EXTRA Dementia and Geriatric Cognitive Disorders

Dement Geriatr Cogn Disord Extra 2018;8:306–320
DOI: 10.1159/000490380
Received: November 17, 2017
Accepted: May 24, 2018
Published online: September 10, 2018

Original Research Article

The Neuropsychiatric Inventory-Diary Rating Scale (NPI-Diary): A Method for Improving Stability in Assessing Neuropsychiatric Symptoms in Dementia

Francesca Morganti[a] Alex Soli[b] Paola Savoldelli[a] Gloria Belotti[b]

[a]Department of Human and Social Sciences, University of Bergamo, Bergamo, Italy;
[b]Fondazione Casa di Ricovero Santa Maria Ausiliatrice (CaRiSMA) onlus, Bergamo, Italy

Keywords

Dementia · Neuropsychiatric symptoms · Neuropsychiatric Inventory-Nursing Home · Neuropsychiatric Inventory-Diary · Health-care nursing homes

Abstract

Background: In health-care settings, the use of the Neuropsychiatric Inventory-Nursing Home (NPI-NH) may not always be consistent with the authors' guidelines, which affects its reliability. To avoid this bias, a diary version of the NPI (NPI-Diary) was developed. ***Aims:*** This study aimed to evaluate the psychometric properties (internal consistency and reliability) of the NPI-Diary, and examined its convergence with the NPI-NH. ***Methods:*** Two raters administered the NPI-NH and NPI-Diary to 40 participants with Alzheimer's disease, selected randomly from a hospital's weekly turnover. ***Results:*** The NPI-Diary exhibited adequate internal consistency (total: $\alpha = 0.581$) and test-retest reliability (total: $\rho = 0.711$; $p < 0.01$). The interrater reliability values (ICC) for the NPI-NH and NPI-Diary differed significantly (Total: NPI-NH ICC = 0.506, NPI-Diary ICC = 0.879; Frequency: NPI-NH ICC = 0.51, NPI-Diary ICC = 0.798; Severity: NPI-NH ICC = 0.491, NPI-Diary ICC = 0.809). The convergent validity between the two inventories was also significant (total: $\rho = 0.48$; $p < 0.01$). ***Conclusions:*** The NPI-Diary showed more appropriate validity and reliability compared to the NPI-NH, when administered in a highly variable sample, as is generally the case in the current health-care setting.

Introduction

In the last few decades, the Neuropsychiatric Inventory (NPI) [1] has been largely considered as the standard measure of neuropsychiatric symptoms in several clinical population-based international studies on dementia. The NPI has been found to be capable of

Francesca Morganti
Dipartimento di Scienze Umane e Sociali, Università di Bergamo
Piazzale S. Agostino 2
IT–24129 Bergamo (Italy)
E-Mail francesca.morganti@unibg.it

detecting, quantifying, and tracking neuropsychiatric symptoms in people with dementia [2] during the treatment period, and it consists of a structured assessment of 12 behavioral domains (i.e., hallucinations, agitation, apathy, irritability, disinhibition, etc.) based on an interview with a caregiver. Problematic behaviors within domains are evaluated based on a general screening questionnaire (detailed in several items with yes/no response options) and a rating scale that assesses the frequency and occurrence of the behaviors in each specific domain. Alongside the NPI, a modified version (NPI-Questionnaire [3]) was developed. It consists of a 5-min caregiver report questionnaire that has exhibited good test-retest reliability and optimal convergent validity with the complete NPI.

Although the NPI is a well-standardized inventory, some variability was recently detected in the administration and scoring of the instrument in clinical trials, which could have introduced some measurement biases influencing the psychometric properties of the tool. For example, NPI raters could have been trained to administer the scale in different ways, which could have decreased the intratrial reliability of the instrument [4].

To overcome several shortcomings of the original NPI, a revised version, the Neuropsychiatric Inventory Clinician (NPI-C [5]) has been introduced recently. The NPI-C is now a commonly used scale, as it incorporates clinician ratings to mitigate the reliance on caregiver-provided information. Moreover, it appears to systematize the way in which clinical observation and caregiver reports are integrated. The NPI-C has exhibited high interrater reliability between different examiners [6]. As done in all previous NPI studies, the NPI-C was tested between trained raters who judged the same caregiver's observations [7]. However, the institutional care context is characterized by the presence of many formal caregivers (e.g., physicians, registered nurses, etc.) shifting around inpatients. Each of them could be, from time to time, involved in the NPI assessment. Sometimes, personal factors of the raters, such as distress, exhaustion, anxiety, irritability, and cultural biases, could lead to misinterpretation of the referred neuropsychiatric symptoms. Using the NPI-C, for example, significant disagreements have been found between caregivers' reports and clinician's ratings of agitation, depression, anxiety, apathy, irritability, and aberrant motor behavior. Generally, caregivers tended to overestimate or underestimate the patient's symptoms in comparison with the clinician's rating [8].

An alternative version of the NPI, the Neuropsychiatric Inventory Nursing Home (NPI-NH) [9] was therefore designed for use in hospital settings. Although the NPI-NH was specifically designed to be used in professional health-care settings, the use of this inventory in such contexts (such as a geriatric hospital or an Alzheimer's nursing home) is often not in accordance with the original authors' guidelines. Even when their recommendations were followed, at least the following four critical points can still be detected within daily health-care practices, which have already been discussed partially in some previous works [4, 5]:

1. Some shortcomings on the behavior disorders classification are associated with the lack of staff training in neuropsychiatry or psychogeriatrics, which can influence the inventory rating.
2. Due to raters' tight work schedule, in contrast with guidelines [4], there is little use of the sub-questions provided in the original protocol of the NPI-NH. Instead, an arbitrary evaluation of symptoms is made directly in the summary sheet, based on the general domains.
3. Frequently, there is an insufficient continuous monitoring of the person during the survey period (4 weeks) due to the ordinary and extraordinary staff roster system of a typical professional hospital. Further, because the NPI is a retrospective (up to 1 month) caregiver-informant rating, the problem of recall bias has been raised [7].
4. Often, there is an overlap in the roles of the rater and caregiver, which can result in a completely self-referential evaluation. Further, it has not yet been confirmed that the NPI can be used as a caregiver-administered tool. [8].

Owing to these reasons, a diary version of the NPI, called NPI-Diary was developed and tested in the present study. Using the NPI-NH, we developed an easy-to-use inventory for the detection of neuropsychiatric symptoms to improve the original structure of the NPI-NH. This new tool can be considered as more appropriate for the typical work context of an Italian nursing home.

Aims of the Present Study

The main aim of this study was to develop a modified version of the NPI for use in the daily work in a nursing home and to assess its internal consistency and reliability. The study also aimed to evaluate the convergence of the NPI-Diary with the NPI-NH regarding the detection of neuropsychiatric symptoms in people suffering from dementia. The appropriateness of the NPI-Diary and NPI-NH for nursing home work schedule was also examined.

Methods

Instruments

The Neuropsychiatric Inventory Nursing Home

The Italian version of NPI-NH [10] was used in this study. Raters recorded neuropsychiatric symptoms using a 1–4 scale for frequency and a 1–3 scale for severity for each item in the instrument. In determining the score, screening sub-questions can be used according to each rater's confidence. The scores obtained in each of the two dimensions are then multiplied (Frequency × Severity = 12 points maximum), and the result obtained provides the NPI-NH total score (144 points maximum) derived from the sum of all the behavioral domains.

In an Alzheimer's Nursing Home, according to Italian regional directives, the same nurses can complete the NPI-NH during the multidimensional geriatric assessment. Thus, in contrast with original guidelines [4], every nurse fills out inventories on the basis of his/her observations and, as participants are inpatients, the principal caregiver is missing.

NPI-Diary

The NPI-Diary was developed in the present study by building on the Italian version of the NPI-NH. The team comprised five experts (neuropsychologist, registered nurse, and geriatricians) operating at the CaRiSMA Alzheimer's Nursing Home and at the University of Bergamo. All the experts attending the working group on NPI-Diary have many years of proven experience in the detection of neuropsychiatric symptoms.

The team met repeatedly to make changes to the NPI-NH using the Delphi Panel methodology. Without changing the structure of the NPI-NH questionnaire (12 symptoms domains, each with a 0–12 range), in the first round, for each symptom domain, the experts individuated a series of sub-items (SIs) that were considered appropriate or to be more useful for the assessment of neuropsychiatric symptoms by nursing staff. In particular, for each of the 12 neuropsychiatric symptoms defined in the NPI-NH, the original sub-questions were first drafted by a linguist into sentence form in order to be quantified as SIs. In the second round, all experts eventually proposed new SIs according to their experience, expanding the total number of SIs. At this point, every expert independently indicated the SIs that were duplicate or misleading. If over 50% of experts agreed, those SIs were eliminated. In the third round, every expert proposed possible unifications of related SIs when referring to behaviors that frequently occur together. If over 50% of experts agreed, every expert proposed her/his own version of the merged SI. These merged SI versions were independently rated (1–3) by each expert. The best-rated version of the merged SIs become new SIs. In the last round, each SI was compared to SIs of the other 11 neuropsychiatric symptoms to identify SIs that were

overlapping and were not clearly differentiated from others. If over 50% experts agreed, new versions of the overlapping SIs were proposed, and they were subsequently rated by everybody. The best-rated SI versions were retained as the final SI. At the end of the Delphi Panel, 41 SIs were agreed upon. These SIs fell under the 12 original NPI-NH domains. Appendix presents the SIs for each of the neuropsychiatric symptoms and their scoring system.

The version of the NPI-Diary that was developed was discussed by the team and was considered to be compatible with the everyday work of nurses and the amount of professional resources available each week in the Alzheimer's Nursing Home. Thus, the NPI-Diary was proposed to clinicians and nurses in the working team in order to familiarize them with the immediate reading of the inventory scores before using it.

The following main innovations were introduced in the NPI-Diary:

1. The introduction of SIs for any of the NPI domains aimed to provide nurses with a description of the presence of specific behaviors referable to neuropsychiatric symptoms instead of technical categories. The NPI-Diary presents 41 SIs that are clustered into and scored according to the 12 domains of the original NP-NH items. For each of these items, the final score is the product of the most severe (1–3) SIs' weekly frequency (1–4) during the previous 4 weeks, particularly during the last one. See Appendix for a detailed explanation. In this way, the NPI-Diary maintains the same scoring rules as the classic NPI.
2. The NPI-Diary is structured into a weekly inventory in which each rater records (by monitoring the patient consistently) the presence/severity of each behavior related to a neuropsychiatric symptom each day, for 7 continuous days. This structure supports the possibility of recording the accurate frequency for each neuropsychiatric symptom during the week even if the observation is made by different nursing staff (as is generally the case in an Alzheimer's Nursing Home).

Participants

In total, 40 participants (25 female/15 male) were selected from the inpatients in an Alzheimer's Nursing Home in Northern Italy. All patients were of Caucasian ethnicity and Italian nationality, and their mean age was 81.85 (61–90) years. Their average score on the Mini-Mental State Examination (MMSE) was 5.19 (0–19.4); thus, all patients had some form of cognitive impairment. Twenty of the participants were diagnosed with Alzheimer's disease, according to the criteria of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) [11]. Participants' details have been presented in Table 1.

Ethical Considerations

All participants had moderate to severe dementia, with high vulnerability in their capacity to understand, reasoning, and decision-making. Most of them had cognitive decline and were judged to be unable to consent. Further, nobody had signed consent for them in advance regarding their preference for participation in dementia research. In such cases, for individuals under legal guardianship, the investigator was required to obtain informed consent from the legally authorized representative in accordance with the applicable law. In general, a dual or double consent procedure is advocated, in which the researchers acquire assent of patients who are judged not (fully) capable to consent, in addition to consent by proxy [12]. For each participant of our study, the informed consent and agreement were verbally discussed directly with the person involved, according to their residual understanding ability. Due to memory loss, for each step of the assessment procedure, a new verbal agreement was requested from the patient. They could withdraw from the study at any time without any consequence to their health or care service. Meanwhile, written informed consent was provided by the respective legal executor (support administrator, legal guardian, or the nearest family member) in accordance with Alzheimer Europe's ethical issues guidelines.

Table 1. Sample and rater characteristics

a Sample characteristics (*n* = 39)	
Gender	
Male	15
Female	24
Type of dementia, n	
AD	20
VaD	3
Mixed	8
FTD	3
PPA	2
ARD	1
LBD	1
PDD	1
Age, years	
Inferior-superior	61–90
Mean ± SD	81.85±6.66
Schooling, years	
Inferior-superior	3–18
Mean ± SD	7.92±4.46
MMSE	
Inferior-superior	0–19.4
Mean ± SD	5.19±6.23
b Rater characteristics (*n* = 12)	
Gender	
Female	10
Male	2
Age, years	
Mean ± SD	29.5±5.8
Experience of working in an Alzheimer Nursing Home, months	
Mean ± SD	37.33±65.8
Shifts with inpatients before assessment, n	
Mean ± SD	48.5±45.6

AD, Alzheimer's dementia; VaD, vascular dementia; mixed, mixed dementia; FTD, frontotemporal dementia; PPA, primary progressive aphasia; ARD, alcohol-related dementia; LBD, Lewy body dementia; PDD, Parkinson's disease dementia.

ASST Papa Giovanni XXIII's ethics committee approved the informed consent procedure and the experimental procedure in July 2016 (protocol ID: CaRiSMA001).

Participants' inclusion criteria were the following:

1. Dementia diagnosis according to DSM-V.
2. Cognitive decline assessed in the last 6 months (MMSE <24).
3. Continuing residence at the same Alzheimer's nursing home for at least 2 weeks.
4. Implementation of psychoactive therapies (including anti-dementia drugs) regularly for at least 2 weeks, and they were unmodified until the end of the assessment at T1 (see the procedure section).
5. Stable clinical conditions until the end of the T1 assessment, absence of any sudden variation of behavioral symptoms indicated by medical evaluation or overall differences in the NPI-NH at T1.

Raters

All 12 raters were chosen from the Alzheimer's Nursing Home staff.

Raters' inclusion criteria were the following:

1. Italian registered nurses
2. Working in the Alzheimer's Nursing Home for at least 1 month
3. Having had at least one work shift with the participant during the week before her/his assessment.

Every active nurse from the Alzheimer Nursing Home's staff who met these criteria was offered to join this study after explaining the research procedure and timing. All the contacted nurses were enrolled in the study, and no one dropped out during the process. Table 1b describes the rater sample.

Procedure

According to the sample size guidelines proposed by Donner and Eliasziw [13], all the 40 participants were assessed from July to November 2016. The first step was to determine a baseline for each participant. The baseline evaluation (T0) was conducted by 2 independent raters: R1 and R2 for the NPI-NH, and RD1 and RD2 for the NPI-Diary. R were single raters, while RD were raters' teams made up of 4–5 nurses, alternating during the diary assessment. Every participant was evaluated at T0 with both the NPI-NH and NPI-Diary. For each participant, raters were randomly assigned to the NPI-NH or NPI-Diary condition following simple randomization procedures (computerized random numbers), such that R1–R2 were never a part of the RD1–RD2 teams for the same participant.

The second step was to verify the test-retest reliability of the NPI-Diary. According to previous similar studies [9, 14], at T1, the same raters (R1 and RD1) selected for T0 administered both inventories. On the completion of the procedure for each participant, we had 15 dependent variables (Total score, 12 neuropsychiatric symptoms, Frequency, and Severity) for both the NPI-NH and NPI-Diary.

Data Analysis

With regard to the NPI-NH, the overtime stability of neuropsychiatric symptoms has been established by investigating the absence of a significant difference on comparing (Spearman ρ) T1 and T0.

The interrater reliability of the two inventories (NPI-NH and NPI-Diary) was examined using the data collected at T0 by comparing the respective detection of the NPI-NH and NPI-Diary pair raters with the calculation of the intraclass correlation coefficient (ICC) One-Way Random, "Absolute Agreement" [14]. Moreover, to verify the average reliability difference in the NPI-NH and NPI-Diary, the single measure values obtained for the two inventory parameters were compared (one-tailed *t* test).

The convergent validity of the NPI-Diary was calculated by correlating (Spearman ρ) the scores obtained at T0 in the corresponding NPI-NH dependent variables (Total score, 12 neuropsychiatric symptoms, Frequency, and Severity). The internal consistency of the NPI-Diary for the values collected at T0 was calculated using the Cronbach's α value. The test-retest reliability of the NPI-Diary was calculated by correlating (Spearman ρ) the respective scores recorded for each inventory at T0 and T1. Finally, a factor analysis was conducted on the NPI-Diary using the Principal Component Analysis method. The factors with eigenvalues >1 were selected. The collected data were analyzed using the SPSS software (22.0 version).

Table 2. ICC values for the NPI-NH and NPI-Diary

Item	NPI-NH	NPI-Diary
Total	0.506	0.879
Frequency	0.510	0.798
Severity	0.491	0.809
Delusions	0.539	0.852
Hallucination	0.533	0.856
Agitation/aggression	0.446	0.778
Dysphoria/depression	0.474	0.795
Anxiety	0.292	0.813
Euphoria	0.117	0.778
Apathy	0.047	0.786
Disinhibition	0.580	0.759
Irritability	0.426	0.834
Aberrant motor behavior	0.585	0.884
Sleep and night-time behavior change (12-item version only)	0.249	0.870
Appetite and eating change (12-item version only)	0.582	0.947

Results

One woman was excluded from the 40 participants originally included in the study owing to changes in the psychopharmacological treatment during the recording period. For the remaining 39 participants, the mean time for the NPI-NH administration (one shot) was about 20–25 min, the same found in other studies [8], and that for the NPI-Diary was about 1 min (every day for 1 week). With regard to the NPI-NH, a Spearman's correlation test revealed no significant behavioral changes ($\rho = 0.733$; $p < 0.01$) between the T0 and T1 evaluations.

Interrater Reliability Comparison

According to Cicchetti's recommended ICC values [15], the reliability of the detection between different raters was sufficient for the NPI-NH (ICC = 0.506), and it was excellent for the NPI-Diary total values (ICC = 0.879). Similarly, for Frequency and Severity, the reliability was sufficient for the NPI-NH (ICC = 0.51 and 0.491, respectively) and excellent for the NPI-Diary (ICC = 0.798 and 0.809, respectively).

The reliability values for each neuropsychiatric symptom were sufficient overall for the NPI-NH and excellent for the NPI-Diary, as reported in Table 2. Particularly noteworthy is the reliability value for the Sleep item, which was not acceptable for the NPI-NH (ICC = 0.249) and excellent for the NPI-Diary (ICC = 0.87), as well as those for the items Apathy (NPI-NH = 0.047; NPI-Diary = 0.786), Euphoria (NPI-NH = 0.177; NPI-Diary = 0.778), and Anxiety (NPI-NH = 0.292; NPI-Diary = 0.813). The average difference between the ICC values of the two inventories was significant, with $t(28) = -9.88$, p (one-tailed) <0.001.

Convergent Validity

The convergent validity between the two inventories was significant for the Total value ($\rho = 0.48$; $p < 0.01$), Frequency ($\rho = 0.539$; $p < 0.01$), and Severity ($\rho = 0.509$; $p < 0.01$). The two inventories also showed acceptable convergent validity for all the NPI items.

Internal Consistency

The coherence values for the NPI-Diary items were consistent with those previously found in the literature for the NPI-NH [10], with $\alpha = 0.581$ for the Total score, $\alpha = 0.366$ for Frequency, and $\alpha = 0.632$ for Severity.

Table 3. Matrix of rotated components for NPI-Diary

Component	Factors				
	1	2	3	4	5
Delusions	0.366	0.743	0.154	−0.090	−0.048
Hallucination	0.135	0.122	−0.163	0.011	0.883
Agitation/aggression	0.856	−0.052	0.150	−0.155	0.008
Dysphoria/depression	0.254	0.234	0.677	0.051	−0.172
Anxiety	−0.155	0.699	−0.011	0.451	−0.098
Euphoria	−0.031	−0.031	−0.001	0.937	0.061
Apathy	−0.351	0.000	0.292	0.281	0.439
Disinhibition	0.729	0.128	0.360	−0.205	0.054
Irritability	0.889	0.115	−0.033	0.217	−0.027
Aberrant motor behavior	0.207	0.105	−0.587	0.166	−0.342
Sleep/night behavior change	0.228	0.135	0.555	0.057	−0.085
Appetite/eating change	−0.020	0.787	0.175	−0.155	0.325

Test-Retest Reliability

The NPI-Diary measurements at T0 and T1 showed significant reliability scores for the Total value ($\rho = 0.711$; $p < 0.01$), Frequency ($\rho = 0.732$; $p < 0.01$), and Severity ($\rho = 0.737$; $p < 0.01$).

Factor Analysis

The main components of the instrument were analyzed. The first examination of the scree plot suggested the presence of 5 components. Even the Mineigen criterion suggested a 5-component solution that can explain 69.387% of the total variance. Subsequent analyses confirmed the presence of 5 unrelated components. With an oblique rotation (oblimin criterion), a maximum correlation between the components of $r = |0.19|$ emerged; therefore, an orthogonal rotation (varimax criterion) was conducted. Only Factor 1 ("hyperactivity"), which included the items Agitation/aggression ($r = 0.856$), Irritability ($r = 0.889$), Motor activity aberrant ($r = 0.207$), and Disinhibition ($r = 0.729$), appeared comparable to the factors observed in the recent work of Baranzini et al. [10]. Factor 2 ("anxiety") included the items Delusions ($r = 0.743$), Anxiety ($r = 0.699$), and Appetite and eating disorders ($r = 0.787$). Factor 3 ("depression") included the items Depression ($r = 0.677$) and Sleep disorders ($r = 0.555$). Factor 4 ("euphoria") only comprised the item Euphoria ($r = 0.937$). Factor 5 ("withdrawal") included the items Apathy ($r = 0.439$) and Hallucinations ($r = 0.883$). The components of the rotated matrix have been presented in Table 3.

Methodological Considerations

The construct validity, consistency, and reliability of the NPI-NH re-test, widely demonstrated for both the original inventory [1, 16] and for its Italian version [10], have not been discussed in the present study. Nevertheless, some doubts about the reliability of inexperienced raters among the nursing home staff had already been pointed out. Wood et al. [9] showed a moderate correlation with measurements of trained researchers, limiting the efficacy of the NPI-NH to monitor behavioral changes. The present study seems to confirm this finding. When administered by not complying with original guidelines, the reliability values of the NPI-NH decreased to the limit of sufficiency in almost all the examined parameters. In particular, for some neuropsychiatric symptoms (Sleep, Apathy, Exhilaration, and Anxiety), the ICC values for the interrater reliability did not reach the optimal threshold (0.4)

considered in the present study, as suggested by Cicchetti [15]. It is important to also consider that, more recently, an even more selective threshold (0.7) has been recommended for such scales [17]. This could be translated as a real bias of the measurements. Our study also highlighted the fact that the measurement made with the NPI-NH and with a different manner of administration, as in the NPI-Diary, revealed significantly different average values, with $t(76) = -4{,}50$, p (one-tailed) <0.01. Thus, the utilization of the NPI-NH seems to lead to an underestimation of the incidence of neuropsychiatric symptoms in patients. Moreover the NPI-Diary showed how the evaluation of neuropsychiatric symptoms can be disentangled from one specific rater, and it can be conducted during an entire week without modifying the structure of the NPI-NH (including the scoring).

Discussion and Conclusions

As psychometric reliability is a main condition required in the last Lombardy regional decrees on health care [18], which impose the use of "validated instruments for assessments," the more obvious solution could be to respect the original guidelines for the administration of the NPI-NH. However, guidelines often clash with the logistical, organizational, and economic constraints of the daily activity in an Alzheimer's Nursing Home.

The present work revealed the insufficient interrater reliability of the NPI-NH when using it without complying with the original guidelines within the limitations imposed by the daily routine of Italian nursing homes. This finding has important implications for the clinical and the experimental detection of neuropsychiatric symptoms and strongly discourages the improper use of the NPI-NH, as it can compromise its psychometric reliability.

The alternative proposed here, the NPI-Diary, has therefore been designed as a tool that can bypass the inherent obstacles encountered in a health-care institution, without altering the structure of the NPI-NH, which remains the most popular and well known tool for the assessment of neuropsychiatric symptoms.

On being used in a real Alzheimer's Nursing Home workflow within a residential context, the NPI-Diary showed optimal psychometric properties and a good interrater reliability, simultaneously allowing an accurate detection of neuropsychiatric symptoms.

Our study revealed different ICC values for the two inventories. These data provide us with evidence on how the involvement of different raters in the NPI-NH can result in different evaluation profiles, while the NPI-Diary can be more consistent during weekly observations even if they are conducted by different raters.

Within tight deadlines, the NPI-Diary enables simple and rapid symptom detection that is preferably subdivided among the staff during the weekly turnover. Often, in fact, the questionnaire is completed in several stages, which are handled and shared by the entire professional team. Moreover, generally nurses and operators use the NPI-Diary inventory to also identify early signs of behavioral disorders, attributing the severity according to the level of assistance needed rather than using a more objective method. Some raters involved in the present trial positively evaluated the SIs of the NPI-Diary as easy for less expert raters to use, thus suggesting the considerable training potential of the NPI-Diary.

Compared to previous versions of the NPI, the NPI-Diary showed more detailed detection of neuropsychiatric symptoms, which allows professionals to plan therapeutic interventions more accurately. For example, the SI "He avoids some places/people with an unjustified fear" instead of the summary item "Anxiety" can lead to targeted environmental intervention, instead of the use of generic pharmacological therapy.

The factor analysis results, with 5 factors for just 12 items, may justify the modest internal consistency of the inventory (already observed in previous forms of the NPI). The macro-

category of neuropsychiatric symptoms gathers behaviors that are quite heterogeneous, placing the NPI-Diary between multidimensional scales.

The major limitation of the NPI-Diary could be the inability of unique administration, as it needs 7 days of continuous observation. However, the detection of neuropsychiatric symptoms in a nursing home rarely happens in emergency situations and can usually be properly planned in advance. Thus, from our experience, the NPI-Dairy appears to fit well with the routine of the Alzheimer's Nursing Home staff.

Another limitation of the present study is the representativeness of the sample, which was limited to the Italian geriatric population with dementia [19]. Further investigations would be desirable to extend the applicability in other areas, such as for mentally healthy individuals.

Conclusions

In a sample of geriatric patients with dementia, the NPI-Diary proved to be capable of restoring credibility to the process of assessing neuropsychiatric symptoms through the use of its SIs. It also exhibited optimal psychometric properties. It can therefore be considered as an improvement of the NPI-NH as it is an easy-to-use inventory for the assessment of neuropsychiatric symptoms in the typical work context of a nursing home.

The NPI-Diary is suitable for use in the broad-spectrum screening of the major behavioral alterations in dementia patients, in conducting repeated assessments during patients' internship, and in the evaluation of neuropsychiatric symptoms for multicenter or repeated measures for experimental purposes. In particular, the NPI-Diary is suitable for use by healthcare professionals working in shifts, as they often lack highly specialized training in assessing behavioral disorders. In this regard, future investigations could verify the reliability of the NPI-Diary when it is implemented by unspecialized nurses or even informal caregivers, for example inside home environment.

Finally, we recommend that NPI-Diary raters share guidelines before starting the evaluation, in order to enable an in-depth analysis of the psychiatric issues identified. Accordingly, we recommend that appropriate coaching and training be provided to nursing staff prior to the administration of the NPI-Diary.

Acknowledgements

The authors would like to thank all the CaRiSMA Foundation staff for their cooperation and willingness to support this study, particularly Fabrizio Lazzarini, for allowing us to conduct the study; Estella Cuni, for supporting the experimental procedure; and Sylvie Andreoletti, Ilaria Brevi, Cristian De Blasi, Monica Egman, Ilaria Milesi, Nava Maria, Ludmila Odajiu, Erika Ongaro, Aurora Pascu, Valentina Personeni, Mara Suardi, and Matteo Torreano for collaborating in the experimental trial.

Disclosure Statement

The authors declare that there are no conflicts of interest regarding the publication of this article.

Dementia and Geriatric Cognitive Disorders EXTRA

Appendix

The NPI-Diary
Items, sub-items, use guidelines and scoring procedures

NPI-Diary										
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Best value	Frequency	Score
Deliri *(Delirium)*										
1 Accusa altri di essere bugiardi, di avercela con lui (*She/he accuses others of lying or to have it in for her/him*)										
2 Dice di essere derubato, rovinato, avvelenato (*She/he says that she/he was robbed, ruined, or poisoned)*										
3 Dice di esser tradito, abbandonato, non curato (*She/he claims to have been betrayed, abandoned, or untreated*)										
4 Dice di trovarsi in un altro luogo, di dover fare altre cose (*She/he claims to have been in another place or having to do other things*)										
5 Dice che persone vive sono morte e/o viceversa (*She/he says that living people were dead and/or vice versa*)										
Allucinazioni *(Hallucinations)*										
1 Sente delle voci inesistenti, risponde ad esse (*She/he hears non-existent voices and responds to them*)										
2 Vede cose inesistenti, interagisce con esse (*She/he sees unreal things and interacts with them*)										
3 Riferisce odori o sapori irreali (*She/he reports unreal smells or tastes)*										
Aggressività/Agitazione *(Aggressivness/Agitation)*										
1 Si oppone attivamente a proposte di cura, si dimena (*She/he actively opposes nursing proposals; she/he squirms)*										
2 Grida, insulta o bestemmia per inezie (*She/he shouts, offends, or blasphemes for trifles*)										
3 Sbatte le porte, lancia, rompe o percuote oggetti (*She/he slams the door and throws, breaks, or strikes objects*)										
4 Morde, graffia, spinge, picchia (incluso sé stesso) (*She/he bites, scratches, pushes, and hits her/himself or others*										

Dement Geriatr Cogn Disord Extra 2018;8:306–320
DOI: 10.1159/000490380

Dement Geriatr Cogn Disord Extra 2018;8:306–320
DOI: 10.1159/000490380

Dementia and Geriatric Cognitive Disorders EXTRA

NPI-Diary										
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Best value	Frequency	Score
Disforia/Depressione *(Disphoria/Depression)*										
1 Piange o esprime tristezza (in probabile assenza di dolore fisico) (*She/he cries or expresses sadness (in likely absence of physical pain)*)										
2 Dice di sentirsi solo, dimenticato, fallito (*She/he says he feels alone, forgotten, or failed*)										
3 Dice di esser cattivo, meritevole di punizione (*She/he claims to be bad or deserving of punishment*)										
4 Dice di non aver speranze, ricerca la morte (*She/he expresses absence of hopes; she/he is looking for death*)										
Ansia *(Anxiety)*										
1 Allarmato, si spaventa per stimoli irrilevanti (*She/he is alarmed or startled by irrelevant stimuli*)										
2 In attesa costante di qualche visita/evento (*She/he is constantly waiting for some visit/event*)										
3 Lamenta palpitazioni, sudore, nausea, vertigini (causa non nota) (*She/he complains about palpitations, sweating, nausea, and dizziness (cause unknown)*)										
4 Evita alcuni luoghi/persone per paura immotivata (*She/he avoids certain places/people for groundless fear*)										
5 Ricerca esagerata di attenzione, pedinamento (*She/he exhibits exaggerated attention pursuit or shadowing*)										
Esaltazione/Euforia *(Elation/Euphoria)*										
1 Ride, scherza, canta, urla inappropriatamente (*She/he laughs, jokes, sings, or cries inappropriately*)										
2 Esagera in azioni incompatibili con i suoi limiti fisici (*She/he exaggerates actions that are incompatible with her/his physical limitations*)										
Apatia/Indifferenza *(Apathy/Indifference)*										
1 Seduto a lungo (sveglio), poco coinvolgibile, senza interessi (*She/he sits down (awake), is shortly engageable, or is without interest*)										
2 Silenzioso, non chiede, risponde con poche parole (*Being silent, she/he does not ask questions or responds with few words*)										
3 Non mostra emozioni di fronte a stimoli affettivi (*She/he does not show emotions when encountering emotional stimuli*)										

Dement Geriatr Cogn Disord Extra 2018;8:306–320
DOI: 10.1159/000490380

E X T R A
Dementia
and Geriatric
Cognitive Disorders

NPI-Diary										
	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Best value	Frequency	Score
Disinibizione *(Disinhibition)*										
1 Usa un linguaggio scurrile o volgare (*She/he uses obscene or vulgar language*)										
2 Avances sessuali, verbali e/o fisiche, eccessive, si sveste (*She/he exhibits advances that are sexual, verbal, and/or physical exaggerated; she/he takes off her/his clothes*)										
3 Si avvicina e conversa o si prende troppe libertà con sconosciuti (*She/he approaches and converses with or talks too freely with strangers)*										
Irritabilità/Labilità *(Irritability/Lability)*										
1 Si arrabbia con chi lo assiste (nell'igiene, vestizione) (*She/he gets angry with the caregiver (during hygiene activities and dressing*)										
2 Si innervosisce facilmente con tutti, alza la voce (*She/he easily get nervous in the presence of others; she/he raises her/his voice*)										
3 Si lamenta di ogni cosa, borbotta, continua insofferenza (*She/he complains about everything; she/he mutters and continually expresses intolerance*)										
4 Si irrita se interrotto in attività afinalizzate (es., wandering) (*She/he gets irritated if stopped from engaging in non-task-focused activities (e.g. wandering*))										
Attività motoria *(Motor disturbance)*										
1 Vagabondaggio continuo senza meta (wandering) (*She/he wanders continously and aimlessly*)										
2 Tenta di uscire dallo spazio protetto (*She/he tries to exit the protected space*)										
3 Accumula e/o nasconde cose senza motivo (*She/he collects and/or hides things without any reason*)										
4 Si affaccenda in gesti afinalizzati (es. manipolare oggetti) (*She/he remains engaged in goalless gestures (e.g. manipulating objects)*)										
Sonno *(Nightime behaviors)*										
1 Insonne parziale o totale (non se si alza per il bagno) (*She/he experiences partial or total sleeplessness (excluding when she/he gets up to go to the bathroom*))										
2 Ipersonne durante il giorno (oltre il riposo pomeridiano) (*She/he is hypersomnic during the day (over the lunch break*)										
Disturbi dell'appetito e dell'alimentazione *(Appetite/Eating)*										
1 Scarso appetito, si lamenta dei piatti, rifiuta, mangia pochi cibi (*She/he has a poor appetite, complains about the dishes, refuses to eat, or eats few foods*)										
2 Richiede troppo cibo, mastica voracemente (anche oggetti) (*She/he requires too much food; she/he chews food or objects voraciously*)										

Compilation Instructions

If the listed behaviors are observed every day, indicate their manifestation in the appropriate box and then insert appropriate values according to the following rating scale:

1 for mild (little disturbing for the person and those nearby; it is not necessary to intervene, or only a brief intervention is required).

2 for moderate (relevant action is required, such as a long and repeated interview and/or other compensatory strategies).

3 for severe (problematic both for the individual and for those nearby; the intervention seems to be poor or ineffective).

If during the same day several episodes of varying severity are repeated, consider only the highest value detected (e.g., if you observe an intensity of 2 in the morning but 3 in the afternoon, record a 3 in the same cell; if a new registration in the evening is a 2, retain the 3 rating).

Scoring Instructions

Severity

Once you have completed the diary, for every cluster of sub-items (corresponding to one of the 12 domains), consider the most serious encountered absolute value (1–3) for the week and report it in the corresponding "Best value" cell. Only if no episodes are reported in a domain in the past week but caregivers remember the occurrence of one or more episodes in the previous 3 weeks, the value (1–3) of the major severe episode can be directly reported in the "Best value" cell, or the "Best value" is recorded as 0.

Frequency

For each domain, only consider the sub-item (e.g., a row) with more severe episodes (for a row with more than 3 values, if a value of 3 is absent or is present in more than one row, consider values 2, at least values 1). Only for this sub-item, count how many days in the row the episode appears with any severity (count all nonempty cells in the row) and then mention the respective value in the "Frequency" box according to the following classification:

1 = less than once a week (it was not detected in the last week, but it was observed in the last month). This score is used when no episodes are reported on the diary in a domain's cluster, but raters know about episodes that have occurred previously (in that case, refer to the above for the corresponding severity rating).

2 = once a week

3 = two or more times a week (up to a maximum of 6 and not each day of the week)

4 = every day

Otherwise the frequency is 0.

References

1 Cummings JL, Mega M, Gray K, Rosenberg-Thomson S, Carusi DA, Gornbein J: The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994;44:2308–2314.
2 Lyketsos CG, Carrillo MC, Ryan JM, Khachaturian AS, Trzepacz P, Amatniek J, Cedarbaum J, Brashear R, Miller DS: Neuropsychiatric symptoms in Alzheimer's disease. Alzheimers Dement 2011;5:532–539.
3 Kaufer D, Cummings JL, Ketchel P, Smith V, MacMillan A, Shelley T, Lopez OL, DeKosky ST: Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. J Neuropsychiatry Clin Neurosci 2000;12:233–239.
4 Connor DJ, Sabbagh MN, Cummings JL: Comment on administration and scoring of the Neuropsychiatric Inventory in clinical trials. Alzheimer Dement 2008;4:390–394.
5 De Medeiros K, Robert P, Gauthier S, Stella F, Politis A, Leoutsakos J: The Neuropsychiatric Inventory-Clinician rating scale (NPI-C): reliability and validity of a revised assessment of neuropsychiatric symptoms in dementia. Int Psychogeriatr 2010;22:984–994.
6 Zaidi S, Martin GK, Jos de Jonghe FM: Clinician and caregiver agreement on neuropsychiatric symptom severity: a study using the Neuropsychiatric Inventory-Clinician rating scale (NPI-C). Int Psychogeriatr 2014;26:1139–1145.
7 Stella F, Forlenza OV, Laks J, Pires de Andrade L, de Castillo Cacao J, Govone JS, de Medeiros K, Lyketsos CG: Caregiver report versus clinician impressions: disagreements in rating neuropsychiatric symptoms in Alzheimer's disease patients. Int J Geriatr Psychiatry 2015;30:1230–1237.
8 Lay CK: The merits and problems of Neuropsychiatric Inventory as an assessment tool in people with dementia and other neurological disorders. Clin Interv Aging 2014;9:1051–1061.
9 Wood S, Cummings JL, Hsu MA, Barclay T, Wheatley MV, Yarema KT , Schnelle JF: The use of the neuropsychiatric inventory in nursing home residents. Characterization and measurement. Am J Geriatr Psychiatry 2000;8:75–83.
10 Baranzini F, Grecchi A, Berto E, Colombo D, Costantini C, Ceccon F, Cazzamalli S, Poloni N, Callegari C: Factor analysis and psychometric properties of the Italian version of the Neuropsychiatric Inventory-Nursing Home in an institutionalized elderly population with psychiatric comorbidity (in Italian). Riv Psichiatr 2013;48:335–344.
11 American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, ed 5. Washington, American Psychiatric Association, 2013.
12 Meulenbroek O, Vernooij-Dassen M, Kessels RPC, Graff MJL, Sjögren MJC, Schalk BWM, Hoogsteen-Ossewaarde ME, Claassen JAHR, Melis RJF, Olde Rikkert MGM: Informed consent in dementia research. Legislation, theoretical concepts and how to assess capacity to consent. Eur Geriatr Med 2010;1:58–63.
13 Donner A, Eliasziw M: Sample size requirements for reliability studies. Stat Med 1987;6:441–448.
14 Landers RN: Computing intraclass correlations (ICC) as estimates of interrater reliability in SPSS. Winnower DOI: 10.15200/143518.81744.
15 Cicchetti DV: Guidlines, criteria and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6:284–290.
16 Cummings JL: The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. Neurology 1997;48:10–16.
17 Barrett P: Conventional interrater reliability: definitions, formulae, and worked examples in SPSS and Statistica. Tech Whitepap Ser 2001;1:1–28.
18 Regione Lombardia, DGR 2014;2569;allegato 1.
19 Binetti G, Mega M, Padovani A, Rozzini L, Bianchetti A, Trabucchi M, Cummings JL: Behavioral disorders in Alzheimer disease: a transcultural prospective. Arch Neurol 1998;55:539–544.