-Osterrieth Complex Figure Test^{19,21}).

Neuroimaging. All participants underwent a 3T volumetric MRI using a standardized protocol and standard preprocessing corrections for intensity inhomogeneity and gradient warping were applied as previously described.⁴ Patterns of gray matter loss were assessed in each PPA group compared to a cohort of 40 age- and sex-matched healthy participants using voxel-based morphometry²² in SPM5. All healthy controls underwent the same imaging protocol. All scans were normalized and segmented using customized priors created using all PPA participants and healthy controls in the study. Gray matter images were modulated and smoothed at 8 mm full-width at half maximum. Statistical comparisons were performed using t tests with age and sex included as covariates. All analyses were assessed after correction for multiple comparisons at familywise error of $p < 0.05$ and uncorrected at $p < 0.001$.

Statistical analysis. Statistical analysis was performed using JMP 9.0.1 (SAS Institute Inc., Cary, NC). Nonparametric Kruskal-Wallis testing was used to compare demographics and clinical features across groups. A p value of < 0.05 was considered significant. Post hoc comparisons were made using the Dwass-Steel-Critchlow-Fligner test, which corrects for multiple comparisons,²³ when Kruskal-Wallis was significant. A hierarchical cluster analysis was performed using the speech and language test scores for all participants who were unclassifiable. Demographics, speech and language characteristics, and neuropsychological profiles were compared across clusters, as described above.

RESULTS Subject classification. Eighty-four participants were analyzed for this study while the other 21 PPAOS participants served as controls. Of the 84 participants, 58 met the minimum diagnostic features for at least 1 of the 3 PPA variants. Twenty-three participants (27%) met criteria for agPPA, 6 (7%) for svPPA, and 31 (36%) for lvPPA. The remaining 26 participants (31%) did not meet minimum diagnostic criteria for classification of any variant and hence were labeled as unclassified. One participant classified as agPPA and another as svPPA simultaneously met criteria for lvPPA. Utilizing a cutoff of 1.0 resulted in identical classification for 92% of the participants: 2 agPPA and 2 lvPPA participants became unclassified while 3 unclassified cluster 1 participants fulfilled criteria for lvPPA.

Speech and language profiles. There were no demographic differences across the groups (table 1). The svPPA participants, followed by lvPPA, showed the greatest aphasia severity as measured by the WAB

Table 1 Demographic and speech and language data by primary progressive aphasia (PPA) classification

	agPPA (n = 23)	svPPA (n = 6)	lvPPA (n = 31)	Unclassified (n = 26)	PPAOS (controls) (n = 21)	p Value
N (%)	23 (27)	6 (7)	29 (35)	26 (30)	21	
% Male	70	33	45	50	29	NS
Age at onset, y	66.0 (61.0–69.5)	59.0 (52.3–66.5)	65.0 (56.0–70.0)	65.0 (59.3–68.8)	71.0 (63.0–75.0)	NS
Duration, y	3.5 (2.5–4.3)	3.8 (3.5–4.8)	3.0 (2.3–5.0)	2.5 (2.0–3.9)	3.0 (2.5–4.3)	NS
Education, y	16 (13–17)	16 (14–18)	16 (14–18)	14 (12–18)	15 (13–17)	NS
Western Aphasia Battery						
Aphasia quotient (/100)	84.1 (74.7–88.6) [‡]	75.1 (72.7–88.7) [‡]	81.6 (72.6–85.6) ^{‡,b}	90.8 (85.8–94.6) [‡]	96.7 (95.8–98.7)	<0.0001
Informational content (/10)	9.0 (9.0–9.0) [‡]	8.5 (7.3–9.8) [‡]	8.0 (7.0–9.0) [‡]	9.0 (8.0–10.0) [‡]	10.0 (10.0–10.0)	<0.0001
Fluency (/10)	6.0 (5.0–9.0) [‡]	9.0 (6.8–9.0)	8.0 (7.5–9.0) [‡]	7.8 (3–10) [‡]	10.0 (9.0–10)	<0.0001
Repetition (/10)	8.4 (5.7–9.2) [‡]	9.1 (7.7–9.5)	7.8 (7.0–8.4) ^{‡,b}	9.0 (6.5–9.0)	9.7 (9.4–9.9)	<0.0001
Reading irregular words (/10)	9.0 (8.0–10.0) [‡]	8.5 (7.3–9.0) [‡]	8.0 (6.5–9.0) ^{‡,b}	10.0 (8.3–10.0)	10.0 (10.0–10.0)	<0.0001
Reading nonwords (/10)	6.0 (5.0–8.0) [‡]	8.5 (8.0–9.0)	6.0 (3.5–9.0) [‡]	7.0 (3.5–9.8) [*]	10.0 (9.0–10.0)	0.0001
Writing irregular words (/10)	5.0 (4.0–8.5) [*]	5.0 (2.8–7.3)	3.0 (1.0–5.5) [‡]	6.5 (4.0–8.5) [*]	9.0 (7.0–10.0)	0.0002
Writing nonwords (/10)	4.0 (2.0–5.5) [‡]	4.5 (3.3–5.8) [*]	3.0 (2.0–6.5) [‡]	6.0 (2.5–7.0) [*]	8.0 (6.0–9.0)	0.0001
Auditory recognition (/60)	59.0 (59.0–60.0) [*]	53.0 (47.3–57.3) ^{‡,a,b}	59.0 (57.5–60.0) [*]	60.0 (60.0–60.0)	60.0 (60.0–60.0)	<0.0001
Boston Naming Test (/15)	13.0 (9.0–14.5)	1.5 (0.3–2.8) ^{‡,a,b}	6.0 (2.0–9.5) ^{‡,a,b}	10.0 (6.0–13.8) [‡]	14.0 (13.0–15.0)	<0.0001
PPTT (/52)	50.0 (48.0–51.0)	34.0 (22.5–42.5) ^{‡,a,b}	47.0 (44.0–49.0) ^{‡,a}	48.0 (46.0–49.3) [*]	50.0 (49.0–51.0)	<0.0001
Token Test (/22)	14.0 (9.0–18.5) [‡]	8.0 (6.0–15.0) [*]	9.0 (6.5–14.0) ^{‡,b}	15.5 (10.3–19.0) [‡]	21.0 (19.0–21.0)	<0.0001
AOS rating (/64)	19.0 (10.8–27.8)	0.0 (0.0–0.0) ^{‡,a}	2.0 (0.0–3.0) ^{‡,a}	0.5 (0.0–3.0) ^{‡,a}	18.0 (13.0–20.0)	<0.0001
Dysarthria rating (/10)	6.0 (4.5–8.0)	10.0 (10.0–10.0) ^{‡,a}	10.0 (10.0–10.0) ^{‡,a}	10.0 (10.0–10.0) ^{‡,a}	7.0 (6.0–8.0)	<0.0001
Phonologic errors rating (/4)	0.0 (0.0–1.0)	0.0 (0.0–0.8)	1.0 (1.0–2.0) ^{‡,a}	1.0 (0.0–1.0) [‡]	0.0 (0.0–0.0)	<0.0001

Abbreviations: agPPA = agrammatic variant of PPA; AOS = apraxia of speech; lvPPA = logopenic variant of PPA; NS = nonsignificant; PPAOS = primary progressive apraxia of speech; PPTT = Pyramids and Palm Trees Test; svPPA = semantic variant of PPA.

Data shown as median (interquartile range) except where otherwise noted.

* $p < 0.05$ compared with PPAOS; [‡] $p < 0.01$ compared with PPAOS; [‡] $p < 0.0001$ compared with PPAOS.

^a $p < 0.05$ compared with agPPA.

^b $p < 0.05$ compared with unclassified.

aphasia quotient (WAB-AQ), while unclassified participants were the least severe (table 1). As expected, all PPA variants differed significantly from PPAOS controls on the WAB but did not differ from each other except on the repetition, naming, and reading irregular words subtests, where lvPPA performed worse than unclassified participants.

Performance on the BNT differed across groups, with worse performance in svPPA and lvPPA (table 1). Object knowledge, as measured by the PPTT, was most severely impaired in the svPPA participants, although unclassified and lvPPA groups also showed impairment. All variants showed impairment in verbal comprehension, compared to PPAOS, as measured by the Token Test. AOS and dysarthria were noted most commonly in agPPA (table 1). Phonologic errors were noted mostly in lvPPA participants, with unclassified participants also demonstrating increased phonologic errors compared to controls.

Neurologic and neuropsychological findings. The agPPA participants showed more parkinsonian features than

the svPPA, lvPPA, and unclassified participants (table 2). The svPPA, lvPPA, and unclassified participants performed worse on the MoCA compared to the PPAOS controls. Compared to PPAOS controls, performance on the DKEFS was impaired in agPPA, svPPA, lvPPA, and unclassified participants, while memory was impaired in lvPPA and unclassified participants (table 2). Visuospatial and visuoperceptual tests did not differ across variants except on the letters test, where svPPA and lvPPA participants performed worse than PPAOS controls. The lvPPA participants performed worse on the MoCA and letters test than agPPA.

Unclassified subjects. The diagnostic features of each of the 26 unclassified participants are shown in table e-2. Five participants fulfilled major criteria for agPPA, svPPA, and lvPPA but did not fulfill minor criteria for any of the 3 variants. One additional participant fulfilled major criteria for agPPA only. The remaining 20 participants did not fulfill major criteria for any of the variants.

The hierarchical cluster analysis identified 2 clusters within the unclassified participants. The first

Table 2 Neurologic and neuropsychological profiles of subjects by primary progressive aphasia (PPA) classification

	agPPA (n = 23)	svPPA (n = 6)	lvPPA (n = 29)	Unclassified (n = 26)	PPAOS (controls) (n = 21)	p Value
MoCA (/30)	24.0 (20.0–24.5)*	17.0 (9.3–18.0)*	15.0 (10.5–18.0)* ^{b,c}	22.0 (18.3–24.0)*	28.0 (25.0–29.0)	<0.0001
FBI (/72)	13.0 (8.5–16.0)	15.0 (2.5–20.8)	10.0 (8.0–21.5)	8.5 (4.0–16.8)	6.0 (3.0–12.0)	NS
Limb praxis (/60)	54.0 (45.0–57.5)	53.5 (46.5–57.5)	56.0 (53.0–60)	59.0 (53.3–59.8)	58.0 (51.0–59.0)	0.0355 ^a
MDS-UPDRS (/132)	13.0 (5.5–20.5)	3.0 (0.5–4.0)* ^b	4.0 (2.0–8.5)* ^b	2.0 (0.0–7.5)* ^b	13.0 (5.0–22.0)	<0.0001
NPI (/36)	2.0 (1.0–3.5)	2.0 (2.0–3.5)	2.0 (1.0–3.5)	2.0 (1.0–3.0)	1.0 (0.0–5.0)	NS
Trails A (time), ^e s	51.0 (40.0–67.5)	68.0 (36.0–92.0)* ^{b,c}	59.5 (39.3–93.0) ^d	42.0 (32.3–51.0)	48.0 (43.0–55.0)	0.0005
Trails B (time), ^f s	135.5 (108.0–152.8)	90.0 (79.0–112.5)	168.5 (87.0–205.5)	113.0 (90.5–160.0)	107.5 (72.0–132.0)	NS
DKEFS (scaled score)	8.0 (7.0–10.0)*	6.5 (3.8–7.8)*	6.0 (3.8–7.0)*	8.0 (4.0–11.0)*	11.0 (9.0–13.0)	<0.0001
WMS-LM % recall	84.5 (70.8–97.8)	46.5 (21.3–68.8)	50.0 (22.5–88.0)	77.0 (56.0–92.0)	89.0 (74.0–90.5)	0.0218 ^a
WMS-VR % recall	65.0 (36.5–84.5)	42.0 (12.5–55.0)	37.0 (4.0–64.0)	37.0 (19.3–59.3)*	74.0 (56.0–83.0)	0.0111
AVLT % recall	70.0 (50.0–89.0)	67.0 (33.0–92.0)	0.0 (0.0–63.0)* ^b	60.0 (18.8–80.5)*	89.0 (75.0–92.0)	0.0001
VOSP Letters raw (/20)	20.0 (19.0–20.0)	13.5 (5.0–18.3)*	19.0 (18.0–19.0)* ^b	19.0 (18.3–20.0)	20.0 (20.0–20.0)	0.0004
VOSP Cube raw (/10)	9.0 (6.5–9.5)	7.0 (3.8–8.8)	9.0 (5.0–10.0)	9.0 (9.0–10)	10.0 (9.0–10.0)	0.0402 ^a
Rey-O drawing (/36)	23.8 (19.1–30.9)	20.5 (8.1–25.8)	20.0 (5.5–29.3)	27.0 (15.0–31.0)	29.0 (22.5–31.0)	NS

Abbreviations: agPPA = agrammatic variant of PPA; AVLT = Auditory Verbal Learning Test; DKEFS = Delis-Kaplan Executive Function System Card Sort; FBI = Frontal Behavioral Inventory; LM = logical memory; lvPPA = logopenic variant of PPA; MDS-UPDRS = Movement Disorders Society-sponsored version of the Unified Parkinson's Disease Rating Scale part III; MoCA = Montreal Cognitive Assessment battery; NPI = brief questionnaire form of Neuropsychiatric Inventory; NS = nonsignificant; PPAOS = primary progressive apraxia of speech; Rey-O = Rey-Osterrieth Complex Figure Test; svPPA = semantic variant of PPA; VOSP = Visual Object and Space Perception; VR = visual reproduction; WMS = Wechsler Memory Scale.

Data shown as median (interquartile range).

* $p < 0.05$ compared with PPAOS; † $p < 0.01$ compared with PPAOS; ‡ $p < 0.0001$ compared with PPAOS.

^a Omnibus significant but comparisons nonsignificant.

^b $p < 0.05$ compared with agPPA.

^c $p < 0.05$ compared with unclassified.

^d $p < 0.05$ compared with svPPA.

^e Trails A was discontinued or not administered in 1 agPPA, 1 svPPA, 1 lvPPA, and 2 unclassified participants as the participants were unable to comprehend the task.

^f Trails B was discontinued or not administered in 3 agPPA, 3 svPPA, 14 lvPPA, and 4 unclassified participants as the participants were unable to comprehend the task.

cluster consisted of 12 participants with features in the direction of lvPPA (lvPPA-like), with trends toward more phonologic errors (table 3). The second cluster of 8 participants had features resembling those of svPPA (svPPA-like), with trends toward greater difficulty with naming and impaired object knowledge (table 3). The 2 clusters did not differ in demographics, behavioral, or neuropsychiatric measures.

Imaging findings. The agPPA participants showed gray matter loss in left greater than right frontal lobe, particularly inferior and posterior frontal and insular regions (figure). The svPPA participants showed loss localized to left anterior temporal lobe, while lvPPA participants showed loss throughout the left lateral temporal lobe, with additional involvement of left parietal, occipital, and frontal lobes, and right temporal lobe (figure). The unclassified participants in cluster 1 showed gray matter loss in left lateral temporal lobe, similar to lvPPA, although less widespread. The unclassified participants in cluster 2 showed loss in left anterior temporal lobe, similar to svPPA, although

less severe (figure). The unclustered subjects showed prominent left-sided posterior inferior frontal and temporal atrophy. Findings in agPPA and lvPPA survived correction for multiple comparisons.

DISCUSSION Our findings demonstrate that quantitative application of the diagnostic criteria for PPA only classified 69% of PPA participants into the 3 syndromic variants, hence leaving a large proportion of participants unclassified.

One postulate for the large number of unclassified participants may be the insensitivity of some of the administered tests, particularly since the majority of the unclassified participants were less severe than those classified. For example, the Auditory Word Recognition subtest of the WAB that was utilized to assess single word comprehension and the Informational Content subtest of the WAB that was used as a measure of impairment of single-word retrieval in spontaneous speech were usually associated with perfect scores. Our selected tests, however, satisfy the qualitative test descriptions provided by the

Table 3 Characteristics of unclassified primary progressive aphasia (PPA) participants by clusters

	Cluster 1 lvPPA-like	Cluster 2 svPPA-like	Unclassified	p Value
N (%)	12 (46)	8 (31)	6 (23)	
% Male	58	50	33	NS
Age at onset, y	65.0 (62.0–68.3)	64.0 (56.9–69.8)	66.0 (61.0–68.0)	NS
Duration, y	2.5 (2.0–3.5)	2.0 (1.5–3.3)	3.3 (2.3–3.9)	NS
Education, y	15 (13–17)	19 (16–22)	12 (12–12) ^{b,c}	0.0020
Speech and language				
Western Aphasia Battery				
Aphasia quotient (/100)	90.8 (89.5–94.5)	94.8 (92.9–98.5)	55.2 (47.5–57.8) ^{b,c}	0.0004
Informational content (/10)	9.5 (8.0–10.0)	10.0 (9.0–10.0)	5.5 (4.3–6.8) ^{b,c}	0.0021
Fluency (/10)	9.0 (8.8–9.0) ^c	9.0 (9.0–10.0) ^b	4.5 (4.0–5.0) ^{b,c}	<0.0001
Repetition (/10)	9.2 (9.0–9.7)	9.7 (9.6–9.9)	5.8 (4.4–6.8) ^{b,c}	0.0028
Reading irregular words (/10)	9.5 (9.0–10.0)	10.0 (10.0–10.0)	7.5 (4.0–9.5)	0.0496 ^a
Reading nonwords (/10)	6.5 (4.5–8.3) ^c	10.0 (9.8–10.0) ^b	2.5 (0.5–4.5) ^c	0.0007
Writing irregular words (/10)	5.5 (4.0–7.0) ^c	10.0 (8.5–10.0) ^b	0.5 (0.0–1.0) ^{b,c}	0.0004
Writing nonwords (/10)	6.0 (4.0–7.0) ^c	7.5 (7.0–9.0) ^b	0.0 (0.0–0.0) ^{b,c}	0.0003
Auditory recognition (/60)	60.0 (60.0–60.0)	60.0 (60.0–60.0)	52.5 (49.8–54.5) ^{b,c}	0.0002
Boston Naming Test (/15)	13.0 (9.8–14.0)	10.5 (5.3–15.0)	4.5 (1.8–5.8) ^b	0.0249
PPTT (/52)	49.0 (47.5–50.3)	47.5 (47.0–49.5)	39.0 (33.5–44.0) ^b	0.0279
Token Test (/22)	14.0 (12.5–16.0) ^c	20.0 (19.5–21.3) ^b	2.0 (1.3–5.0) ^c	0.0002
AOS rating (/64)	0.5 (0.0–2.0)	0.0 (0.0–0.0)	5.0 (5.0–5.8) ^{b,c}	0.0004
Dysarthria rating (/10)	10.0 (10.0–10.0)	10.0 (10.0–10.0)	10.0 (7.8–10.0)	0.0312 ^a
Phonologic errors rating (/4)	1.0 (1.0–1.0) ^c	0.0 (0.0–0.3) ^b	1.5 (1.0–2.0)	0.0066
Neurologic				
MoCA (/30)	22.0 (19.8–23.0) ^c	24.5 (24.0–27.0) ^b	5.5 (4.0–9.3) ^{b,c}	0.0002
FBI (/72)	8.5 (6.3–13.3)	4.0 (2.8–7.5)	18.0 (17.0–28.8) ^c	0.0075
Limb praxis (/60)	59.0 (57.0–60.0)	59.0 (59.0–59.3)	43.0 (35.8–46.5) ^{b,c}	0.0018
MDS-UPDRS (/132)	1.5 (0.0–5.5)	1.0 (0.0–2.0)	11.5 (5.3–19.3) ^c	0.0157
NPI (/36)	1.5 (1.0–3.0)	2.0 (1.0–4.0)	2.0 (1.3–2.0)	NS
Neuropsychological				
Trails A (time), ^d s	45.5 (33.8–59.0)	35.0 (28.5–41.5)	67.5 (58.8–76.3)	NS
Trails B (time), ^e s	134.0 (104–160)	105 (70.3–124.8)	221.0 (181.5–260.5)	NS
DKEFS (scaled score)	8.0 (6.3–9.3)	11.0 (7.0–12.3)	3.0 (2.5–3.5)	0.0390 ^a
WMS-LM % recall	60.0 (45.0–75.0) ^c	93.0 (82.3–100.0) ^b	50.0 (25.0–75.0)	0.0386
WMS-VR % recall	36.5 (19.3–52.0)	64.5 (53.5–83.5)	12.0 (0.0–26.8)	0.0101 ^a
AVLT % recall	46.5 (0.0–63.8)	80.0 (73.5–89.5)	25.0 (0.0–60.0)	NS
VOSP Letters raw (/20)	20.0 (19.0–20.0)	20.0 (19.0–20.0)	15.0 (10.8–17.0) ^{b,c}	0.0006
VOSP Cube raw (/10)	9.5 (9.0–10.0)	9.5 (9.0–10.0)	7.0 (4.0–9.0) ^c	0.0332
Rey-O drawing (/36)	26.0 (23.5–31.5) ^c	32.5 (30.9–34.3) ^b	1.5 (1.5–4.5) ^{b,c}	0.0011

Abbreviations: AOS = apraxia of speech; AVLT = Auditory Verbal Learning Test; DKEFS = Delis-Kaplan Executive Function System Card Sort; FBI = Frontal Behavioral Inventory; LM = logical memory; lvPPA = logopenic variant of PPA; MDS-UPDRS = Movement Disorders Society-sponsored version of the Unified Parkinson's Disease Rating Scale part III; MoCA = Montreal Cognitive Assessment battery; NPI = brief questionnaire form of Neuropsychiatric Inventory; NS = nonsignificant; PPTT = Pyramids and Palm Trees Test; Rey-O = Rey-Osterrieth Complex Figure Test; svPPA = semantic variant of PPA; VOSP = Visual Object and Space Perception; VR = visual reproduction; WMS = Wechsler Memory Scale.

Data shown as median (interquartile range) unless otherwise noted.

^aOmnibus significant but comparisons nonsignificant.

^b $p < 0.05$ compared with cluster 1.

^c $p < 0.05$ compared with cluster 2.

^dTrails A was discontinued or not administered in 2 unclassified participants as the participants were unable to comprehend the task.

^eTrails B was discontinued or not administered in 4 unclassified participants as the participants were unable to comprehend the task.

Figure Patterns of gray matter loss in each primary progressive aphasia (PPA) variant compared to healthy controls

Results are shown uncorrected for multiple comparisons at $p < 0.001$. Renders were generated using the BrainNet Viewer (<http://www.nitrc.org/projects/bnv/>). agPPA = agrammatic variant of PPA; lvPPA = logopenic variant of PPA; svPPA = semantic variant of PPA.

consensus criteria.² The proposed PPA criteria do not stipulate specific tests to use in classifying variants, despite their encouragement to use quantitative data. It may not be an issue of simply needing more sensitive tests, however, since the 6 unclassified and unclustered participants, as a group, were slightly more severe than those classified as agPPA, svPPA, and lvPPA. In addition, a recent study utilizing a different battery of tests was also unable to classify a large proportion of their participants.⁶ Nevertheless, although the majority of tests in this study are well-recognized and widely utilized, we would not recommend that they are considered a gold standard without modification to the criteria, given the less than desirable sensitivity and specificity.

Another postulate for the large unclassifiable group in this study was the relatively short disease duration of some of our participants. It has been

suggested that a diagnosis of PPA be made when aphasia has been present for at least 2 years.²⁴ Some of our participants had 1-year disease duration. However, mean disease duration was greater than 2 years for all groups and did not differ between the unclassified and classified participants, suggesting that disease duration was unlikely the main driver of misclassification. Furthermore, one study was able to classify 80% of subjects with early or mild disease.²⁵ Our cohort, although highly educated, did not differ in education level across groups and hence high education is another unlikely explanation for the unclassifiable participants. Although a quantitative approach will inevitably fail to capture qualitative features that expert clinicians render, qualitative application of the criteria, while likely more sensitive, does not allow for study reproducibility, comparison of findings across studies, or longitudinal analyses.

Therefore, if the criteria are to be useful in these regards, modifications are necessary to provide greater sensitivity; a standardized battery of tests with high sensitivity and specificity should be proposed or developed.

Detailed examination of the unclassified participants revealed 2 groups: a group resembling lvPPA and another resembling svPPA. Therefore, it appears that the consensus criteria amply capture agPPA as previously reported,⁶ although there are other issues regarding this variant that will be discussed later on. It thus appears that it is mostly the milder cases of lvPPA and svPPA that are not being captured by the criteria. The lvPPA-like participants who were not classified often had phonologic errors and spared object knowledge without impaired word retrieval or anomia. The svPPA-like participants who were not classified had anomia without evidence of loss of word meaning. In one study, analysis of early aphasic participants showed that follow-up evaluation of isolated anomic subjects may later lead to svPPA classification,²⁵ suggesting that our svPPA-like participants in fact may represent early svPPA. The imaging findings support this contention. Although the remaining 6 unclassified unclustered participants did not form a homogenous group, they more often had severe global aphasia, satisfied major criteria for all variants, and were more likely to have cognitive deficits in other nonlanguage domains at enrollment. All 6, however, had presented with aphasia as their chief complaint and it was the aphasia that caused impairment in activities in daily life. Hence these subjects meet criteria for PPA²⁴ and did not have dementia with an embedded pronounced aphasia. It is possible that some of the unclustered participants were not clustered due to the relatively small sample size, although a larger sample size could have yielded an even higher percentage of unclustered participants.

Two studies have now demonstrated that a large number of subjects cannot be classified with the consensus criteria.⁶ In both studies the main issue with classification seems to arise predominantly with the lvPPA group. And although in our study we were able to identify participants who could be classified as lvPPA using the consensus criteria, it appears that the criterion for lvPPA is problematic and harsh. Sensitivity would be improved if only one of the core features is required for diagnosis, and further improved if the first core feature required only one of either impaired single word retrieval in spontaneous speech or impaired naming, instead of both. To capture the more severe cases, a modification to the minor features is needed. It must be noted that although we were able to classify a large group of participants as agPPA using the consensus criteria, in order to do so we had to ignore the original PPA

criterion, which states that aphasia has to be the most prominent feature of the presenting syndrome.²⁴ It is important to note that 57% of the participants classified as agPPA in fact had AOS as the dominant symptom of the presenting syndrome, with aphasia being less prominent. Therefore, many participants classified as agPPA in this study would in fact not meet PPA, and hence agPPA, criteria if the consensus recommendations had been strictly adhered to. In fact, we have previously proposed separating out such patients and instead classifying them as progressive AOS,⁵ not as PPA. Another concern with this consensus agPPA group is the fact that the imaging findings were more bilateral than expected. This requires further investigation.

All 84 participants in our study had variable degrees of aphasia. Although some participants had a WAB-AQ greater than the standard cutoff of 94.9, which reliably distinguishes normal from aphasic subjects,²⁶ there can be overlap in the WAB-AQ in high-functioning aphasic subjects and normal controls.²⁷ However, many of the subjects with high WAB-AQs did have impaired sentence comprehension or naming and hence demonstrated specific language impairments that allowed for classification first as PPA²⁴ and secondly as a PPA variant.²

Quantitative application of the current PPA criteria yielded a large group of unclassifiable participants and although we considered different explanations for this problem, it appears that the current consensus criteria need to be modified in order to be applicable to more PPA subjects. Our participants were recruited from all over the United States and hence the findings should be generalizable. This article represents one step in the iterative process between initial diagnostic guidelines and their subsequent quantitative and qualitative application that will lead to guideline refinements based on the gathered data.

AUTHOR CONTRIBUTIONS

Dr. Wicklund: drafting/revising the manuscript for content, study concept or design, analysis or interpretation of the data, statistical analysis. Dr. Duffy: drafting/revising the manuscript for content, study concept or design, analysis or interpretation of the data, acquisition of data. Dr. Strand: drafting/revising the manuscript for content, analysis or interpretation of the data, acquisition of data. Dr. Machulda: drafting/revising the manuscript for content, analysis or interpretation of the data, acquisition of data. Dr. Whitwell: drafting/revising the manuscript for content, analysis or interpretation of the data, acquisition of data. Dr. Josephs: drafting/revising the manuscript for content, study concept or design, analysis or interpretation of the data, acquisition of data, study supervision, obtaining funding.

ACKNOWLEDGMENT

The authors thank Drs. Ahlskog, Boeve, Drubach, Keegan, Knopman, and Petersen for subject referral; Sarah Boland for performing the neuropsychometric testing and organizing all subject test schedules; and Dr. Jack for providing image analysis tools.

STUDY FUNDING

Supported by NIH R01-DC010367.

DISCLOSURE

M. Wicklund reports no disclosures. J. Duffy is a coinvestigator for R01-DC012519 and R01-DC010367. E. Strand is a coinvestigator for R01-DC012519 and R01-DC010367. M. Machulda is funded by R01-AG037491 (Co-I), R01-DC012519 (Co-I), and the Alzheimer's Association (Co-I). J. Whitwell is funded by R01-DC012519 (PI), R01-DC010367 (Co-I), R01-AG037491 (Co-I), and the Alzheimer's Association (PI). K. Josephs is funded by NIH R01-DC010367 (PI), R01-AG037491 (PI), R01-DC012519 (Co-I), and the Alzheimer's Association (Co-I). Go to Neurology.org for full disclosures.

Received July 5, 2013. Accepted in final form November 19, 2013.

REFERENCES

1. Mesulam MM. Slowly progressive aphasia without generalized dementia. *Ann Neurol* 1982;11:592–598.
2. Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. *Neurology* 2011;76:1006–1014.
3. Mesulam MM, Wieneke C, Hurley R, et al. Words and objects at the tip of the left temporal lobe in primary progressive aphasia. *Brain* 2013;136:601–618.
4. Josephs KA, Duffy JR, Strand EA, et al. Characterizing a neurodegenerative syndrome: primary progressive apraxia of speech. *Brain* 2012;135:1522–1536.
5. Josephs KA, Duffy JR, Strand EA, et al. Syndromes dominated by apraxia of speech show distinct characteristics from agrammatic PPA. *Neurology* 2013;81:337–345.
6. Sajjadi SA, Patterson K, Arnold RJ, Watson PC, Nestor PJ. Primary progressive aphasia: a tale of two syndromes and the rest. *Neurology* 2012;78:1670–1677.
7. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. 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;34:939–944.
8. Kertesz A. Western Aphasia Battery (Revised), San Antonio: PsychCorp; 2007.
9. De Renzi E, Vignolo LA. The Token Test: a sensitive test to detect receptive disturbances in aphasics. *Brain* 1962;85:665–678.
10. Lansing AE, Ivnik RJ, Cullum CM, Randolph C. An empirically derived short form of the Boston Naming Test. *Arch Clin Neuropsychol* 1999;14:481–487.
11. Howard D, Patterson K. The Pyramids and Palm Trees Test: A Test of Semantic Access from Words and Picture. Bury St. Edmunds: Thames Valley Test Company; 1992.
12. Nasreddine ZS, Phillips NA, Bedirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005;53:695–699.
13. Kertesz A, Davidson W, Fox H. Frontal behavioral inventory: diagnostic criteria for frontal lobe dementia. *Can J Neurol Sci* 1997;24:29–36.
14. Kaufer DI, Cummings JL, Ketchel P, et al. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *J Neuropsychiatry Clin Neurosci* 2000;12:233–239.
15. Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord* 2008;23:2129–2170.
16. Spreen O, Strauss E. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. New York: Oxford University Press; 1991.
17. Delis D, Kaplan E, Kramer J. Delis-Kaplan Executive Function System (DKEFS): Examiner's Manual. San Antonio: The Psychological Corporation; 2001.
18. Wechsler D. Wechsler Memory Scale-III. New York: Psychological Corporation; 1997.
19. Rey A. L'Examen Clinique En Psychologie. Paris: Presses Universitaires de France; 1964.
20. Warrington E, James M. The Visual Object and Space Perception Battery. Bury St Edmunds: Thames Valley Test Company; 1991.
21. Osterrieth PA. Le test de copie d'une figure complexe. *Arch de Psychol* 1944;30:206–356.
22. Ashburner J, Friston KJ. Voxel-based morphometry: the methods. *Neuroimage* 2000;11:805–821.
23. Critchlow D, Fligner M. On distribution-free multiple comparisons in the one-way analysis of variance. *Commun Stat Theor Methods* 1991;20:127–139.
24. Mesulam MM. Primary progressive aphasia. *Ann Neurol* 2001;49:425–432.
25. Mesulam MM, Wieneke C, Thompson C, Rogalski E, Weintraub S. Quantitative classification of primary progressive aphasia at early and mild impairment stages. *Brain* 2012;135:1537–1553.
26. Shewan CM, Kertesz A. Reliability and validity characteristics of the Western Aphasia Battery (WAB). *J Speech Hear Disord* 1980;45:308–324.
27. Kertesz A, Poole E. The aphasia quotient: the taxonomic approach to measurement of aphasic disability: 1974. *Can J Neurol Sci* 2004;31:175–184.