International Psychogeriatrics (2013), 25:9, 1503–1511 © International Psychogeriatric Association 2013
doi:10.1017/S1041610213000811

The Brazilian version of the Neuropsychiatric Inventory-Clinician rating scale (NPI-C): reliability and validity in dementia

Florindo Stella,[1,2] Orestes Vicente Forlenza,[2] Jerson Laks,[3] Larissa Pires de Andrade,[1] Michelle A. Ljubetic Avendaño,[2] Elisandra Villela Gasparetto Sé,[4] João de Castilho Cação,[5] Constantine G. Lyketsos[6] and Kate de Medeiros[7]

[1] *Biosciences Institute, Universidade Estadual Paulista (UNESP), Rio Claro, SP, Brazil*
[2] *Laboratory of Neurosciences (LIM-27), Department and Institute of Psychiatry, Faculty of Medicine, University of São Paulo, SP, Brazil*
[3] *Centre for Alzheimer's Disease and Related Disorders, Institute of Psychiatry, Federal University of Rio de Janeiro, RJ, Brazil*
[4] *Clinic of Geriatric Psychiatry, Faculty of Medicine, State University of Campinas, Campinas, SP, Brazil*
[5] *FAMERP – Faculty of Medicine, Geriatric Unity, São José do Rio Preto, Brazil*
[6] *Department of Psychiatry, Johns Hopkins Bayview, Johns Hopkins University, Baltimore, Maryland, USA*
[7] *Department of Sociology and Gerontology, Miami University, Oxford, Ohio, USA*

ABSTRACT

Background: Patients with dementia may be unable to describe their symptoms, and caregivers frequently suffer emotional burden that can interfere with judgment of the patient's behavior. The Neuropsychiatric Inventory-Clinician rating scale (NPI-C) was therefore developed as a comprehensive and versatile instrument to assess and accurately measure neuropsychiatric symptoms (NPS) in dementia, thereby using information from caregiver and patient interviews, and any other relevant available data. The present study is a follow-up to the original, cross-national NPI-C validation, evaluating the reliability and concurrent validity of the NPI-C in quantifying psychopathological symptoms in dementia in a large Brazilian cohort.

Methods: Two blinded raters evaluated 312 participants (156 patient-knowledgeable informant dyads) using the NPI-C for a total of 624 observations in five Brazilian centers. Inter-rater reliability was determined through intraclass correlation coefficients for the NPI-C domains and the traditional NPI. Convergent validity included correlations of specific domains of the NPI-C with the Brief Psychiatric Rating Scale (BPRS), the Cohen-Mansfield Agitation Index (CMAI), the Cornell Scale for Depression in Dementia (CSDD), and the Apathy Inventory (AI).

Results: Inter-rater reliability was strong for all NPI-C domains. There were high correlations between NPI-C/delusions and BPRS, NPI-C/apathy-indifference with the AI, NPI-C/depression-dysphoria with the CSDD, NPI-C/agitation with the CMAI, and NPI-C/aggression with the CMAI. There was moderate correlation between the NPI-C/aberrant vocalizations and CMAI and the NPI-C/hallucinations with the BPRS.
Conclusion: The NPI-C is a comprehensive tool that provides accurate measurement of NPS in dementia with high concurrent validity and inter-rater reliability in the Brazilian setting. In addition to universal assessment, the NPI-C can be completed by individual domains.

Key words: neuropsychiatric symptoms, dementia, Alzheimer's disease, scale, neuropsychiatric assessment, Brazil

Introduction

In addition to cognitive and functional impairment, neuropsychiatric symptoms (NPS) are a near universal aspect of a dementia diagnosis (Gauthier *et al.*, 2010; Dubois *et al.*, 2010; Lyketsos *et al.*, 2011). NPS contribute to increases in the patient's clinical deterioration and caregiver burden, and lead to diminished quality of life for patients and family members. NPS also predict likelihood of earlier institutionalization, interact with other comorbidities, accelerate disease progression, and increase mortality risk (Gauthier *et al.*, 2010; Lyketsos *et al.*, 2011; Lyketsos and Miller, 2012; Wadsworth *et al.*, 2012). In addition, multiple co-occurring NPS (e.g., depression with apathy, irritability with aggression,

Correspondence should be addressed to: Florindo Stella, Biosciences Institute, Universidade Estadual Paulista (UNESP), Rio Claro 13506-900, SP, Brazil. Phone: 55-19-3526-4245; Fax: 55-19-3526-4246. Email: fstelldo@gmail.com.
Received 20 Feb 2013; revision requested 8 Apr 2013; revised version received 6 May 2013; accepted 10 May 2013. First published online 14 June 2013.

or delusions with agitation) complicate the clinical picture, creating challenges for diagnosis and treatment (Brodaty *et al.*, 2001; Lopez *et al.*, 2003; de Medeiros *et al.*, 2010; Lyketsos *et al.*, 2011; Benoit *et al.*, 2012). In Brazil, several cohort studies found high prevalences of NPS in dementia, namely apathy, agitation, aggression, depression, sleep disturbances, anxiety, and aberrant motor behavior (Tatsch *et al.*, 2006; Camozzato *et al.*, 2008; Truzzi *et al.*, 2013).

Even in prodromal stages of dementia, NPS may also be present; indeed, recent studies have shown that the occurrence of NPS in patients with mild cognitive impairment (MCI) may increase the risk of subsequent progression to dementia (Taragano *et al.*, 2009; Di Iulio *et al.*, 2010; Lyketsos *et al.*, 2011). Sperling and colleagues (2011) postulated that cognitive and behavioral changes could represent an early stage of a progressive dementia in individuals with evidence of a long asymptomatic period characterized by neuropathological Alzheimer's disease (AD) biomarkers. Accurate measurement of NPS, even at early stages of dementia, is therefore critical for diagnosis and treatment (Gauthier *et al.*, 2010; Lyketsos *et al.*, 2011; Sperling *et al.*, 2011).

Despite approaches to improve accurate measurement of NPS in dementia, there are several major challenges for clinicians and researchers. Many existing measurements are based solely on patient's or caregiver's inputs, leading to significant measurement limitations. The patient may be unable to provide reliable information due to cognitive decline (e.g., forgetfulness) or lack insight; caregivers frequently suffer emotional burden that interferes with appropriate judgment of the patient's behavior (Rosenberg *et al.*, 2005; de Medeiros *et al.*, 2010; Lyketsos *et al.*, 2011). Reports by patients and caregivers alone may, therefore, not provide a complete or accurate picture of NPS. In contrast, clinicians with training in NPS screening in patients with dementia should be able to incorporate caregiver and patient information along with their clinical judgment to achieve a better understanding of the behavioral syndrome. This strategy of using clinical impression ratings can improve accuracy in measuring the clinical relevance of each symptom, as well as in distinguishing neuropsychiatric conditions when symptoms overlap, for instance deciding whether a given symptom is due to apathy or depression, a decision usually difficult for the non-clinician (de Medeiros *et al.*, 2010; Benoit *et al.*, 2012).

The Neuropsychiatric Inventory-Clinician rating scale (NPI-C) is a comprehensive and versatile psychometric scale that has been designed to measure NPS both in clinical and research settings. As compared to similar instruments designed for the same purpose, the NPI-C rating incorporates the expert clinician's impressions (i.e., all relevant information according to his/her clinical judgment and patient records) to the data provided by patients and caregivers. The NPI-C output may be readily compared across distinct investigation sites, which renders it particularly useful for clinical trials (de Medeiros *et al.*, 2010; Lyketsos *et al.*, 2011).

The present study is a follow-up to the original, cross-national NPI-C validation (de Medeiros *et al.*, 2010), evaluating the reliability and concurrent validity of the NPI-C in quantifying psychopathological symptoms in dementia in a large Brazilian cohort.

Although Brazil was one of the nine international sites that participated in original NPI-C validation study, the present Brazilian cohort study is justified for several reasons. First, sociocultural context may influence caregivers' beliefs, perceptions, and attitudes with regional variations in Latin America (Peluso and Blay, 2004). This was not possible to consider in the international validation study in which 15 patient/caregiver dyads participated. Also, the lack of socioeconomic resources associated with recent political decisions has led to shortages in long-stay nursing homes for people with dementia (Creutzberg *et al.*, 2007), requiring families to provide care at home, often with increasing emotional burden. Low levels of education still are also common in Brazil, and have become an additional problem for the poorest people. Taken together, these factors contribute to caregiver difficulties in interpreting NPS and point to the importance of a large NPI-C validation study in Brazil.

Another important rationale for the current study is apathy assessment. The original validation of the NPI-C used the Apathy Evaluation Scale (AES; Marin *et al.*, 1991) for convergent validity. However, as noted in the original article, data on the AES for the Brazilian subgroup were excluded from analyses due to rater discrepancies in the interpretation of some AES items (de Medeiros *et al.*, 2010). In the current study, the Apathy Inventory (AI; Robert *et al.*, 2002; 2010), an instrument already validated for Brazilians community (Stella *et al.*, 2013), was used instead of the AES.

Methods

Structure of the NPI-C

The development and cross-national validation of the NPI-C was led by de Medeiros and Lyketsos with the participation of an international group including researchers from Argentina, Australia, Brazil, Canada, France, Greece, Hungary, Italy, and the USA. This group worked closely with the original developer of the Neuropsychiatric

Inventory (NPI; Cummings, 1994). De Medeiros *et al.* (2010) published the final version of the NPI-C, making it available for researchers and clinicians.

The NPI-C encompasses all traditional NPI domains (agitation and aggression are split into individual domains in the NPI-C), includes more items and one new domain (aberrant vocalizations), and uses clinician judgment to rate the severity of each neuropsychiatric symptom (de Medeiros *et al.*, 2010). These improvements rectified important weaknesses of the traditional NPI such as possible biases caused by emotional stress and excessive burden to family members or caregivers, and possible cognitive decline when the caregiver is older and may also suffer from memory loss, which could compromise the response reliability and accuracy. In each domain of the NPI-C, the clinician ultimately decides on the clinical value of each symptom considering responses from a knowledgeable informant, the direct interview with the patient, and additional relevant information from the patient's record or from direct observation of the patient's behaviors. This clinician judgment provides more accurate assessment of NPS in dementia. The NPI-C is available to be used as a broad-spectrum scale or as a single tool driven to selected neuropsychiatric domains.

The NPI-C was translated into Portuguese (Brazil) (*Inventário Neuropsiquiátrico – Avaliação do Clínico – NPI-C*): An independent expert in English performed the back-translation into English to verify the reliability of the text to be used in Brazil.

As mentioned earlier, one limitation of the traditional NPI concerns scoring that is based solely on caregiver reports provided during a clinical interview (Lyketsos *et al.*, 2011). Caregivers provide a global rating of frequency, severity, and caregiver distress for each of 12 NPS domains. Two relevant changes were incorporated in the NPI-C. First, rather than provide a global rating, caregivers are asked to rate individual symptoms within a domain for frequency, severity, and caregiver distress. Since it is well known that caregiver informants often suffer emotional burdens that may lead to biased reporting of the patient's symptomatology, the second important change is a clinical judgment rating of NPS. The clinician now provides a rating for each item in each domain based on an interview with the caregiver (and or other informants), direct observation of and interaction with the patient, and additional information from the patient record. The clinician scoring assessment minimizes potential inaccuracies or misinterpretation of symptoms (e.g., confusing apathy with depression) by the caregiver (de Medeiros *et al.*, 2010; Lyketsos *et al.*, 2011). Thus, changes in the NPI-C represent a rating approach that can reduce inappropriate influences from caregivers or family members when they report the nature or severity of NPS.

Inclusion criteria

In the present study, the inclusion criteria were the same applied to the original validation of the NPI-C (de Medeiros *et al.*, 2010), and were divided into two parts: (a) knowledgeable informants and (b) patients. For knowledgeable informants (caregiver or family members), inclusion criteria were capacity to identify and to report on NPS in the patient over the past month, and having maintained regular verbal contact with the patient at least three times a week during the past three months. Patients were included if they had a medical diagnosis of probable AD as well as a knowledgeable informant (caregiver or family member).

Sample

We studied 312 participants from five Brazilian centers (156 patient-knowledgeable informant dyads), who completed the NPI-C with two blinded raters, totaling 624 observations. In these centers (São Paulo, Rio de Janeiro, Campinas, São José do Rio Preto, and Rio Claro), the raters determined dementia severity by the Mini-Mental State Examination (MMSE; Folstein *et al.* 1975), the Clinical Dementia Rating (CDR; Hughes *et al.*, 1982), the Pfeffer Functional Activities Questionnaire (Pfeffer, 1982), and the Global Deterioration Scale (GDS; Reisberg *et al.*, 1982). Table 1 summarizes participant demographic and clinical data including the MMSE, CDR, and GDS.

Ethics review boards approved the study, and family members or legal representatives signed written informed consent, as did as patients who were able to understand the purpose of investigation. The research was conducted according to the principals of the Helsinki Declaration.

Procedures

We performed a cross-sectional investigation of patients and respective caregivers to estimate inter-rater reliability and convergent validity of the Brazilian translation of the NPI-C. For the validation of the NPI-C, all participants (100%) completed the scale.

Inter-rater reliability

At each center, two independent and trained raters interviewed each knowledgeable informant/patient dyad at different times (in general, in the same day or at least within the same week) in order to estimate the inter-rater reliability of the NPI-C. Before administering the measure, an investigator (Florindo Stella) who previously had participated in the

Table 1. Mean for selected demographic features and scales according to the Clinical Dementia Rating

DEMOGRAPHIC AND CLINICAL FEATURES	CDR 1 $N = 60$ (F:42/M:18)	CDR 2 $N = 53$ (F:44/M:9)	CDR 3 $N = 43$ (F:29/M:14)	TOTAL $N = 156$ (F:115/M:41)	p VALUE
Age (years)	77.4	76.8	75.6	76.7	0.45
Education (years)	5.9	5.8	4.4	5.5	0.15
Mini-Mental State Examination	22.8	17.4	9.2	17.2	<0.01
Pfeffer Functional Activities Questionnaire	8.3	18.3	26.2	16.6	<0.01
Global Deterioration Scale	2.7	4.3	6.0	4.2	<0.01
Apathy Inventory	2.9	5.9	10.1	5.9	<0.01
Brief Psychiatric Rating Scale	10.6	26.0	55.3	28.1	<0.01
Cohen-Mansfield Agitation Index	7.7	25.4	56.4	27.1	<0.01
Cornell Scale for Depression in Dementia	5.6	6.5	7.9	6.6	0.09

$p < 0.05$ (one-way ANOVA).
CDR = Clinical Dementia Rating; F = female; M = male.

original validation of the NPI-C trained the raters on completion of all study scales. Eligible clinician raters for this study were working regularly at specialized centers with patients with dementia observing or making diagnostic procedures, providing appropriate care, counseling caregivers or family members, and establishing or following pharmacological or non-pharmacological treatment.

Each independent rater completed the NPI-C. At first, caregivers were asked to rate the frequency and severity of each item in each domain, as well as estimate their distress level due to the patient's behavior. In addition, raters interviewed the patient. This provided them an opportunity to interact with the patient and, when possible, to obtain the patient's insight into their recent experiences that may be related to NPS. Based on caregiver answers, patient interview, and any additional clinical information (e.g., a specific report from the patient record), raters used their experience and judgment to determine the final score for each symptom of the NPI-C. Symptom scores in each of the 14 domains were added up to produce total clinician rating domain scores. Inter-rater reliability was determined through calculation of intraclass correlation coefficients for the NPI-C domain scores, as well as for those of the traditional NPI that was administered as part of the NPI-C (Shrout and Fleiss, 1979).

CONVERGENT VALIDITY

In addition to estimating the convergent validity of specific domains of the NPI-C and to make appropriate correlations, one of the two raters completed four selected scales: the Brief Psychiatric Rating Scale (BPRS; Ventura *et al.*, 1993); the Cohen-Mansfield Agitation Index (CMAI; Cohen-Mansfield *et al.*, 1989); the Cornell Scale for Depression in Dementia (CSDD; Alexopoulos *et al.*, 1988); and the AI (Robert *et al.*, 2002; 2010). In order to obtain precise information about NPS, the scales were completed regardless of the caregiver's response to the screening question for each of the NPI-C domains.

Pearson correlations of these data were used to provide statistical support for convergent validity, as well as for test–retest reliability.

Within the convergent validity strategy, we expected moderate to strong association of NPI-C domains with selected scales, with significant correlations between correspondent assessments. Furthermore, at the end of the interview, each rater estimated the assessment of respondent's reliability using a short questionnaire based on distinct questions requiring sequential levels (0 = poor; 1 = fair; 2 = good; 3 = excellent). Considering the quality of information provided during the interview, each rater scored the response reliability from the same caregiver.

Data analyses

Descriptive analyses including mean, standard deviation, and frequency were estimated for general demographic data such as age, gender, and years of education, as well as clinical performance on scores from performance on the MMSE, Pfeffer Functional Activities Questionnaire, CDR, GDS, CMAI, BPRS, CSDD, and AI. Inter-rater reliabilities for each NPI-C and NPI domain were estimated by calculating intraclass correlations (ICC).

Convergent validity was estimated by calculating Pearson correlations coefficient between NPI-C domains and selected scales (NPI-C/delusions vs. BPRS; NPI-C/hallucinations vs. BPRS; NPI-C/agitation vs. CMAI; NPI-C/aggression vs. CMAI; NPI-C/aberrant vocalizations vs. CMAI; NPI-C/depression vs. CSDD; and NPI-C/apathy vs. AI). We also examined convergent validity by dementia severity based on CDR scores. A one-way ANOVA was applied to identify differences between scores from clinical data, relative to CDR severity levels. All statistical analyses were carried out using SPSS 20.0.

Table 2. Inter-rater reliability (ICCs) with 95% confidence limits for NPI-C domains (156 patients)

NPI-C, TRADITIONAL NPI, AND INTER-RATER RELIABILITY	ICC CI (95%)
NPI – Traditional total score	0.923 (0.897–0.944)
NPI-C domains:	
Delusions	0.937 (0.914–0.954)
Hallucinations	0.777 (0.707–0.832)
Agitation	0.903 (0.869–0.928)
Aggression	0.879 (0.838–0.910)
Depression/dysphoria	0.812 (0.751–0.860)
Anxiety	0.826 (0.769–0.870)
Elation/euphoria	0.916 (0.887–0.938)
Apathy/indifference	0.865 (0.819–0.900)
Disinhibition	0.947 (0.928–0.961)
Irritability/lability	0.904 (0.870–0.929)
Aberrant motor behavior	0.899 (0.864–0.925)
Sleep disorders	0.844 (0.792–0.884)
Appetite and eating disorders	0.877 (0.835–0.909)
Aberrant vocalizations	0.915 (0.885–0.937)

ICC = intraclass correlations; NPI-C = Neuropsychiatric Inventory-Clinician rating scale; NPI = Neuropsychiatric Inventory; CI = confidence interval.

Results

Demographic and clinical features

Table 1 displays demographic and clinical data from patients ($n = 156$) according to severity levels of dementia based on CDR scores: 60 had mild dementia (CDR 1), 53 moderate (CDR 2), and 43 severe (CDR 3).

As expected, the scores of the MMSE were progressively worse according to clinical deterioration in the CDR severity levels (one-way ANOVA, $p < 0.01$). Likewise, the Pfeffer Questionnaire showed greater impairment of daily living activities among patients with more severe CDR (one-way ANOVA; $p < 0.05$). These data are compatible with the global impairment measured by the GDS relative to the CDR (one-way ANOVA, $p < 0.05$). In turn, psychopathological symptoms assessed by the AI, BPRS, and CMAI were progressively severe according to the CDR (one-way ANOVA, $p < 0.01$). Conversely, depressive symptoms, assessed by the CSDD, remained without significant changes in the three groups of the CDR (one-way ANOVA, $p = 0.09$). Results from these analyses are displayed in Table 1.

Inter-rater reliability

Inter-rater reliability was strong (ICCs with 95% confidence limits) for all NPI-C domains. In addition, inter-rater reliability was strong for the NPI traditional score ($r = 0.923$). Data are included in Table 2.

Convergent validity

Table 3 reports Pearson correlations coefficients with their 95% confidence intervals for convergent validity. There were significant correlations of specific NPI-C domains with selected scales.

Based on dementia severity according to CDR levels, Table 4 includes correlations between specific NPI-C domains and selected scales that measure the same psychopathological syndromes with 95% confidence intervals for concurrent validity.

Hallucinations had the lowest correlation within the validity analysis. The correlations were not significant for hallucinations or aberrant vocalizations at any CDR level.

Concerning the respondents' assessment of reliability, 109 caregivers completed the short questionnaire with each rater at ending of the interview. According to Pearson correlations, the caregiver reliability was moderate ($r = 0.601$; 95% CI = 0.47, 0.71).

Discussion

The present study aimed to estimate the convergent validity and inter-rater reliability of the Brazilian version of the NPI-C. In doing so, the study covers an important gap in the measurement of psychopathological symptoms in dementia in Brazilians community.

As mentioned, although the Brazilian group participated in the original validation of the NPI-C, different sociocultural aspects support the current Brazilian cohort study. Regional features, mainly associated with low levels of education and distinct perceptions of behavioral disturbances in demented patients, as well as emotional exhaustion in familial caregivers are different than caregivers and patients from developed countries (Peluso and Blay, 2004; Creutzberg *et al.*, 2007; Truzzi *et al.*, 2013).

Inter-rater reliability

Inter-rater reliability was strong for all NPI-C domains. In general, correlations from this study are consistent with data from the original NPI-C validation by de Medeiros *et al.* (2010). The achieved correlation coefficients of the scores from independent raters can be considered reliable. Concerning total scores from the traditional NPI, in our study inter-rater reliability was higher than the moderate caregiver reliability. The discrepancy in reliability between raters and caregivers on interpretation of patients' symptoms reinforces the need for the clinician judgment such as the rating method used in the NPI-C regarding NPS in dementia.

Table 3. Pearson correlations and confidence intervals (95% confidence intervals) for concurrent (convergent) validity involving the NPI-C and selected scales (patients = 156) by the strength of correlation

RATER 1 PLUS RATER 2	PEARSON CORRELATION R ICC CI (95%)
NPI-C/Apathy-Indifference × Apathy Inventory	0.942 (0.921–0.958)
NPI-C/Agitation × Cohen-Mansfield Agitation Index	0.772 (0.700–0.829)
NPI-C/Aggression × Cohen-Mansfield Agitation Index	0.769 (0.696–0.826)
NPI-C/Depression–Dysphoria × Cornell Scale for Depression in dementia	0.736 (0.655–0.801)
NPI-C/Delusions × Brief Psychiatric Rating Scale-delusions	0.713 (0.626–0.783)
NPI-C/Aberrant Vocalizations × Cohen-Mansfield Agitation Index	0.684 (0.591–0.760)
NPI-C/Hallucinations × Brief Psychiatric Rating Scale-hallucinations	0.432 (0.295–0.552)

NPI-C = Neuropsychiatric Inventory–Clinician rating scale; CI = confidence interval.

Table 4. Pearson correlations (95% confidence intervals) for concurrent (convergent) validity involving the NPI-C and selected scales by the CDR (patient severity) rating

	CDR 1 (MILD) $N = 60$	CDR 2 (MODERATE) $N = 54$	CDR 3 (SEVERE) $N = 43$
NPI-C Apathy × AI	0.89 (0.82–0.93)	0.910 (0.85–0.95)	0.796 (0.65–0.89)
NPI-C Agitation × CMAI	0.691 (0.53–0.80)	0.619 (0.42–0.76)	0.745 (0.57–0.85)
NPI-C Aggression × CMAI	0.615 (0.43–0.75)	0.717 (0.56–0.83)	0.710 (0.52–0.83)
NPI-C Depression × CSDD	0.701 (0.54–0.81)	0.648 (0.46–0.78)	0.579 (0.34–0.75)
NPI-C Delusion × BPRS	0.514 (0.300 – 0.68	0.53 (0.30–0.69)	0.645 (0.43–0.79)
NPI-C Aberrant Vocalizations × CMAI	0.650 (0.47–0.77)	0.631 (0.44–0.77)	0.09 (−0.22–0.38)
NPI-C Hallucinations × BPRS	0.146(−0.11 – 0.39)	0.080 (−0.19– 0.34)	0.258 (−0.05–0.52)

Convergent validity

The convergent validity appears to be close to that of the original version of the NPI-C (de Medeiros *et al.*, 2010). There was a strong correlation between most NPI-C domains and the selected validation scales. The highest correlation was observed between the NPI-C/apathy domain and the AI, much stronger than data reported in the original validation study. In the original NPI-C validation, the NPI-C/apathy domain had weaker correlation with the AES (Marin *et al.*, 1991). We correlated this domain with the AI (Robert *et al.*, 2002; 2010), which is more suitable to the dementia setting, and convergent validity was strongest for this domain and corresponding measures. Nevertheless, while we observed the strongest correlation between the apathy domain and the AI (0.942), apathy had the weakest correlation in the original study when the AES instrument was used (0.31). Conversely, we found the lowest correlation for the hallucinations domain and the BPRS (0.432), while in the original study this correlation was higher (0.60). Therefore, the major discrepancies between both studies were related to the NPI-C/apathy and NPI-C/hallucinations domains.

At least in part, the discrepancy regarding hallucinations domain could be related to conflicting responses by caregivers or their interpretation of patient's psychopathological manifestations. The occurrence of distinct types of hallucination-related behaviors, described by caregivers or observed in patients, may have influenced the understanding of symptoms and reporting them. These factors would affect the ICC value concerning the correlation between the NPI-C/hallucinations and the BPRS. Concerning apathy, as we mentioned earlier in the Methods section, the AES was problematic for the Brazilian site in the original NPI-C study. To address this issue in the current study, we replaced the AES by the AI (Robert *et al.*, 2002; 2010), which was already adapted and validated for Brazilians community with good specificity (97.3%) and sensitivity (99.2%) (Stella *et al.*, 2013).

We also examined concurrent validity by dementia severity. We note some differences in the strength of correlation by NPS. Specifically, there was exceptionally strong correlation between the NPI-C/apathy and the AI and NPI-C/depression and the CSDD in mild-to-moderate dementia. The strength of correlation declined in more severe patients. As observed in the Results section, the NPI-C/hallucinations domain had the lowest correlation with the BPRS. In addition, the correlations were not significant for NPI-C/hallucinations or NPI-C/aberrant vocalizations at any CDR level.

Perhaps the biggest surprise was the difference in correlation for NPI-C aberrant vocalizations and the CMAI for mild and moderate patients ($r = 0.65$ and 0.63, respectively) compared to the severe patients ($r = 0.09$). This is likely due

to lack of communicative ability in late stages of dementia.

Patients with more clinical severity of AD showed higher scores on the CMAI (in which the domain NPI-C/aberrant vocalizations was inserted) in the original version and in the present study. In Brazil, we still need studies comparing the prevalence or intensity of NPS, including aberrant vocalizations, with severity levels of AD. In long-stay nursing homes, where there tends to be a high number of patients with advanced dementia, it would be possible to identify high prevalence and intensity of this clinical syndrome.

Psychopathological features

In agreement with the original study in the current convergent validity, most NPI-C domains had high correlations with selected scales (NPI-C/apathy vs. AI; NPI-C/agitation vs CMAI; NPI-C/aggression vs. CMAI; NPI-C/delusions vs. BRPS; NPI-C/depression vs. CSDD; and NPI-C/aberrant vocalizations vs. CMAI).

The question if depression and apathy integrate into the same syndrome or belong to the distinct psychopathological condition has been a continuous debatable matter (Aalten *et al.*, 2007). The NPI-C includes specific items, which appropriately discriminate one syndrome from another. While the core of apathy relates to reduced goal-directed behavior, reduced goal-directed cognition, and emotional blunting (Robert *et al.*, 2010; Benoit *et al.*, 2012); depression essentially concerns emotional disturbances by sadness and decreased pleasure in daily activities.

In estimating convergent validity for the NPI-C against the CMAI, the strength of correlation was high concerning the agitation (0.772) and aggression (0.769) domains. Whereas agitation may emerge in combination with different neuropsychiatric syndromes, this domain can be targeted as a stand-alone measure as was established in the original version of the NPI-C (de Medeiros *et al.*, 2010). Agitation includes uncommon physical or verbal non-threatening behavior, such as wandering, uncooperative attitudes, resistance to care, or unusual non-threatening communication. Conversely, aggression comprises angry behavior, including physical or verbal threatening attitudes, intentional attempts to hit or hurt people or things, and uncommon threatening screams toward self or other person. Based on specific psychopathological features and different prevalence rates (Lyketsos *et al.*, 1999; 2011; Brodaty *et al.*, 2001), the NPI-C considered agitation and aggression as two separate domains with insertion of new items to each (de Medeiros *et al.*, 2010).

Aberrant vocalizations was a new domain added to the NPI-C and also was compared with the CMAI since this scale comprises some items concerning verbal behavioral disturbances with moderate correlation (0.684). Aberrant vocalizations constitute a prominent framework in advanced dementia, and this domain was added to the NPI-C in order to capture symptoms present in this dementia stage (de Medeiros *et al.*, 2010).

The delusions domain from the NPI-C highly correlated with the BPRS (0.713). Hallucinations domain from the NPI-C presented only a moderate correlation with the BPRS (0.432).

The reliability of responses to each rater during the interview, available for 109 caregivers, was moderate ($r = 0.601$). This is an important point to be considered because it reveals some degree of change in the quality of information provided by the caregiver to each rater. During the interview, emotional reactions involving the caregiver and rater relationship, such as anxiety or difficulties in interpersonal relationship, and patient's irritability, could interfere with the quality of responses.

In general, the concurrent validity was supported by good coefficient correlations of NPI-C domains with selected scales measuring the same neuropsychiatric syndromes. Furthermore, the validity and reliability of the scale suggest it would be a useful tool for Brazilian communities to assess neuropsychiatric syndromes through both clinical practice and research.

Low quality of accuracy in assessing of psychopathological manifestations in dementia is an important problem for diagnosis and treatment. Appropriate methodological strategies are critical to assess NPS in patients with dementia through valid, sensitive, and reliable instruments. These resources provide crucial information to the clinician in order to adequately interpret the patient's psychopathological syndromes. In the NPI-C, the clinician rating approach minimizes the caregiver's misidentification concerning the patient's behaviors. It improves the accuracy in each NPI-C domain by direct observation of and interview with the patient, and by other relevant clinical information in addition to the caregiver's report. Likewise, improved diagnosis accuracy may facilitate appropriate potential treatment.

In Brazil, efforts are needed to achieve reliable and valid measures of cognition, function, and NPS in dementia (Tatsch *et al.*, 2006; Memória *et al.*, 2013; Truzzi *et al.*, 2013). The availability of the Brazilian-validated version of the NPI-C could represent a useful contribution for research and clinical trials on neuropsychiatric syndromes in

dementia. The results demonstrate strong correlations of NPI-C domains with selected scales, which measure the same psychopathological syndromes. The high convergent validity and inter-rater reliabilities provide support for the use of the NPI-C across clinical and research settings, as well as in clinical trials.

Conclusion

The NPI-C is a comprehensive tool, which provides accurate measurement of NPS in dementia with high concurrent validity and inter-rater reliability in the Brazilian setting. In addition to universal assessment, the NPI-C can be completed by individual domains, such as delusions, hallucinations, agitation, depression, apathy, sleep disorders, etc. Additional efforts to use the NPI-C in Brazil targeting a comprehensive approach to psychopathological symptoms in dementia or to apply an individual NPI-C domain to investigate a distinct neuropsychiatric syndrome should be encouraged across clinical trials and research settings, as well as clinical practice.

Conflict of interest

None.

Description of authors' roles

F. Stella designed the study, trained the research team, was responsible for the study coordination, and wrote the first draft of the manuscript. K. de Medeiros and C. Lyketsos contributed to the study content and design, supervised the statistical procedures and data analysis, contributed to drafting the manuscript and revising it critically for intellectual content, and assisted with writing the paper. O. V. Forlenza, J. Laks, and J. C. Cação contributed to the study design, participated in the raters, training at their sites, contributed to drafting the manuscript and revising it critically for intellectual content, and assisted with writing the paper. L. P. Andrade, M. A. L. Avendaño, and E. V. G. Sé contributed to the study design, participated in raters training and data collection in their sites, contributed to drafting the manuscript and revising it critically for intellectual content, and assisted with writing the paper.

Acknowledgments

This study was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; grant no. 2010/18999-6). Jerson Laks is supported by Conselho Nacional de Pesquisa (CNPq) as a researcher 2.

References

Aalten, P. *et al.* (2007). Neuropsychiatric syndromes in dementia. Results from the European Alzheimer Disease Consortium: Part I. *Dementia and Geriatric Cognitive Disorders*, 24, 457–463.

Alexopoulos, G. S., Abrams, R. C., Young, R. C. and Shamoian, C. A. (1988). Cornell Scale for Depression in Dementia. *Biological Psychiatry*, 23, 271–284.

Benoit, M. *et al.* (2012). Apathy and depression in mild Alzheimer's disease: a cross-sectional study using diagnostic criteria. *Journal of Alzheimer's Disease*, 31, 325–334.

Brodaty, H. *et al.* (2001). Psychosis, depression and behavioral disturbances in Sydney nursing home residents: prevalence and predictors. *International Journal of Geriatric Psychiatry*, 16, 504–512.

Camozzato, A. L. *et al.* (2008). Reliability of the Brazilian Portuguese version of the Neuropsychiatric Inventory (NPI) for patients with Alzheimer's disease and their caregivers. *International Psychogeriatrics*, 20, 383–393.

Cohen-Mansfield, J., Marx, M. S. and Rosenthal, A. S. (1989). A description of agitation in a nursing home. *Journal of Gerontology: Medical Sciences*, 44, M77–M84.

Creutzberg, M., Gonçalves, L. H. T., Sobottka, E. A. and Ojeda, B. S. (2007). Long-term care institutions for elders and the health system. *Revista Latino-Americana de Enfermagem*, 15, 1144–1149.

Cummings, J. L., Mega, M., Gray, K., Rosenberg-Thomson, S., Carusi, D. A. and Gornbein, J. (1994). The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology*, 44, 2308–2314.

De Medeiros, K. *et al.* and the NPI-C Research Group. (2010). The Neuropsychiatric Inventory-Clinician rating scale (NPI-C): reliability and validity of a revised assessment of neuropsychiatric symptoms in dementia. *International Psychogeriatrics*, 22, 984–994.

Di Iulio, F. *et al.* (2010). Occurrence of neuropsychiatric symptoms and psychiatric disorders in mild Alzheimer's disease and mild cognitive impairment subtypes. *International Psychogeriatrics*, 22, 629–640.

Dubois, B. *et al.* (2010). Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurology*, 9, 1118–1127.

Folstein, M. F., Folstein, S. E. and McHugh, P. (1975). Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.

Gauthier, S. *et al.* (2010). Management of behavioral problems in Alzheimer's disease. *International Psychogeriatrics*, 22, 346–372.

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A. and Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, 140, 566–572.

Lopez, O. L. *et al.* (2003). Psychiatric symptoms vary with the severity of dementia in probable Alzheimer's disease.

The Journal of Neuropsychiatry and Clinical Neurosciences, 15, 346–353.

Lyketsos, C. G. *et al.* (1999). Physical aggression in dementia patients and its relationship to depression. *American Journal of Psychiatry*, 156, 66–71.

Lyketsos, G. C. *et al.* (2011). Neuropsychiatric symptoms in Alzheimer's disease. *Alzheimer's & Dementia*, 7, 532–539.

Lyketsos, C. G. and Miller, S. S. (2012). Addressing the Alzheimer's disease crisis through better understanding, treatment, and eventual prevention of associated neuropsychiatric syndromes. *Alzheimer's & Dementia*, 8, 60–64.

Marin, R. S., Biedrzycki, R. C. and Firinciogullari, S. (1991). Reliability and validity of the Apathy Evaluation Scale. *Psychiatry Research*, 38, 143–162.

Memória, C. M., Yassuda, M. S., Nakano, E. Y. and Forlenza, O. V. (2013). Brief screening for mild cognitive impairment: validation of the Brazilian version of the Montreal cognitive assessment. *International Journal of Geriatric Psychiatry*, 28, 34–40.

Peluso, E. T. P. and Blay, S. L. (2004). Community perception of mental disorders: a systematic review of Latin American and Caribbean studies. *Social Psychiatry and Psychiatric Epidemiology*, 39, 955–961.

Pfeffer, R. I. (1982). Measurement of functional activities in older adults in the community. *Journal of Gerontology*, 37, 323–329.

Reisberg, B., Ferris, S. H. and De Leon, M. J. (1982). The global deterioration scale for assessment of primary degenerate dementia. *American Journal of Psychiatry*, 139, 1136–1139.

Robert, P. H. *et al.* (2002). The apathy inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment. *International Journal of Geriatric Psychiatry*, 17, 1099–1105.

Robert, P. H., Mulin, E., Malléa, P. and David, R. (2010). Apathy diagnosis, assessment, and treatment in Alzheimer's disease. *CNS Neuroscience & Therapeutics*, 16, 263–271.

Rosenberg, P. B., Mielke, M. and Lyketsos, C. G. (2005). Caregiver assessment of patients' depression in Alzheimer disease: longitudinal analysis in a drug treatment study. *American Journal of Geriatric Psychiatry*, 13, 822–826.

Rosenberg, P. B., Mielke, M. M., Appleby, B., Oh, E., Leoutsakos, J. M. and Lyketsos, C. G. (2011). Neuropsychiatric symptoms in MCI subtypes: the importance of executive dysfunction. *International Journal of Geriatric Psychiatry*, 26, 364–372.

Shrout, P. E. and Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428.

Sperling, R. A. *et al.* (2011). Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dementia*, 7, 280–292.

Stella, F. *et al.* (2013). Validation of the Brazilian version of the Apathy Inventory. *International Journal of Geriatric Psychiatry*; in press.

Taragano, F. E. *et al.* (2009). Mild behavioral impairment and risk of dementia. *Journal of Clinical Psychiatry*, 70, 584–592.

Tatsch, M. F. *et al.* (2006). Neuropsychiatric symptoms in Alzheimer disease and cognitively impaired, nondemented elderly from a community-based sample in Brazil: prevalence and relationship with dementia severity. *American Journal of Geriatric Psychiatry*, 14, 438–445.

Truzzi, A. *et al.* (2013). Patterns of neuropsychiatric sub-syndromes in Brazilian and Norwegian patients with dementia. *International Psychogeriatrics*, 25, 228–235.

Ventura, J., Green, M. F., Shaner, A. and Liberman, R. P. (1993). Training and quality assurance with the brief psychiatric rating scale: "the drift buster." *International Journal of Methods in Psychiatric Research*, 3, 221–224.

Wadsworth, L. P. *et al.* (2012). Neuropsychiatric symptoms and global functional impairment along the Alzheimer's continuum. *Dementia and Geriatric Cognitive Disorders*, 34, 96–111.