Published in final edited form as:
Am J Geriatr Psychiatry. 2013 July ; 21(7): 607–622. doi:10.1016/j.jagp.2012.10.027.

Subscale Validation of the Neuropsychiatric Inventory Questionnaire (NPI-Q): Comparison of ADNI and NACC Cohorts

Paula T. Trzepacz, MD[1,2], **Andrew Saykin, PhD**[2], **Peng Yu, PhD**[1], **Phani Bhamditipati, MS**[1], **Jia Sun, MS**[1], **Ellen B. Dennehy, PhD**[1,3], **Brian Willis, PhD**[1], and **Jeffrey L. Cummings, MD**[4] **for the Alzheimer's Disease Neuroimaging Intiative**[5]

[1]Eli Lilly and Company, Indianapolis, IN

[2]Indiana University School of Medicine, Indianapolis, IN

[3]Purdue University, West Lafayette, IN

[4]Cleveland Clinic, Lou Ruvo Center, Las Vegas, NV

Abstract

Background—Neuropsychiatric symptoms (NPS) are common in Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD) and often measured using the Neuropsychiatric Inventory (NPI). Development of validated subscales that measure clinically meaningful symptom clusters would improve capacity for individualized treatment and assessment of treatment interventions. We report preliminary validation of three NPI Questionnaire (NPI-Q) subscales derived from examination of the existing exploratory literature and clinical knowledge.

Methods—The validity of subscales that assess Frontal, Agitation/Aggression, and Mood symptoms (based on NPI-Q-10 item scores) was ascertained by comparison of cross-sectional data from amnestic MCI and AD dementia cases from the National Alzheimer's Coordinating Center (NACC) and Alzheimer's Disease Neuroimaging Initiative (ADNI) databases. The statistical approach was confirmatory unrotated principal component analysis.

Please address correspondence to Dr. Paula T. Trzepacz, Neurosciences Research, Lilly Corporate Center, Indianapolis, IN 46285; PTT@lilly.com; 317 433 5391.

[5]Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete list of ADNI investigators is available at http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Disclosures: Drs. Trzepacz, Dennehy, Yu and Willis, and Mr. Bhamditipati and Mr. Jia Sun are full-time salaried employees and minor shareholders of Eli Lilly and Company.

Dr. Saykin receives support from the NIH National Institute on Aging (R01 AG19771 as PI; P30 AG10133 Indiana ADC as Imaging Core leader; U01 AG024904 and RC2 AG036535 as ADNI Genetics Core leader; and U01 AG032984 Alzheimer's Disease Genetics Consortium as site PI).

Dr. Cummings has provided consultation to Abbott, Acadia, Acerra, ADAMAS, Anavex, Astellas, Avanir, Baxter, Bristol-Myers Squibb, Eisai, Elan, EnVivo, Forest, Genentech, GlaxoSmithKline, Janssen, Lilly, Lundbeck, Medivation, Medtronics, Merck, Neurokos, Novartis, Pfizer, Prana, QR, Sonexa, Takeda, and Toyama pharmaceutical companies and Bayer, Avid, GE, MedAvante, Neurotrax, and UBC Assessment Companies. Dr. Cummings owns stock in ADAMAS, Prana, Sonexa, MedAvante, Neurotrax, Neurokos, and QR pharma and owns the copyright of the Neuropsychiatric Inventory scale. He is a speaker/lecturer for Eisai, Forest, Janssen, Novartis, Pfizer, Lundbeck.

Results—ADNI contributed 103 MCI, 90 MCI-converters and 112 AD dementia cases while NACC had 1042, 763, and 3048. Baseline mean age was higher in NACC (74.6 vs 75.7). Patients in NACC were significantly more impaired at last visit on MMSE (mean scores 19.5 vs 22.4) and NPI-Q-10 (5.0 vs 4.3), as well as for each of the three subscales (NPI-Q-4-Frontal, NPI-Q-4-Agitation/Aggression, and NPI-Q-3-Mood than ADNI (at month 24). Medians were not different for Agitation/Aggression or Mood subscales, however. Each item on all scales contributed variance in PCA Pareto plots. All items in Factor (F) 1 for each scale projected in a positive direction on biplots (coherence), while F2 and F3 items showed more spatial separation (independence). Scale analyses showed remarkable similarities between ADNI and NACC cohorts for factor loadings and spatial patterns of item projections, though factor item identities varied somewhat, especially beyond F1.

Conclusions—The similar pattern of results across two cohorts of patients support the validity of these constructs. These subscales are worthy of further psychometric evaluation in patients with MCI and AD dementia and preliminary application in clinical settings.

Keywords

Neuropsychiatric symptoms; Alzheimer's disease; Neuropsychiatric Inventory; subscales

Objective

Alzheimer's Disease(AD) is looming as a public health problem of epidemic proportions and it comprises cognitive, language and other neuropsychiatric symptoms (NPS) (1). The disease process begins decades prior to the diagnosis of dementia due to AD, and NPS are apparent even in the Mild Cognitive Impairment (MCI) stage where certain NPS predict conversion to AD dementia (2,3). Noncognitive NPS include affective, psychotic and behavioral symptoms and are associated with decreased function and quality of life, increased caregiver burden, and institutionalization (4-7).

The Neuropsychiatric Inventory (NPI) is the most widely used rating scale for NPS in patients with dementias and other neurological disorders (8,9). The original version included 10 symptom domains: delusion, hallucination, depression, anxiety, agitation/aggression, euphoria/elation, disinhibition, irritability/lability, apathy, and aberrant motor activity. It was later amended adding items assessing night-time behavior disturbances and changes in appetite/eating behaviors (the NPI-12). The NPI is administered during an interview with a caregiver who has daily contact with the patient. A positive response to each screening question is followed by a series of subquestions which are rated for frequency and severity of that symptom domain per the caregiver. Each positive domain is also rated for its severity of caregiver distress. The NPI has satisfactory validity and reliability in outpatient settings (8) with high scores for content validity across all items, and acceptable concurrent validity compared to standard instruments used to measure NPS in neurologically impaired patients.

Several alternate versions of the NPI are used in research and other settings. The NPI-Nursing Home version (NPI-NH) was developed for ratings by caregivers in institutional settings (10). The NPI Questionnaire (NPI-Q) is a briefer version which includes ratings of domain severity only (not frequency) and was developed for use in general clinical practice or epidemiological research settings (11). The NPI-Q is a caregiver-report questionnaire, versus a clinician administered caregiver interview, and can be completed in 5 minutes or less. The NPI-Q has good test-retest reliability and convergent validity, correlating with the full NPI at 0.91 (11). More generally, the measure has acceptable psychometric properties and is a good brief measure of NPS and associated caregiver distress (11).

There have been extensive efforts to identify clusters or syndromes of NPS that may identify meaningful subgroups of patients, or that might predict treatment course or outcomes for patients with dementia. The Alzheimer's Association's Professional Interest Area committee on NPS in AD reviewed diagnostic and translational data for five proposed syndromes – depression, psychosis, agitation, apathy and sleep (12). Some clusters have been studied in great detail, in particular domains that address symptoms of psychosis (13) and verbal and physical agitation and aggression (14).

There are a number of published exploratory analyses of the NPI, in diverse samples that include different types and acuity of dementias. These typically yield solutions ranging from 3-5 factors. Despite some differences among reports, common factors emerge for symptom clusters that encompass items describing agitation and aggression, psychosis, and mood, with apathy usually either loading alone on a factor or combined with mood. Most studies have relied on exploratory factor analysis or principal component analysis of NPI scores rather than confirmatory modeling of symptom domains or psychometric validation of derived subscales. One exception is the development of an agitation-aggression subscale within the NPI-NH composed of the agitation/aggression, irritability/lability, disinhibition, and aberrant motor behavior domains (10). This subscale accounted for 60% of the variance in Cohen-Mansfield Agitation Inventory total scores (10). This was a small study (n=69) and to our knowledge, the scale has not been utilized in other research applications.

In the interest of furthering our knowledge of neurological substrates and management of NPS syndromes, we analyzed NPI-Q data from MCI and AD dementia patients in two publically available databases, National Alzheimer's Coordinating Center (NACC) and Alzheimer's Disease Neuroimaging Initiative (ADNI). We developed NPI-Q subscales for agitation/aggression, mood, and frontal syndromes, whose compositions were determined based on an overall interpretation of the exploratory factor analysis literature and knowledge of functional neuroanatomy. We hypothesized that the NPI-Q subscales measure valid NPS constructs and utilized findings from two large cohorts to demonstrate that point. We used PCA to ascertain how the components of these subscales related to each other, including if they contributed separately to the subscale as well as whether the items reflected a common phenotypic dimension. Comparisons between two cohorts were used to further confirm validity of the subscales. We also applied the Gabriel biplots method to PCA findings, a well-described method in literature (15,16), to visually interpret multivariate data and discern item relationships between factors.

Methods

Subjects and Procedures

ADNI is a multisite, multi-study program funded by a public-private partnership to investigate whether the combination of neuroimaging, biological markers, and clinical and neuropsychological assessments can accurately track disease progression in AD. We analyzed the ADNI data released in Sept 2010. At that time, 819 subjects had been recruited, with 229 elderly controls, 402 amnestic MCI (aMCI) patients, and 188 AD dementia patients. Those with aMCI who converted to dementia were 172 of the 402 cases. Subjects were assessed every 6 months for the first 2 years and every 12 months thereafter. Inclusion criteria for AD dementia, aMCI, and control subjects can be found at ADNI Website (http://www.loni.ucla.edu/ADNI). We analyzed the NPI-Q and Mini Mental State Exam (MMSE) data collected in AD dementia and aMCI patients at the 24-month visit, along with their demographic data at entry into the study, using the diagnostic assignments made by ADNI. A subset of those aged >50 and with NPS (NPI-10>0) were selected for this report.

We analyzed data provided by NACC from a longitudinal study at 29 National Institute on Aging (NIA)-funded Alzheimer's Disease Centers (ADCs) across the United States. Data were transferred to our study site in June 2011. Data are collected at approximately annual visits for up to 7 years, in three samples of older adults: normal, cognitively impaired, and with various dementias including due to AD. In total, there were 8462 subjects. NPI-Q and MMSE scores were measured at the last visit and demographic data at entry into the study. We first excluded nonMCI cognitively impaired patients, nonamnestic MCI and normal persons before analyzing the NPI-Q data in AD dementia, MCI converters, and aMCI patients collected at their last visit (N Visits: Mean=4, Min=2, Max=7). However, because some diagnoses could change across visits, we assigned diagnoses based on each individual's progression path using the following algorithm:

1. If the patient did not start with AD (or AD contributing), was at least aMCI or nonamnestic MCI once, and ended with AD (or AD contributing), then that patient was categorized as an MCI converter.
2. Otherwise, we used the diagnosis at the last visit. However, we excluded patients whose last status was aMCI, but had AD (or AD contributing) in the previous visit (e.g., progression of AD-AD-aMCI).

After this process, we had 3710 AD dementia, 978 MCI Converter, 1792 aMCI, 324 Cognitively Impaired without MCI, 496 nonamnestic MCI, and 550 normal patients. The total number of AD dementia, MCI Converter, and aMCI cases that were analyzable in our NACC cohort was 6480, from which a subset aged >50 and with NPS was selected.

While both databases included MCI and AD dementia patients, the ADNI cohort is, in general, less advanced in disease stage than is the NACC cohort, and analyzing both databases lends more breadth to the understanding of how subscales perform across these patients. In both samples we chose to analyze a later point in time within the available longitudinal data in anticipation of capturing a greater severity of symptoms. We chose to include both aMCI and AD dementia (including MCI converters) so we had a larger range of scores to analyze [hereafter aMCI is referred to only as MCI]. We hypothesized that the scale and subscales should perform similarly across the two databases with the caveat that the proportion of more advanced cases in the NACC database might influence some aspects of the results. We acknowledge that some of the MCI cases might not have eventually progressed to AD dementia but since we chose later visits to analyze, it raised the likelihood that longitudinal evaluations would lend more clarity to the diagnoses.

Measures

The **MiniMental State Exam (MMSE)** (17) was used as a general proxy for staging illness severity. Scores range from 0 to 30 with those under 24 reflecting cognitive impairment.

The **Neuropsychiatric Inventory Questionnaire (NPI-Q)** is a 10- or 12-item scale, derived from the original NPI-10 and 12-item versions(9), and has a high correlation with the original NPI (r= 0.91 overall, r= 0.90 for subgroup with high MMSE scores and r=0.95 for low MMSE scores) (11). Each NPI-Q item is rated by the caregiver as 0-3 points according to levels of increasing severity and the score maximums are determined by multiplying the number of items by 3. Directions for completing the questionnaire, and anchor points for ratings, are provided to guide the respondent. The NPI-Q continues to use screening questions for each behavioral symptom domain, although they are somewhat shortened. If “yes” is circled in response to the stem question, the person is guided to rate presence and severity of behaviors present in the past 4 weeks.

The items selected for the subscales are based on literature for descriptive population data, a composite impression from exploratory factor analyses including a very large sample by Aalten et al (18,19), as well from known phenotypic and neuroanatomical relationships among symptoms drawn from the larger neuropsychiatric literature. Items for NPI-Q-4-Agitation/Aggression, NPI-Q-3-Mood, and NPI-Q-4-Frontal subscales are listed in Table 2. While apathy and psychosis (delusions and hallucinations) are common NPS in patients with AD, subscales for these two dimensions are not evaluated in this report because they have insufficient number of items for analysis.

Statistical Procedures

Data were analyzed using JMP v. 9 and R v. 2.13.0 software. Cases were selected for this report only if they had at least one point on the NPI-Q-10. Demographic, MMSE and scale score variables were evaluated using descriptive statistics and boxplot distributions. NPI-Q scores from ADNI and NACC databases were analyzed separately to allow for assessing impressions of pattern similarity or difference. Analysis of each NPI-Q scale or subscale was conducted on the pooled AD dementia, MCI Converter, and aMCI patients whose age was >50 and who had at least one non-zero item score on that scale/subscale at the designated visit in the respective database. Brown-Forsythe test for equality of variances was applied to all variables. Based on these results, age and MMSE were compared between cohorts using Welch's ANOVA test for unequal variances for NPI-Q-10 and all subscale groups. Sex was compared using Pearson's Chi-squared test. NPI-Q scores were compared using ANOVA. Wilcoxon rank-sum test was applied to compare medians between ADNI and NACC samples for boxplots of rating scales.

PCA is mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. Although rotated PCA yields more easily interpretable factors (components), we did not conduct rotation. Rather, our predefined goal was to compare the principal component structures of these NPI-Q scale/subscales in two different datasets, where each principal component is expressed by a linear combination of NPI-Q items and each principal component is orthogonal to the others.

Since our purpose was to fully evaluate the NPI, we examined factors generated by PCA equal to the number of items in each scale/subscale. We determined and graphed the variance explained by each factor for each NPI-Q scale/subscale using Pareto plots. Using JMP software, we then produced visual representations (Gabriel biplots) of item loadings of the first three factors for each scale. This produced a series of 2-dimensional graphs showing item projections between pairs of the first three factors. Only three factors were graphed, because these factors explained at least half of the variance in each particular scale/subscale. This allowed visual interpretation of spatial relationships among the items (15,16) to evaluate their contributions to the scale/subscale for each cohort.

Results

4853 individuals (75%) of the NACC cohort and 305 (37%) from ADNI had at least one point on the NPI-Q and met the other entry criteria to be included in these analyses. Group sizes for each of the NPI-Q subscale analyses differed according to how many cases had at least one point on the particular subscale to be included in those analyses (see Table 1). The NACC cohort was significantly older and more % female at baseline. The NACC cohort's mean scores for MMSE and every NPI-Q scale were significantly worse than in ADNI, determined at the last visit for NACC and 24-month visit for ADNI. Mean MMSE scores for both samples were in the mildly impaired stage for AD.

Boxplots (see Figure 1) reveal significant differences for median scores between cohorts for MMSE (ADNI = 24, NACC = 21; $p<0.0001$), NPI-Q-10 (ADNI = 3, NACC = 4; $p<0.008$) and NPI-Q-4-Frontal (ADNI = 2, NACC = 2; $p<0.041$), but not for NPI-Q-4-Agitation/Aggression or NPI-Q-3-Mood (all analyses conducted using Wilcoxon rank-sum test for 2 samples with normal approximation). The NACC cohort had proportionally more advanced stage cases than the ADNI cohort (on MMSE and NPI-10). Median scores are in the milder range, though distributions also extend into higher severities.

Pareto plots of NPI-Q and NPI-Q subscale variances for each factor, as explained by an ordered PCA, are shown for NACC and ADNI cohorts in Figure 2. By definition the factors calculated by the PCA as shown in the Pareto plots are independent and therefore contributing uniquely. In each plot the first factor comprises proportionately more variance, with the subsequent factors each contributing additional variance. Pareto plots for ADNI and NACC cohorts are very similar.

Table 2 shows the factor (F) loadings for each item of NPI-Q-10 and NPI-Q subscales for NACC and ADNI cohorts with shadings for values meeting the cutoff of ≥0.40, the generally accepted cutoff for meaningfulness (20). The item composition of individual factors on the NPI-10 were largely the same for F1 between the two cohorts, but had a more random pattern across the other factors. On the NPI-Q-4 Agitation/Aggression subscale, all four items loaded onto F1 in each cohort (except irritability/lability trended in ADNI) and therefore measure something in common. Irritability/lability and aberrant motor behavior play opposite roles in F2, and item loadings in F3 and F4 are the same but with these factors flipped between the cohorts. Nearly all items loaded onto F1 in the NPI-Q-3-Mood subscale in both cohorts, where F2 loaded depression/dysphoria and F3 loaded anxiety and irritability. Loadings for the NPI-Q-4-Frontal subscale revealed similarities and differences between the cohorts. Apathy did not load onto F1 in either cohort and F2, F3 and F4 loadings showed item pattern similarities when direction of loadings were considered including items that did not meet the 0.40 cutoff. F2 showed opposite direction for item loadings (apathy and irritability) between the cohorts.

Figure 3 shows Gabriel biplots for the first three factors' item loadings for each NPI-Q scale/subscale for the ADNI and NACC cohorts (separately by column). Gabriel biplots display 2-dimensional (x and y axes) projections plotted between pairs of factors' item loadings, where the direction and length of a vector represents a spatial representation of the *relationship* for items between two given factors. This produces a visual display - not of a single factor loading as listed in Table 2, but rather of the relationship of item factor loadings between every two factors. For example, for the ADNI cohort's NPI-Q-3 Mood subscale in Table 2, item D loaded onto F1 with a value of 0.30 and onto F2 at 0.95, so that in Figure 3's topmost graph for the NPI-Q-3 Mood subscale item D is positioned based on the x axis (F1) loading value of 0.30 and on the y-axis (F2) loading value of 0.95 which produces a vector ending at the intersection for those two values. The spatial patterns reveal information about item relationships across factors which allow comparisons between ADNI and NACC cohorts.

The biplots first compare item projection relationships for F1 with either F2 or F3, then between F2 and F3, so that each factor pair permutation is graphed. When F1 item values were graphed against values for F2 or F3, plots for all scales/subscales showed that all items projected in a positive direction in the factor space with only a small degree of separation between items (see the top four biplots for each scale/subscale in Figure 3 where all item values project toward the right on the x-axis). This indicates they are measuring something in common for that clinical phenotype domain (coherence). However, in plots for F2 item values with those for F3 (see bottom two graphs for each scale/subscale in Figure 3), items are more separated from each other and project in positive or negative directions along the

axis of the factor, indicating that items contribute differentially (ie, uniquely)to the phenotype domain being measured by the scale (independence). Therefore biplots reveal both coherence and independence of items for each subscale, though with F1 revealing coherence on biplots and also accounting for the most variance of any factor on Pareto plots.

Notably, the projection pattern for each scale graph is very similar between ADNI and NACC cohorts and their comparability supports validity of the subscales, even though there was some variation as to which item comprised a given projection in some of the biplots.

Conclusion

Our predefined purpose was to characterize how the NPI-Q-10 and three prespecified NPI-Q subscales performed in unrotated PCA in over 5000 MCI and AD dementia patients from two public databases to evaluate their psychometric attributes for research and clinical applications. Such NPI subscales have not been previously analyzed for their potential use as independent subscales in clinical and research applications, given the growing recognition of the importance of understanding clusters of NPS. Each individual patient's neurodegenerative course is unique as it affects certain brain regions and circuitry so that identification of meaningful NPI-Q subscales could allow researchers and clinicians more opportunities to individualize care. Further, because the FDA (21,22) requires measurement of definable syndromes in AD, such subscales could be useful in drug development. Comparing data from two different cohorts – ADNI and NACC – and finding a high degree of comparability in PCA results contributes to the validity of these subscales, in addition to measurement of a common phenotype domain where each item contributed uniquely to the subscale in our PCA analyses.

We determined subscale item composition based on the AD literature especially large exploratory PCA analyses (19,20) and on knowledge of symptoms and their functional neuroanatomy. We believe this optimized subscale face validity. Some of the items selected for each of our subscales overlapped with another subscale because our intent was not to simply divide up the NPI-Q scale into subsets but rather to be able to measure meaningful clinical domains, reflecting that clinical NPS syndromes have symptom clusters that are nonexclusive. As expected, F1 explained the majority of the variance for all scales, though all items made some contribution to the variance (see Pareto plots).

Though less often used in clinical literature, Gabriel biplots are a unique method to express spatial relationships between factors' item projections in PCA, information not apparent from simply listing factor loading values. These biplots have been adopted in other scientific fields as a powerful tool to visually interpret multivariate data and discern item relationships between factors. Gabriel biplots illustrated that F1, when plotted with either F2 or F3, revealed coherence which supports that the items together were measuring a symptom domain in common. Taken together, F1's high variance proportion, face validity for item selection, and coherence all suggest the subscale composition was valid in assessing a particular phenotype. While F2 and F3 accounted for less of the total variance than F1, each item was necessary to the subscale composition. Gabriel biplots comparing F2 with F3 item values found considerable spatial separation across the items which suggest that each item also made unique contributions to the underlying phenomenological domain being measured by the subscale. Therefore, the items selected for each subscale were necessary and not redundant while also measuring something together. This is a measure of validity. Interestingly, all scales showed the same item projection spatial patterns between the groups (see Figure 3).

We chose not to report any analyses of a psychosis subscale because most exploratory analyses find delusions and hallucinations loading together and two NPI-Q items was insufficient for our PCA analysis methods. However, psychosis is an important component of any assessment battery for patients, particularly those in the severe range.

We wanted a subscale to assess symptoms that represent prefrontal cortex dysfunction. Such a subscale might be useful in comparisons of AD and FTD (frontotemporal dementia) patients. We included apathy in the frontal subscale to capture mesial frontal pathology. Apathy often loads onto a separate factor and has been shown to be different from depression in AD (23-25), though it can also load with other items depending on the cohort. Apathy is the most persistent NPS throughout all stages of MCI and AD but we assigned it only to our NPI-Q-4-Frontal subscale.

The NPI-Q-4-Agitation/Aggression subscale reflects the literature where these symptoms often cluster together on a factor (19,20,26). Occasionally, and depending on the particular study sample, these symptoms might load with depression or psychosis; however given the burden associated with these disruptive behaviors we believe they need to be measured independent of mood symptoms (4,27,28). The growing recognition of the importance of symptoms of agitation and aggression in AD was recognized at a recent consensus conference, where experts in the field of geriatric mental health stressed the need for specific FDA-approved treatments for symptoms of severe and persistent or recurrent agitation and aggression in patients with dementia, even those who don't experience concomitant psychosis (29).

Although we utilized data from two cohorts (ADNI and NACC), we ascertained a high level of similarity across these two cohorts on each scale that supported our validation. There were overall striking similarities between the two cohorts despite more severe cases of cognitive impairment and NPS in NACC than ADNI. Where subtle differences in factor item loadings or directional projections existed, it may be attributable to differences in the clinical characteristics of these cohorts such as ADNI excluding cases with comorbid Major Depressive Disorder. The NACC cohort is more clinically representative in its case inclusion and also much larger. Our work is also consistent with findings obtained by the European Alzheimer Disease Consortium in another large sample with small differences in sample selection (18,19). Subtle differences might also relate to study design where ADNI is a more demanding study with visits every 6 months and many required tests including neuroimaging, thereby being more burdensome for caregivers to participate in. NACC is less standardized across centers and more observational with annual visits. The frontal subscale revealed more differences than other scales between cohorts in factor loadings which might reflect that frontal degeneration increases with advancing disease. Nonetheless, the biplots revealed much spatial similarity between cohorts.

Whether NPS syndromes are completely separable in each patient is improbable. It is also possible that symptoms are less distinct at different phases of illness. However, our intent was to establish valid subscales that capture a particular constellation of NPS so that clinical assessments can track NPS syndromes over time and research about the evolution of NPS in AD dementia and MCI can be encouraged using these tools. Response to interventions could also be measured. NPS syndromes are considered a viable and meaningful target for pharmaceutical drug development though more research is needed to identify clear syndromes of problematic behaviors (21,22). Currently, proposed diagnostic criteria exist for sleep, depression and psychosis (13,30-32). Application of subscales might be useful in future research on NPS diagnostic criteria.

Limitations of our study include being a retrospective analysis of prospectively collected databases. A replication of our findings in a separate database would increase confidence that these are valid, though the striking similarities between ADNI and NACC PCA results within this report suggests these subscales are likely to represent typical NPS domains. Unfortunately, there is no gold standard for NPS syndromes against which to compare these NPI-Q subscales nor were the existing proposed diagnostic criteria available in ADNI or NACC. Instead, we relied heavily on previously published exploratory factor analyses to select subscale items. Additionally, more psychometric analyses are recommended for these subscales particularly evaluating their use in diverse populations (eg, one with more patients with more severe NPS), especially focusing on divergent and convergent validity and reliability. We relied heavily on prior demonstration of the NPI's psychometric credentials for the total scale and each individual domain (which can stand alone) and therefore focused our validation efforts on the item composition of these subscales themselves where PCA appeared to be a pertinent approach.

Our results are of interest because the NPI-Q is used widely in research, including as an outcome measure in clinical trials. We conclude that these three NPI-Q subscales when used in the context of administering the full NPI scale may be useful tools to assess meaningful NPS domains. This may be advantageous to provide more focused measurements of NPS clusters which can guide treatment or research decisions. Further validation of these subscales could strengthen our findings.

Acknowledgments

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amorfix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

The Principal Investigator of the ADNI initiative is Michael W. Weiner, MD, VA Medical Center and University of California – San Francisco. ADNI is the result of efforts of many co- investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research, approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years and 200 people with early AD to be followed for 2 years.” For up-to-date information, see www.adni-info.org.

We acknowledge the investigators from the National Alzheimer's Coordinating Center (NAAC) who contributed their data to the NACC database for analyses [U01 AG016976].

There was no grant funding for this report. Data were analyzed at Lilly. We appreciate the statistical consultation of Anthony J. Zagar.

References

1. Lyketsos C, Carrillo MC, Ryan JM, et al. Perspectives: Neuropsychiatric symptoms in Alzheimer's Disease. Alzheimers Dement. 2011; 7(5):532–539. [PubMed: 21889116]

2. Geda YE, Roberts RO, Knopman DS, et al. Baseline neuropsychiatric symptoms and the risk of incident Mild Cognitive Impairment: the Mayo Clinic Study of Aging. Presentation at the annual meeting of American Academy of Neurology. 2011
3. Geda YE, Roberts RO, Knopman DS, et al. Prevalence of neuropsychiatric symptoms in Mild Cognitive Impairment and normal cognitive aging. Arch Gen Psychiatry. 2008; 65(10):1193–1198. [PubMed: 18838636]
4. Gaugler JE, Wall MM, Kane RL, et al. The effects of incident and persistent behavioral problems on change in caregiver burden and nursing home admission of persons with dementia. Med Care. 2010; 48:875–883. [PubMed: 20733529]
5. Gaugler JE, Wall MM, Kane RL, et al. Does caregiver burden mediate the effects of behavioral disturbances on nursing home admission? Am J Geriatr Psychiatry. 2011; 19(6):497–506. [PubMed: 21606895]
6. Karttunen K, Karppi P, Hiltunen A, et al. Neuropsychiatric symptoms and Quality of Life in patients with very mild and mild Alzheimers disease. Int J Geriatr Psychiatry. 2011; 26:473–482. [PubMed: 21445998]
7. Rocca P, Leotta D, Liffredo C, et al. Neuropsychiatric Symptoms underlying caregiver stress and insight in Alzheimer's disease. Dement Geriatr Cogn Disord. 2010; 30:57–63. [PubMed: 20689284]
8. Cummings JL, Mega M, Gray K, et al. The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. Neurology. 1994; 44:2308–2314. [PubMed: 7991117]
9. Cummings JL. The Neuropsychiatric Inventory: Assessing psychopathology in dementia patients. Neurology. 1997; 48(5 Suppl 6):S10–S16. [PubMed: 9153155]
10. Wood S, Cummings JL, Hsu MA, et al. The use of the neuropsychiatric inventory in nursing home residents: characterization and measurement. Am J Geriatr Psychiatry. 2000; 8:75–83. [PubMed: 10648298]
11. Kaufer DI, Cummings JL, Ketchel P, et al. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. J Neuropsychiatry Clin Neurosci. 2000; 12:233–239. [PubMed: 11001602]
12. Lyketsos C. Neuropsychiatric Syndromes in Dementia and MCI: Where are we heading? Alzheimers and Dement. 2011:S278.
13. Jeste D, Finkel. Psychosis of Alzheimer's Disease and related dementias. Am J Geriatr Psychiatry. 2000; 8:29–34. [PubMed: 10648292]
14. Cohen-Mansfield J, Marx MS, Rosenthal AS. A description of agitation in a nursing home. J Gerontol. 1989; 44:M77–84. [PubMed: 2715584]
15. Gabriel, KR. Biplot display of multivariate matrices for inspection of data and diagnosis. In: Barnett, V., editor. Interpreting Multivariate Data. John Wiley & Sons; London: 1981.
16. Jacoby, WG. Statistical Graphics for Visualizing Multivariate Data. Thousand Oaks: Sage Publications; 1998.
17. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12(3):189–98. [PubMed: 1202204]
18. Aalten P, Verhey F, Boziki M, et al. Neuropsychiatric syndromes in dementia: Results from the European Alzheimer Disease Consortium: Part I. Dementia Geriatr Cogn Dis. 2007; 24:457–463.
19. Aalten P, Verhey FRJ, Boziki M, et al. Consistency of Neuropsychiatric Syndromes across Dementias: Results from the European Alzheimer Disease Consortium. Dement Geriatr Cogn Disord. 2008; 25:1–8. [PubMed: 18025783]
20. Hair, J.; Anderson, R.; Tatham, R.; Black, W. Multivariate Data Analysis. 5th. Prentice-Hall; Upper Saddle River, NJ: 1988.
21. Laughren T. A regulatory perspective on psychiatric syndromes in Alzheimer's Disease. Am J Geriatr Psychiatry. 2001; 9:340–345. [PubMed: 11739061]
22. Laughren T. FDA perspective in the DSM-V approach to classification of “cognitive” disorders. J Neuropsychiatry Clin Neurosci. 2011; 23:126–131. [PubMed: 21677238]
23. Levy ML, Cummings JL, Fairbanks LA, et al. Apathy is not Depression. J Neuropsychiatry Clin Neurosci. 1998; 10:314–319. [PubMed: 9706539]

24. McPherson S, Fairbanks L, Tiken S, et al. Apathy and executive function in Alzheimer's disease. J Int Neuropsychol Soc. 2002; 8:373–381. [PubMed: 11939696]

25. Onyike CU, Sheppard JE, Tschanz JT, et al. Epidemiology of apathy in older adults: the Cache County study. Am J Geriatr Psychiatry. 2007; 15:365–375. [PubMed: 17463187]

26. Aalten P, deVugt ME, Lousberg R, et al. Behavioral problems in dementia: A factor analysis of the Neuropsychiatric Inventory. Dement Geriatr Cogn Disord. 2003; 15:99–105. [PubMed: 12566599]

27. Davis JD, Tremont G. Impact of frontal systems behavioral functioning in dementia on caregiver burden. J Neuropsychiatry Clin Neurosci. 2007; 19:43–49. [PubMed: 17308226]

28. Rymer S, Salloway S, Norton L, et al. Impaired awareness, behavior disturbance, and caregiver burden in Alzheimer disease. Alzheimer Dis and Assoc Disord. 2002; 16:248–253.

29. Salzman C, Jeste DV, Meyer RE, et al. Elderly patients with dementia-related symptoms of severe agitation and aggression: Consensus statement on treatment options, clinical trials methodology, and policy. J Clin Psychiatry. 2008; 69:889–898. [PubMed: 18494535]

30. Yesavage JA, Friedman L, Ancoli-Isreal S, et al. Development of diagnostic criteria for defining sleep disturbance in Alzheimer's disease. J Geriatr Psychiatry Neurol. 2003; 16:131–139. [PubMed: 12967054]

31. Olin JT, Schneider LS, Kats IR, et al. Provisional diagnostic criteria for depression in Alzheimer's Disease. Am J Geriatr Psychiatry. 2002; 10:125–128. [PubMed: 11925273]

32. Starkstein SE, Jorge R, Mizrahi R, Robinson RG. The construct of minor and major depression in Alzheimer's Disease. Am J Psychiatry. 2005; 162:2086–2093. [PubMed: 16263848]

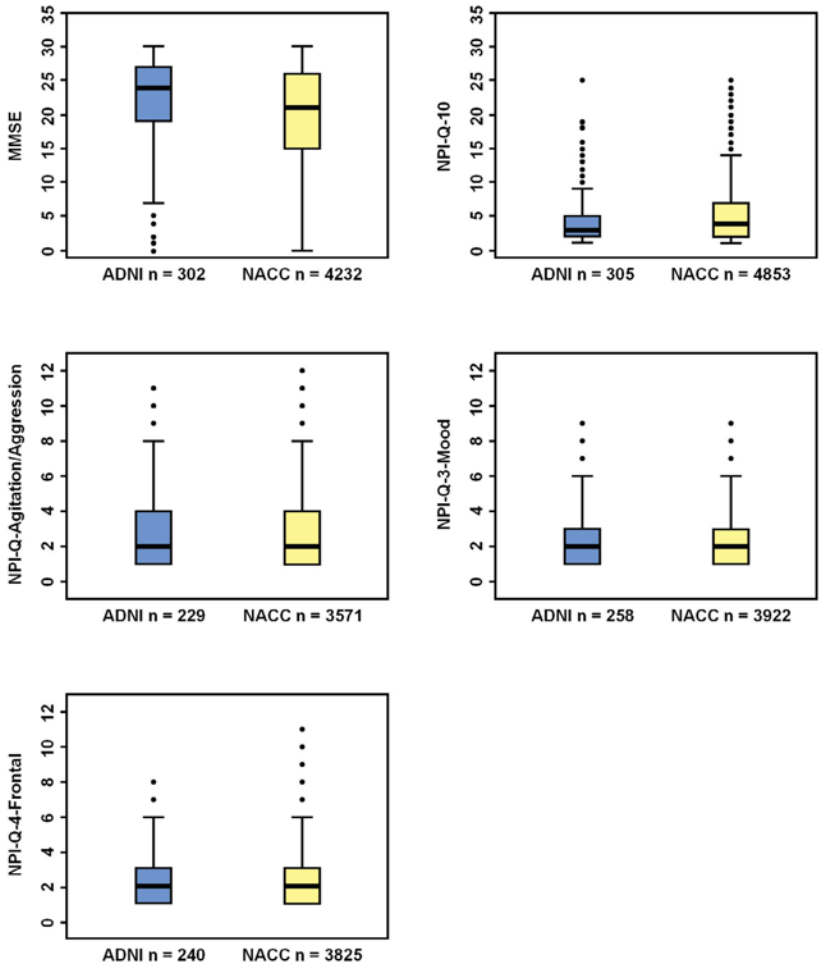


Figure 1.

Boxplots are shown for groups for MMSE, NPI-Q-10 and NPI-Q subscales at the last visit in NACC and 24-month visit in ADNI cohorts. Group sizes varied. The thick black lines show the median; the box is draw between the quartiles (25%-75%); the whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range (IQR) from the box, which correspond to the ±2.698 multiple standard deviation.

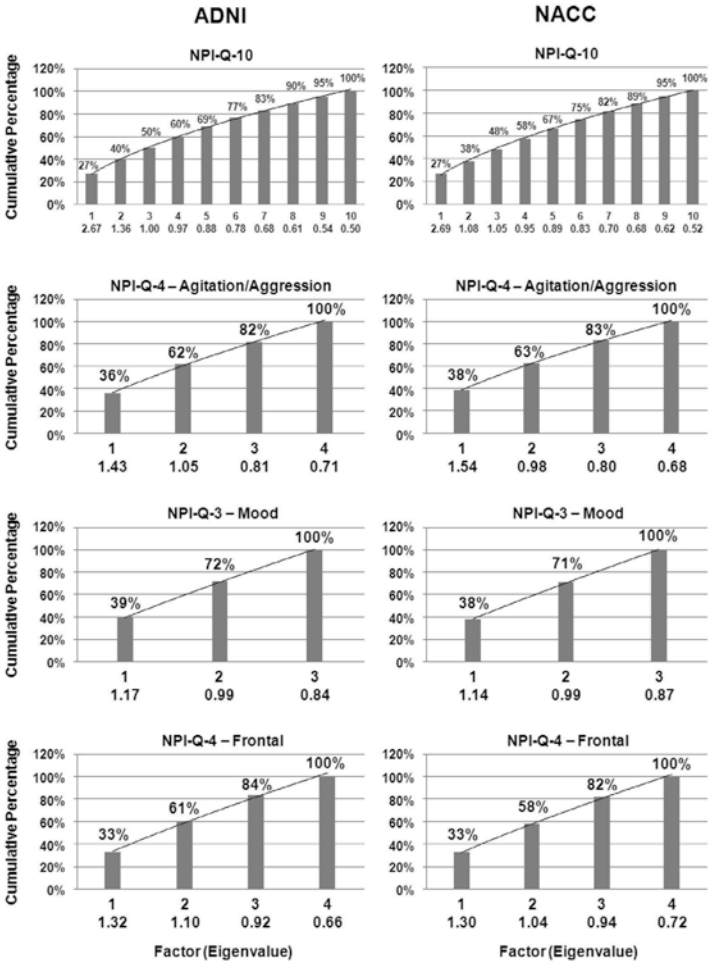
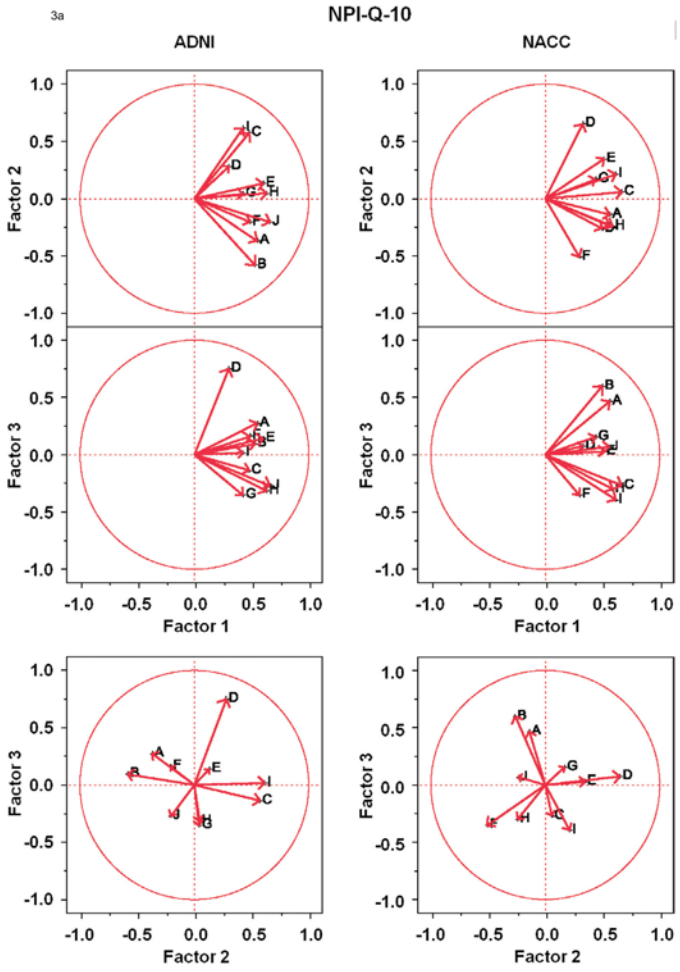
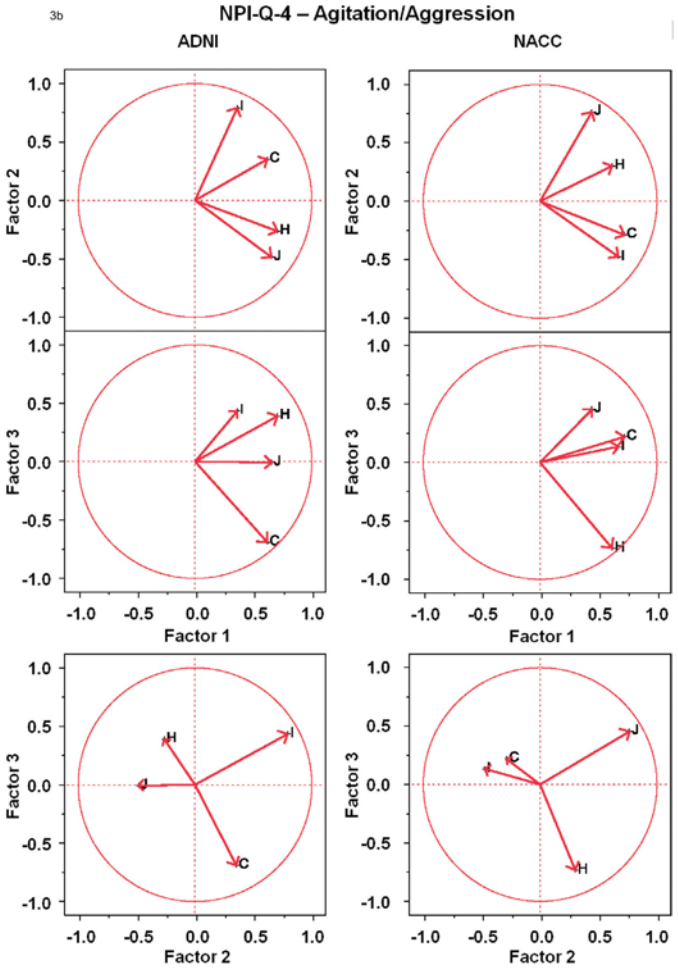
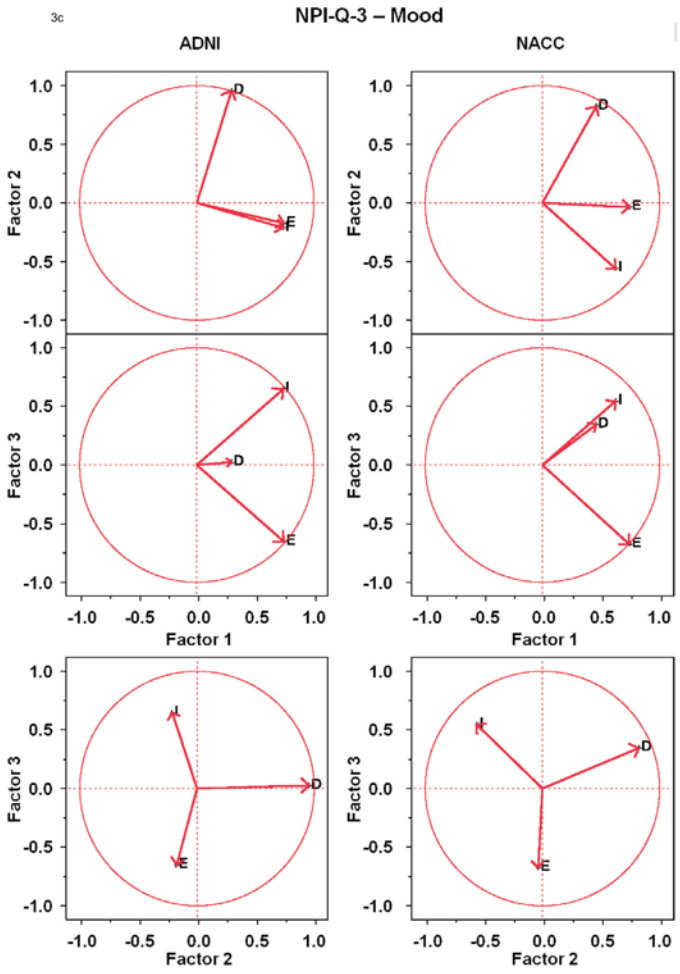


Figure 2.

Pareto plots of the percentage of variance contributed by each factor based on principal component analyses for the ADNI and NACC cohorts for the NPI-Q-10 and NPI-Q subscales.







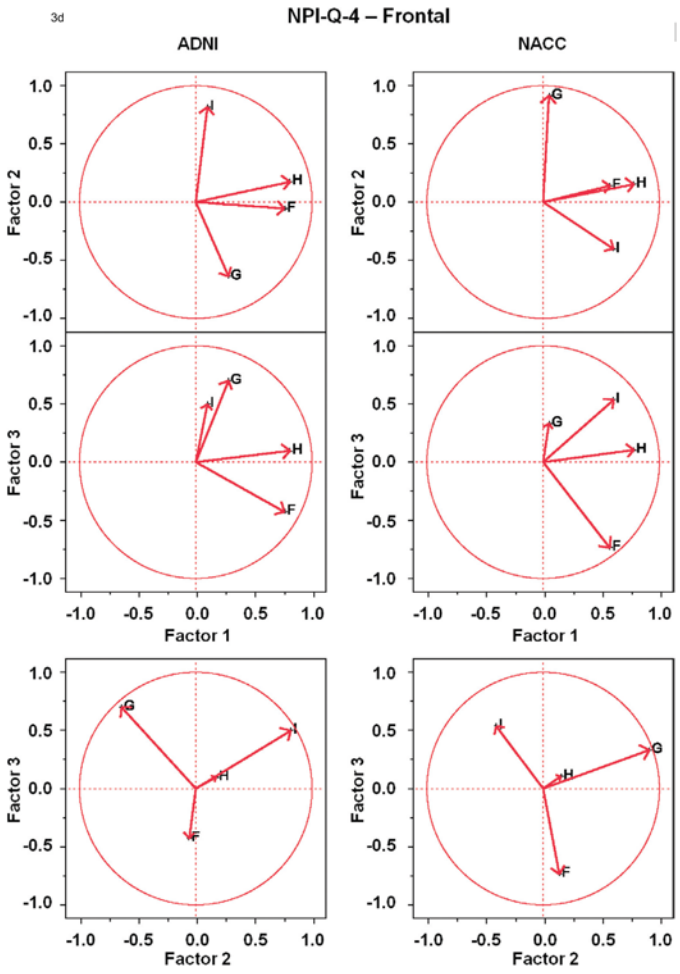


Figure 3.

Gabriel biplots are visual representations of the relationships between item factor loadings when plotted between pairs of factors because directionality offers additional information about scales and subscales. Biplots are shown for each of the NPI-Q-10 and NPI-Q subscales for the ADNI (left) and NACC (right) cohorts. Biplots are 2-dimensional graphic representations of items' factor loadings and display all of the combinations between pairs for the first three factors from PCA analyses. The top two rows of all figures plot item loading values for Factor 1 along the x-axis against those for either Factor 2 or 3 along the y axis. In the third row of all graphs, Factor 2 item values are plotted along the x axis and those for Factor 3 along the y axis. In all graphs of Factor 1's items with either Factor 2 or 3, all values point in a positive direction (ie, pointing within the right half of the circles). However, when Factors 2 and 3 are plotted with each other, the items point in both positive and negative directions. (Abbreviations for NPI-Q items are coded using letters listed in Table 2.)

Table 1

Pooled data for aMCI, MCI converter and AD dementia cases at baseline and the last visit for NACC and ADNI cohorts for MMSE and NPI-Q data. All data are expressed as mean±SD (range) unless otherwise specified. The NPI-Q groups are of different sizes; as inclusion required ≥ 1 point on at least one item in the scale or subscale.

	NPI-Q-10 Group	NPI-Q-4-Agitation/Aggression Group	NPI-Q-3-Mood Group	NPI-Q-4-Frontal Group
NACC				
N	4853	3571	3922	3825
Age	75.7±9.0* (51-110)	75.5±9.0* (51-110)	75.3±9.1* (51-110)	75.7±9.0* (51-110)
Sex (% female)	53%¥	51%¥	53%¥	50%¥
MMSE	19.5±7.7+ (0-30)	18.9±7.9+ (0-30)	19.8±7.7+ (0-30)	19.0±7.8+ (0-30)
NPI-Q	5.0±4.0® (1-25)	3.0±2.2® (1-12)	2.5±1.6 (1-9)	2.6±1.7® (1-11)
ADNI				
N	305	229	258	240
Age	74.6±7.0 (55-91)	74.1±7.1 (55-91)	74.3±7.1 (55-91)	74.5±7.1 (55-91)
Sex (% female)	40%	38%	39%	37%
MMSE	22.4±5.8 (0-30)	22.4±6.0 (0-30)	22.6±5.6 (1-30)	22.1±6.0 (0-30)
NPI-Q	4.3±3.5 (1-25)	2.7±1.9 (1-11)	2.3±1.5 (1-9)	2.3±1.6 (1-8)

*p<0.03 using Welch's ANOVA for unequal variances between NACC and ADNI (df =(1,369.98), (1,277.03), (1,316.15) and (1,289.98) for NPI-Q-10, NPI-Q-Agitation/Aggression, NPI-Q-3-Mood, and NPI-Q-4-Frontal).

¥p<0.0001 using Pearson's Chi-squared test between NACC and ADNI (df = 1 for NPI-Q-10, NPI-Q-Agitation/Aggression, NPI-Q-3-Mood, and NPI-Q-4-Frontal)

+p<0.0001 using Welch's ANOVA for unequal variances between NACC and ADNI (df = (1,380.05),(1,286.29), (1,328.33) and (1,295.48) for NPI-Q-10, NPI-Q-Agitation/Aggression, NPI-Q-3-Mood, and NPI-Q-4-Frontal)

®p≤0.05 using ANOVA (df = (1,5156),(1,3798), (1,4178) and (1,4063) for NPI-Q-10, NPI-Q-Agitation/Aggression, NPI-Q-3-Mood, and NPI-Q-4-Frontal)

Table 2

NPI-Q-10 and NPI-Q subscale items' factor loadings for both NACC and ADNI cohorts. Factors are abbreviated as "F". Letter codes for items match those used in Figure 3 Gabriel biplots. Values ≥0.40 are shaded.

Item		F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	
NPI-Q-10												
						ADNI cohort						
A	Delusions		0.55	-0.37	0.27	0.18	-0.35	0.19	0.40	-0.18	-0.24	-0.23
B	Hallucinations		0.53	-0.58	0.09	0.26	-0.05	-0.22	-0.05	0.12	0.04	0.49
C	Agitation/Aggression		0.48	0.57	-0.14	0.11	0.14	-0.36	0.10	-0.41	-0.26	0.15
D	Depression/Dysphoria		0.30	0.28	0.75	0.22	0.36	-0.13	-0.08	0.23	0.01	-0.14
E	Anxiety		0.60	0.13	0.14	-0.24	-0.25	0.34	-0.57	-0.10	-0.16	0.06
F	Elation/Euphoria		0.48	-0.20	0.16	-0.61	0.37	0.10	0.19	-0.24	0.28	0.06
G	Apathy/Indifference		0.42	0.04	-0.35	0.47	0.45	0.52	0.02	0.01	0.04	0.01
H	Disinhibition		0.63	0.05	-0.31	-0.36	0.09	-0.10	0.14	0.50	-0.29	-0.07
I	Irritability/Lability		0.42	0.61	0.02	0.05	-0.43	0.12	0.24	0.16	0.38	0.13
J	Aberrant Motor Behavior		0.66	-0.20	-0.27	0.15	-0.08	-0.36	-0.25	-0.07	0.29	-0.37
						NACC cohort						
A	Delusions		0.56	-0.14	0.46	-.019	0.25	0.09	-0.21	0.10	-0.53	0.04
B	Hallucinations		0.50	-0.26	0.60	-0.08	0.18	0.07	0.13	0.01	0.52	0.05
C	Agitation/Aggression		0.67	0.06	-0.27	-0.38	0.00	0.10	0.23	0.12	0.00	-0.51
D	Depression/Dysphoria		0.33	0.65	0.07	0.38	0.37	0.25	0.14	-0.33	-0.03	-0.03
E	Anxiety		0.51	0.35	0.03	0.26	0.03	-0.57	-0.29	0.33	0.11	-0.08
F	Elation/Euphoria		0.30	-0.51	-0.35	0.56	0.31	0.16	0.14	0.27	-0.04	0.03
G	Apathy/Indifference		0.44	0.17	0.15	0.26	-0.68	0.42	-0.01	0.22	-0.01	0.06
H	Disinhibition		0.59	-0.24	-0.30	0.00	-0.06	0.13	-0.55	-0.41	0.14	-0.03
I	Irritability/Lability		0.62	0.21	-0.39	-0.38	0.08	0.00	0.14	0.11	0.05	0.49
J	Aberrant Motor Behavior		0.56	-0.24	0.07	0.17	-0.31	-0.44	0.36	-0.38	-0.18	0.04
NPI-Q-4-Agitation/Aggression												
						ADNI cohort						
C	Agitation/Aggression		0.62	0.35	-0.69	-0.15						

Item		F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
NPI-Q-10											
H	Disinhibition		0.70	-0.26	0.39	-0.53					
I	Irritability/Lability		0.36	0.79	0.43	0.25					
J	Aberrant Motor Behavior		0.65	-0.48	-0.01	0.58					
	NACC cohort										
C	Agitation/Aggression		0.72	-0.29	0.22	-0.59					
H	Disinhibition		0.61	0.30	-0.73	0.00					
I	Irritability/Lability		0.67	-0.47	0.14	0.56					
J	Aberrant Motor Behavior		0.44	0.76	0.45	0.12					
NPI-Q-3-Mood											
	ADNI cohort										
D	Depression/Dysphoria		0.30	0.95	0.02						
E	Anxiety		0.74	-0.17	-0.65						
I	Irritability/Lability		0.73	-0.21	0.64						
	NACC cohort										
D	Depression/Dysphoria		0.46	0.82	0.34						
E	Anxiety		0.74	-0.03	-0.67						
I	Irritability/Lability		0.62	-0.56	0.54						
NPI-Q-4-Frontal											
	ADNI cohort										
F	Elation/Euphoria		0.76	-0.06	-0.43	0.49					
G	Apathy/Indifference		0.28	-0.64	0.70	0.17					
H	Disinhibition		0.81	0.17	0.10	-0.55					
I	Irritability/Lability		0.10	0.81	0.50	0.29					
	NACC cohort										
F	Elation/Euphoria		0.57	0.14	-0.73	0.34					
G	Apathy/Indifference		0.05	0.91	0.33	0.24					
H	Disinhibition		0.78	0.15	0.10	-0.60					
I	Irritability/Lability		0.60	-0.40	0.53	0.43					