

Original Research Article

Screening for Aphasia in NeuroDegeneration for the Diagnosis of Patients with Primary Progressive Aphasia: Clinical Validity and Psychometric Properties

Petronilla Battista^{a–c} Eleonora Catricalà^d Marco Piccininni^{a, b}
Massimiliano Copetti^e Valentina Esposito^{f, g} Cristina Polito^c
Antonio Miozzo^h Elena Gobbi^{g, l} Sofia Cuocoⁱ Veronica Boschi^d
Marina Picilloⁱ Sandro Sorbi^{c, j} Paolo Baroneⁱ Sandro Iannaccone^f
Peter Garrard^k Giancarlo Logroscino^{a, b} Stefano F. Cappa^{d, l}

^aDepartment of Basic Medical Sciences, Neuroscience, and Sense Organs, University of Bari Aldo Moro, Bari, Italy; ^bNeurodegenerative Disease Unit, Department of Clinical Research in Neurology, University of Bari Aldo Moro, “Pia Fondazione Cardinale G. Panico”, Tricase, Italy; ^cNEUROFARBA – Dipartimento di Neuroscienze, Psicologia, Area del Farmaco e Salute del Bambino, Università di Firenze, Florence, Italy; ^dCenter for Neurocognition and Theoretical Syntax, Scuola Universitaria Superiore IUSS Pavia, Pavia, Italy; ^eBiostatistics Unit, IRCCS-Ospedale Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy; ^fDepartment of Specialistic Rehabilitation of Neurological, Cognitive, and Motor Disorders, IRCCS San Raffaele Hospital and Vita-Salute University, Milan, Italy; ^gUniversità Vita-Salute San Raffaele, Milan, Italy; ^hDepartment of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy; ⁱDipartimento di Medicina e Chirurgia, Università di Salerno, Salerno, Italy; ^jIRCCS Don Gnocchi Firenze, Florence, Italy; ^kNeuroscience Research Center, St George’s University of London, London, UK; ^lIRCCS S. Giovanni di Dio Fatebenefratelli, Brescia, Italy

Keywords

Language impairment · Screening instrument · Dementia · Diagnostic accuracy · Speech-language assessment

Abstract

Background: We evaluated the psychometric proprieties of the Screening for Aphasia in NeuroDegeneration (SAND) battery in Italian primary progressive aphasia (PPA) and movement disorder (MD) patients. **Methods:** The sample included 30 consecutive PPA and 45 MD patients who completed the SAND battery together with a clinical interview and a neurological/neuropsychological examination and 130 healthy controls (HC). **Results:** The SAND bat-

Prof. Stefano F. Cappa
Institute for Advanced Study-I USS Pavia, Palazzo del Broletto
Piazza Vittoria 15
IT-27100 Pavia (Italy)
E-Mail stefano.cappa@iusspavia.it

tery showed good internal consistency and good convergent and divergent validity. Receiver operating characteristic analysis revealed an area under the curve of 0.978 for PPA versus HC and of 0.786 for PPA versus MD. A cutoff ≥ 3 gave a sensitivity of 0.933% and a specificity of 0.946% for discriminating PPA versus HC, whereas a cutoff ≥ 5 gave a sensitivity of 0.767% and a specificity of 0.667% for discriminating PPA versus MD. **Conclusion:** These results indicate that the SAND battery is an adequate, reliable, and valid diagnostic tool for PPA.

© 2018 S. Karger AG, Basel

Introduction

The language network can be selectively affected by neurodegeneration, leading to progressive language dysfunction (primary progressive aphasia [PPA]). After the seminal description of slowly progressive aphasia by Mesulam and Weintraub [1], many studies have been dedicated to this heterogeneous group of clinical conditions, due to different pathologies in which language impairment is the main sign at onset, and remains the prominent clinical feature during many years of progression. The consensus paper on the definition of PPA and its clinical subtypes recognizes three different variants: nonfluent/agrammatic PPA (nf/a-PPA), semantic variant PPA, and logopenic PPA [2]. Each of these variants presents with a different pattern of speech and language deficits and brain imaging features [3]. The diagnosis of PPA is predominantly based on clinical examination, including detailed history from patient and informant, and neurological and neuropsychological evaluation. Most language tests in common use, such as the Aachen Aphasia Test [4] and the Boston Diagnostic Aphasia Examination [5], have been developed for the evaluation of aphasia due to stroke, which, besides having a different prognosis, is characterized by language profiles that can be very different from those of PPA [6]. Given the dearth of specific tools aimed at assessing language impairments in neurodegenerative diseases [7], we developed the Screening for Aphasia in NeuroDegeneration (SAND) battery [8], based on the Mini Standard Language Examination proposed by Garrard and Ahmed [9]. The SAND battery aims at the detection of language disorders through the assessment of different components of language. It includes nine subtests: picture naming, word and sentence comprehension, word and sentence repetition, reading, semantic association, writing, and picture description. It has been developed on the basis of the recommendations of current diagnostic guidelines [2] and of a comprehensive review of the language deficits in PPA subtypes. The claim that the SAND battery is adequate to detect language disorders requires the assessment of its construct, clinical validity, and diagnostic accuracy. To this aim, we studied a group of participants with PPA, a group of healthy controls (HC), and a group of patients with movement disorders (MDs) (Parkinson's disease [PD] and progressive supranuclear palsy [PSP]). We expected the PPA patients to be the most impaired group, as aphasia is the major symptom in these patients. Healthy individuals were expected to perform close to ceiling. In the case of MD, PD patients were expected to show normal or mildly affected language performance [10]. A similar prediction could be made for PSP patients, since those presenting with a clinical picture of nf/a-PPA were excluded. The psychometric properties of the SAND battery and the discriminatory power in detecting PPA language dysfunction were assessed by comparing PPA with HC and with MD, estimating area under the curve (AUC), sensitivity, and specificity.

Subjects and Methods

Participants

Two patient samples were included in the study: (1) Patients with a clinical diagnosis of PPA according to published guidelines [2], based on audiotapes of language and cognitive testing, history, and review of imaging (magnetic resonance imaging or positron emission tomography). Those patients not fulfilling the diagnostic criteria for a specific variant were assigned the label “unclassifiable.” These patients did not present the typical features of any variant, but only anomia, or were characterized by features of more than one variant (e.g., nonfluent and logopenic). (2) Patients with a clinical diagnosis of PD or PSP. Patients with PD [11] and with probable or possible PSP [12] were clinically diagnosed. Motor impairment severity was assessed using the Unified Parkinson’s Disease Rating Scale [13] for PD patients and Natural History and Neuroprotection in Parkinson Plus Syndromes [14] for PSP patients. All patients were recruited in an MD unit. Features of language impairment could be present, but were not the primary cause of referral for any of these patients. The SAND data were collected as part of a comprehensive study of cognition in MD.

Enrollment started in May 2015 and ended in February 2017. Participants were selected among the outpatients of hospitals located in different Italian areas (Milan, Florence, Bari, and Salerno).

The criteria for inclusion were (1) availability of a audiotaped language examinations in order to allow an offline analysis of connected speech, (2) a Mini-Mental State Examination (MMSE) [15] score of at least 10, (3) Italian native speaker, (4) sufficiently intelligible speech such that the intended target could be determined for the majority of words, (5) intact or corrected auditory and visual functions, and (6) successful completion of the experimental task.

Patients were excluded in the case of (1) major psychiatric disorders [16], (2) organic illness affecting the brain according to the International Classification of Diseases [17], (3) significant history of head injury, (4) major systemic diseases or medical complications, including thyroid disorders and sensory disorders (i.e., blindness or deafness), and (5) history of drug or alcohol addiction.

All study participants underwent extensive studies at baseline, following a diagnostic protocol that included medical, neurological, neuropsychological, and neuroimaging investigations (with magnetic resonance imaging or positron emission tomography). Further examinations, such as cerebrospinal fluid markers (β -amyloid and tau protein concentration), were obtained for a small group of patients in order to improve the accuracy of the diagnosis. Neurologists and neuropsychologists with experience in cognitive disorders and/or MDs examined all patients.

A group of HC was enrolled from a convenience sample of volunteers recruited in the centers included in the study. The group of HC was the same recruited for the collection of normative data. Subjects who reported a history of neurological or psychiatric illnesses or with a corrected MMSE score of <24 were excluded [15].

A total of 92 patients were recruited (43 patients with PPA and 49 patients with MD). We excluded 13 patients with PPA because they were too severely affected to perform a full neuropsychological assessment, 3 patients with MD (2 PSP and 1 PD) who fulfilled the criteria for dementia, and 1 PSP patient with nf/a-PPA. Therefore, the global sample (75 patients, 43 males and 32 females, age range 42–85 years) included three different diagnostic groups: 30 PPA patients, including 8 with semantic variant PPA, 12 with logopenic PPA, 6 with nf/a-PPA, 4 with unclassifiable PPA, and 45 MD patients (24 with PD and 21 with PSP). The HC group included 130 native Italian speakers (54 males), aged >45 years, both sexes with an educational level of >1 year, with a mean age of 63.30 years (standard deviation [SD] = 11.30, range 45–85) and a mean of 10.5 years of education (SD = 4.89, range 2–25).

Data on the demographic and clinical features of all patients and HC included in the study are reported in Table 1.

SAND Procedure

All participants were given the SAND battery, whose normative values have already been published [8]. The entire battery takes <20 min to administer and yields scores for each of the nine subtests. Picture description and written description analysis yields additional subscores, resulting in a total of 23 task-related scores (online suppl. Fig. 1S; for all online suppl. material, see www.karger.com/doi/10.1159/000492632).

Cognitive Assessment

In addition to the SAND battery, PPA patients were administered an extensive battery of standardized neuropsychological tests in a 2-h session. The tests covered the following cognitive domains: attention,

Table 1. Demographic characteristics and clinical data of enrolled subjects

Variables	HC	PPA	MD (PD + PSP)	<i>p</i>	PPA vs. HC	MD vs. HC	PPA vs. MD
Sample size	130	30	45 (24+21)	–	–	–	–
Age, years (mean ± SD)	63.30±11.30	70.9±6.04	66.98±8.05	<0.001	<0.001	0.127	0.196
Education, years (mean ± SD)	10.92±4.90	11.67±4.94	11.11±4.90	0.700	–	–	–
Male sex	41.54%	53.33%	60.00%	0.078	–	–	–
Disease duration, months (mean ± SD)	–	31.63±17.62	54.36±37.97	0.007	–	–	–
MMSE score (mean ± SD)	28.49±1.52	20.48±6.50	25.04±4.81	<0.001	<0.001	<0.001	0.022
SAND global score (mean ± SD)	0.65±0.86	7.87±3.87	3.82±3.50	<0.001	<0.001	<0.001	0.011
UPDRS for PD	–	–	15.80±7.97	–	–	–	–
NNIPPS for PSP	–	–	87.15±38.16	–	–	–	–

The *p* values for the overall differences test and for the post hoc analysis (pairwise comparisons: PPA vs. HC, MD vs. HC, and PPA vs. MD) are reported. HC, healthy controls; MD, movement disorder; MMSE, Mini-Mental State Examination; NNIPPS, Natural History and Neuroprotection in Parkinson Plus Syndromes; PD, Parkinson’s disease; PPA, primary progressive aphasia; PSP, progressive supranuclear palsy; SAND, Screening for Aphasia in NeuroDegeneration; SD, standard deviation; UPDRS, Unified Parkinson’s Disease Rating Scale.

Table 2. Correlation analysis among the subscores of the SAND battery and the state of the art language tests

SAND tasks	Language tests	Spearman	<i>p</i> value	Sample size
Naming	CAGI naming	0.82**	<0.001	28
Sentence comprehension	token test	0.63**	<0.001	24
	auditory sentence comprehension (ENPA)	0.45*	0.007	17
Word comprehension	CAGI word-picture matching	0.67**	<0.001	25
Word/nonword repetition	word/nonword repetition (ENPA)	0.54**	0.009	22
Sentence repetition	sentence repetition (AAT)	0.74**	0.002	14
	total score repetition (AAT)	0.84**	<0.001	16
	sentence repetition (ENPA)	0.85**	<0.001	22
Reading	TIB	–0.59**	0.012	17
Writing – syntactic structure/sentences	sentence writing (ENPA)	0.49*	0.020	22
Semantic association	PPT	0.50*	0.040	17
Picture description – nouns/words	category fluency	0.55**	0.003	26
	CAGI naming	0.62**	<0.001	28

Negative values refer to those tests for which the higher the score the worse the performance. * |0.30| < *r* < |0.50|, moderate correlation; ** *r* > |0.50|, strong correlation. The following correlations were performed: naming with CAGI naming; auditory sentence comprehension with ENPA auditory sentence comprehension; word comprehension with CAGI word-picture matching; word/nonword repetition with ENPA word/nonword repetition; sentence repetition with AAT sentence repetition, AAT total score repetition, and ENPA sentence repetition; reading with TIB total errors score; writing – syntactic structure/sentences with ENPA sentence writing; semantic association with PPT total score; picture description – nouns/words with category fluency and CAGI naming. AAT, Aachen Aphasia Test; ENPA, Esame Neuropsicologico dell’Afasia; PPT, Pyramids and Palm Tree Test; SAND, Screening for Aphasia in NeuroDegeneration; TIB, Test di intelligenza breve (brief intelligence test).

Table 3. Correlation analysis among the subscores of the SAND battery and the nonlanguage tests

Nonlanguage tests	Spearman	p value	Sample size
<i>SAND global score</i>			
MMSE	-0.56**	0.001	29
RAVLT immediate	-0.44*	0.059	19
RAVLT recall	-0.24*	0.319	19
RAVLT recognition	-0.51**	0.073	13
RAVLT false alarms	0.34*	0.304	11
Digit span backward	-0.40*	0.034	28
Digit span forward	-0.29*	0.144	27
Corsi span backward	-0.19*	0.341	27
RCF copy	-0.12*	0.578	25
RCF recall	-0.36*	0.084	24
SCWT time	0.30*	0.323	13
SCWT errors	0.34*	0.251	13
TMT-A	0.34*	0.108	23
TMT-B	0.49*	0.060	15
TMT-AB	0.47*	0.076	15
FAB	-0.33*	0.132	22
CDT	-0.13*	0.585	19

Negative values refer to those tests for which the higher the score the worse the performance. * $|0.30| < r < |0.50|$, moderate correlation; ** $r > |0.50|$, strong correlation. The SAND global score was correlated with the following tests: MMSE, RAVLT (immediate, delayed recall, recognition, and false alarms), digit span (backward and forward), Corsi span backward, RCF (copy and delayed recall), SCWT (time and errors, TMT-AB, FAB, and CDT. CDT, Clock Drawing Test; FAB, Frontal Assessment Battery; MMSE, Mini-Mental State Examination; RAVLT, Rey Auditory Verbal Learning Test; RCF, Rey Complex Figure; SAND, Screening for Aphasia in NeuroDegeneration; SCWT, Stroop Color-Word Test; TMT-A, Trail-Making Test part A; TMT-B, Trail-Making Test part B; TMT-AB, Trail-Making Test parts A and B.

executive functions, memory, language, and visuospatial processing (online suppl. material). For each test score, we report the corresponding adjusted score obtained on the basis of the normative data (i.e., according to age and education corrections).

Statistical Analysis

Demographics and Clinical Variables. Quantitative variables are reported as mean, SD, median, and range, while qualitative variables are reported as frequencies and percentages. Comparison among the diagnostic groups for demographic and clinical variables was made using the Kruskal-Wallis test by ranks and the Nemenyi post hoc test for quantitative variables, or the Pearson χ^2 test for qualitative variables. The disease duration of the two patient groups (PPA and MD) was compared using the Mann-Whitney U test.

Global Diagnostic Score. Before proceeding with statistical analysis, a global score of the SAND battery was calculated. The global score includes the twenty-three task-related scores. Three steps were followed: (1) The raw scores were adjusted by adding or subtracting the influence of age, sex, and education and corrected using normative data. (2) Corrected scores were compared with the corresponding cutoff values obtained from HC. (3) The sum of the twenty-three dichotomous variables (1 = pathological, 0 = normal) represented the global score, with higher scores indicating more severe impairment (range 0–23).

Reliability and Construct Validity. The reliability of the internal consistency of the SAND battery was estimated using Cronbach's alpha coefficient [18]. Construct validity was assessed considering only PPA patients and using corrected scores of the task-related-scores. In order to assess construct validity, a correlation analysis between corrected task-related scores and the other language/nonlanguage tests was carried out. Tables 2 and 3 display the correlation results. All correlations were estimated using Spearman's rank correlation coefficient in order to deal with the nonnormal distributions of the scores.

Receiver Operating Characteristic (ROC) Analyses. The ROC curve for the SAND global score was used to detect the optimal cutoff score, maximizing both sensitivity and specificity. AUC, sensitivity, and specificity were calculated along with their 95% bootstrapped confidence interval (CI), considering 2,000 sampling replications. Positive predictive value (PPV) and negative predictive value (NPV) were also reported for the optimal cutoff. To assess the discriminatory power of the SAND battery (global score), the ROC analysis was used for both comparisons: PPA versus HC and PPA versus MD. All analyses were performed using R (v 3.3.1) and Rstudio (v 1.0.153). A p value <0.05 was considered statistically significant.

Results

Demographics

No significant differences were detected between groups in demographic data, with the exception of age, where HC were younger than the PPA group ($p < 0.001$); however, this variable did not influence the results, since all performances were adjusted for age, sex, and education effects. MD patients had a significantly longer disease duration than PPA patients ($p = 0.007$).

Global Diagnostic Score

The global scores of the SAND battery for HC versus PPA versus MD are displayed in online supplementary Figure 2S. As expected, PPA patients performed worse than HC ($p < 0.001$) and MD patients ($p = 0.011$). The performances of HC were close to ceiling. A significantly worse performance was observed in MD patients compared to HC ($p < 0.001$).

Reliability and Construct Validity

Cronbach's alpha coefficient of the global score of the SAND battery was 0.864 as computed from the whole sample. Table 2 shows the correlation analysis among the subscores of the SAND battery and the language tests. All SAND subscores were correlated with the conventional language test (all $p < 0.005$). The strongest correlations were found for sentence repetition with the Esame Neuropsicologico dell'Afasia sentence repetition task ($r = 0.85$) and for the naming subtest with the CAGI naming task ($r = 0.82$). Table 3 shows the correlations between the global score of the SAND battery and nonlanguage tests, which were not significant, with the exception of tasks with a relevant linguistic component (MMSE, digit span).

ROC Analyses

The ROC curve showed that the optimal cutoff score for the comparison of the PPA sample versus HC was 3, achieving a high sensitivity of 0.933 (95% CI 0.833–1.000), a high specificity of 0.946 (95% CI 0.908–0.985), a PPV of 0.800, and an NPV of 0.984 (Fig. 1a). The overall discriminatory power (AUC) for the SAND battery was 0.978 (95% CI 0.945–0.999). ROC analyses are shown in Figure 1b. A score of 2 classified PPA versus HC with a sensitivity of 0.933 (95% CI 0.833–1.000) and a specificity of 0.854 (95% CI 0.792–0.908). A score of 4 had a sensitivity of 0.833 (95% CI 0.700–0.967) and a specificity of 1.000 (95% CI 1.000–1.000).

For the comparison between PPA and MD (PD + PSP) patients, the ROC curve showed that the optimal cutoff was 5, achieving a high sensitivity of 0.767 (95% CI 0.600–0.900), a specificity of 0.667 (95% CI 0.533–0.800), a PPV of 0.605, and an NPV of 0.811 (Fig. 1c). The overall discriminatory power (AUC) was 0.786 (95% CI 0.670–0.882). ROC analyses are shown in Figure 1d. A score of 6 accurately classified PPA versus MD with a sensitivity of 0.667 (95% CI 0.500–0.883) and a specificity of 0.689 (95% CI 0.555–0.822). A score of 4 had a sensitivity of 0.833 (95% CI 0.667–0.967) and a specificity of 0.578 (95% CI 0.444–0.711).

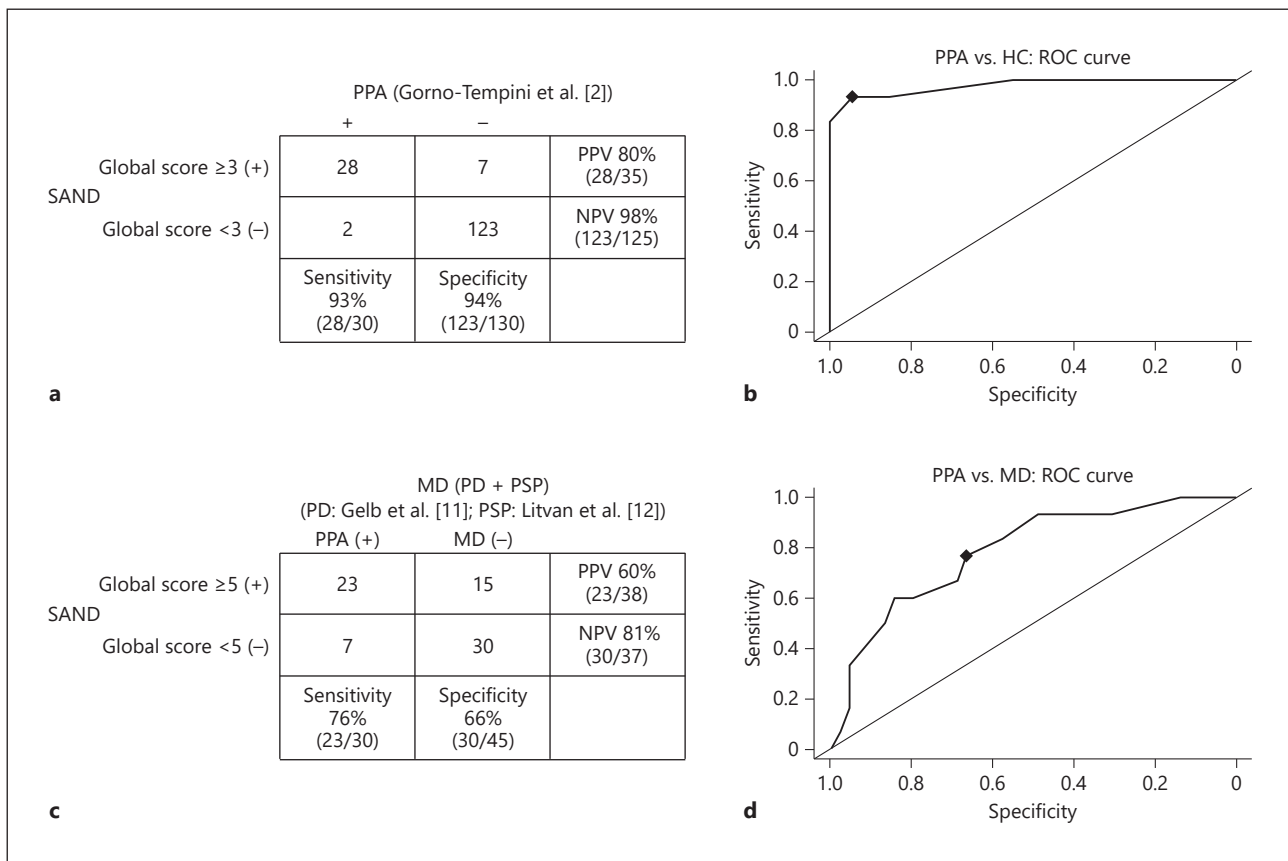


Fig. 1. **a** Summary of the diagnostic accuracy of the SAND battery for the comparison of PPA patients versus HC. **b** ROC curve for the global score of the SAND battery to detect patients with language dysfunction evaluated in the sample of PPA patients versus HC. **c** Summary of the diagnostic accuracy of the SAND battery for the comparison of MD patients versus PPA patients. **d** ROC curve for the global score of the SAND battery to detect patients with language dysfunction evaluated in the comparison of PPA patients versus MD patients. HC, healthy controls; MD, movement disorder; NPV, negative predictive value; PD, Parkinson's disease; PPA, primary progressive aphasia; PPV, positive predictive value; PSP, progressive supranuclear palsy; ROC, receiver operating characteristic; SAND, Screening for Aphasia in NeuroDegeneration.

Discussion

Language assessment plays a critical role in the clinical diagnosis of neurodegenerative diseases, in particular in the case of PPA. However, only a few attempts have been made to develop specific tools to diagnose, clinically classify, and follow up the heterogeneous group of PPA patients. Recently, a systematic review identified the few neuropsychological tests available for the assessment of speech and language disorders in PPA and discussed their limitations [7]. This is the main reason that led to the development of the SAND battery [8]. In this study, we explored its clinical validity and psychometric characteristics. Our findings suggest that the SAND battery provides useful information in the clinical diagnosis of PPA patients.

Consistency and Validity

The good internal consistency, with a Cronbach coefficient of 0.864 for the global score, means that the SAND scores are consistent with each other for the content they measure. The

data about construct validity showed that the SAND battery is a valid measure of language functions. Specifically, the strength of the correlation among the nine scores of the battery and the other measures assessing comparable components of language performance (i.e., naming, comprehension, etc.) was high, compatible with a shared general language function dimension. Divergent validity was also assessed, analyzing the correlation between the SAND global score and nonlanguage tasks. Two measures were significantly associated with the global score: (1) the MMSE, which includes language subtests (e.g., naming, writing, and repetition of sentences), and (2) the digit span backwards, which can be expected to correlate with the two tasks with a relevant working memory load, e.g., repetition of single words/nonwords and of sentences.

Clinical Usefulness

As a screening battery, the main goal of the SAND battery is to determine whether an individual has language impairment. Our results clearly indicate that an impaired performance on the test can be used to identify individuals with clinically significant language impairments. The diagnostic accuracy for specific conditions needs to be confirmed using a reliable gold standard, i.e., neuropathological confirmation, positivity of reliable biomarkers [19], or presence of pathogenetic mutations [20], associated with frontotemporal dementia or with Alzheimer's disease [21]. A diagnosis by independent experts using the current criteria [2] is at the moment the only available option in the case of PPA patients.

Speech and language disorders may occur in corticobasal degeneration and PSP. Patients with these diagnoses can actually fulfill the criteria for nf/a-PPA and apraxia of speech [22, 23] and were excluded from the present MD sample. The SAND battery is thus also able to discriminate nonaphasic MD patients from controls, providing good evidence for its sensitivity in detecting subtle language deficits in neurodegenerative diseases in general.

The SAND battery represents a first step towards a concise multilingual standard language examination, as a fast and simple tool to help clinicians and researchers in the clinical diagnosis of PPA. A further development of this battery will lead to identify the combinations of tests and test items that most reliably and accurately classify patients based on neuropsychological/linguistic features. The use of machine learning algorithms, as reported by several other studies, may represent a useful approach to this aim [e.g., 24–26], when a sufficiently large dataset allowing to identify which test/test item leads to the best discrimination of deficits in each variant will become available.

Limitations

One limitation is that the global score may be biased by the use of cutoff values obtained from the same control group used in the validation process. The group of MD patients was included in order to compare the performance of a different clinical population. Another limitation is the small sample size. Therefore, our findings require replication in larger populations to properly assess generalization. A minor point concerns PSP patients, who were included in the MD group even if they may have clinically presented as nf/a-PPA [27]. The present sample, however, was recruited from MD clinics, and we excluded cases fulfilling the criteria for PPA.

Conclusion

Taken together, our findings encourage the use in clinical practice of the SAND battery, which has been shown to be able to identify patients with PPA. Overall, the SAND battery showed good validity as a screening instrument to detect language impairment in patients

with neurodegenerative disorders. The satisfactory PPVs and NPVs and the limited length of this cognitive and language measure make this battery especially useful in clinical contexts [28]. Cognitive and language tools are inexpensive and low-tech additions to the diagnostic process that can be used in a variety of clinical settings: primary, secondary, or tertiary care and research studies. The SAND battery may be especially useful in settings where a complete neuropsychological/language assessment cannot be proposed because of lack of time or insufficient neuropsychological resources.

Acknowledgments

We acknowledge Alessandra Marcone, Chiara Cerami, Rosanna Tortelli, and Rosa Capozzo for their help in patient recruiting. We wish to thank Monica Consonni and Cristiano Chesi for their helpful comments, Caterina Balduzzi for her help in scoring data, and Laura Sampietro for assistance in data management.

Statement of Ethics

Informed written consent was obtained from all participants. The local hospital ethics committee approved the study protocol. Information collected during the course of the research was kept confidential, the subjects' names and details were removed to prevent identification of participants, and data storage was password protected, with data access restricted to study personnel.

Disclosure Statement

The authors have no conflicts of interest to disclose.

Funding Sources

This study was supported by an MRC Research Grant (Ref MR/N025881/1) and by the "AIRAlzh" Onlus and ANCC-COOP Italia to P. Battista.

References

- 1 Mesulam MM, Weintraub S: Spectrum of primary progressive aphasia. *Baillieres Clin Neurol* 1992;1:583–609.
- 2 Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, Ogar JM, Rohrer JD, Black S, Boeve BF, Manes F, Dronkers NF, Vandenberghe R, Rascovsky K, Patterson K, Miller BL, Knopman DS, Hodges JR, Mesulam MM, Grossman M: Classification of primary progressive aphasia and its variants. *Neurology* 2011; 76:1006–1014.
- 3 Cerami C, Cappa SF: Primary progressive aphasia; in Dickerson B (ed): *Hodges' Frontotemporal Dementia*. Cambridge, Cambridge University Press, 2016, pp 55–67.
- 4 Luzzatti C, Willmes K, De Bleser R, Bianchi A, Chiesa G, De Tanti A, Gonella ML, Lorenzi L, Pozzoli C: New normative data for the Italian version of the Aachen Aphasia Test (A.A.T.). *Arch Psicol Neurol Psichiatr* 1994; 55:1086–1131.
- 5 Goodglass H: Boston diagnostic aphasia examination: short form record booklet; in Goodglass H, Edith K, Barresi B (eds): *The Assessment of Aphasia and Related Disorders*, ed 3. Baltimore, Lippincott Williams and Wilkins, 2000.
- 6 Mesulam MM, Weintraub S: Is it time to revisit the classification guidelines for primary progressive aphasia? *Neurology* 2014;82:1108–1109.
- 7 Battista P, Miozzo A, Piccininni M, Catricalà E, Capozzo R, Tortelli R, Padovani A, Cappa SF, Logroscino G: Primary progressive aphasia: a review of neuropsychological tests for the assessment of speech and language disorders. *Aphasiology* 2017;31:1359–1378.
- 8 Catricalà E, Gobbi E, Battista P, Miozzo A, Polito C, Boschi V, Esposito V, Cuoco V, Barone P, Sorbi S, Cappa SF, Garrard P: SAND: a Screening for Aphasia in NeuroDegeneration. Development and normative data. *Neurol Sci* 2017;38:1469–1483.

- 9 Garrard P, Ahmed S: An abbreviated examination for the assessment of linguistic impairment in primary progressive aphasia. *Eur J Neurol* 2012;19:561.
- 10 Boschi V, Catricalà E, Consonni M, Chesi C, Moro A, Cappa SF: Connected speech in neurodegenerative language disorders: a review. *Front Psychol* 2017;8:269.
- 11 Gelb DJ, Oliver E, Gilman S: Diagnostic criteria for Parkinson disease. *Arch Neurol* 1999;56:33–39.
- 12 Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC, Goetz CG, Golbe LI, Grafman J, Growdon JH, Hallett M, Jankovic J, Quinn NP, Tolosa E, Zee DS: Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP International Workshop. *Neurology* 1996;47:1–9.
- 13 Fahn S, Elton R: Unified Parkinson's Disease Rating Scale; in Fahn S, Marsden CD, Calne DB, Goldstein M (eds): *Recent Developments in Parkinson's Disease*, Vol 2. Florham Park, Macmillan Health Care Information, 1987, pp 153–163.
- 14 Payan CA, Viallet F, Landwehrmeyer BG, Bonnet AM, Borg M, Durif F, Lacomblez L, Bloch F, Verny M, Fermanian J, Agid Y, Ludolph AC, Leigh PN, Bensimon G; NNIPPS Study Group: Disease severity and progression in progressive supranuclear palsy and multiple system atrophy: validation of the NNIPPS – Parkinson Plus Scale. *PLoS One* 2011;6:e22293.
- 15 Measso G, Cavarzeran F, Zappala G, Lebowitz BD, Crook TH, Pirozzolo FJ, Amaducci LA, Massari D, Grigoletto F: The mini-mental state examination: normative study of an Italian random sample. *Dev Neuropsychol* 1993; 9:77–85.
- 16 American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV)*. Washington, DC, American Psychiatric Association, 1994.
- 17 World Health Organization: *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva, World Health Organization, 1992.
- 18 Tavakol M, Dennick R: Making sense of Cronbach's alpha. *Int J Med Educ* 2011;2:53–55.
- 19 Grossman M: Biomarkers in the primary progressive aphasias. *Aphasiology* 2014;28:922–940.
- 20 Rohrer JD: The genetics of primary progressive aphasia. *Aphasiology* 2014;28:941–947.
- 21 Chare L, Hodges JR, Leyton CE, McGinley C, Tan RH, Kril JJ, Halliday GM: New criteria for frontotemporal dementia syndromes: clinical and pathological diagnostic implications. *J Neurol Neurosurg Psychiatry* 2014; 85:866–871.
- 22 Kertesz A, McMonagle P, Blair M, Davidson W, Munoz DG: The evolution and pathology of frontotemporal dementia. *Brain* 2005;128:1996–2005.
- 23 Hodges JR, Davies RR, Xuereb JH, Casey B, Broe M, Bak TH, Kril JJ, Halliday GM: Clinicopathological correlates in frontotemporal dementia. *Ann Neurol* 2004;56:399–406.
- 24 Battista P, Salvatore C, Castiglioni I: Optimizing neuropsychological assessments for cognitive, behavioral, and functional impairment classification: a machine learning study. *Behav Neurol* 2017;2017:1850909.
- 25 Fraser KC, Meltzer JA, Rudzicz F: Linguistic features identify Alzheimer's disease in narrative speech. *J Alzheimers Dis* 2016;49:407–422.
- 26 Tunvirachaisakul C, Supasitthumrong T, Tangwongchai S, Hemrunroj S, Chuchuen P, Tawankanjanachot I, Likitchareon Y, Phanthumchinda K, Sriswasdi S, Maes M: Characteristics of mild cognitive impairment using the Thai version of the Consortium to Establish a Registry for Alzheimer's Disease tests: a multivariate and machine learning study. *Dement Geriatr Cogn Disord* 2018;45:38–48.
- 27 Merello M, Starkstein SE (eds): *Movement Disorders in Dementias*. London, Springer, 2014.
- 28 Prince M, Bryce R, Ferri C: *World Alzheimer Report 2011: The benefits of early diagnosis and intervention*. <https://www.alz.co.uk/research/world-report-2011>.