Biweekly fluctuations of neuropsychiatric symptoms according to the Neuropsychiatric Inventory: Erratic symptoms or scores?

Willem S. Eikelboom1 | Amy den Teuling1 | Daphne E. Pol1 | Michiel Coesmans2 | Sanne Franzen1 | Lize C. Jiskoot1 | Judy van Hemmen1 | Ellen H. Singleton3 | Rik Ossenkoppele3,4 | Frank Jan de Jong1 | Esther van den Berg1 | Janne M. Papma1

1Department of Neurology and Alzheimer Center Erasmus MC, Erasmus MC University Medical Center, Rotterdam, The Netherlands
2Department of Psychiatry, Erasmus MC University Medical Center, Rotterdam, The Netherlands
3Department of Neurology, Alzheimer Center Amsterdam, Amsterdam University Medical Centers, Amsterdam, The Netherlands
4Clinical Memory Research Unit, Lund University, Malmö, Sweden

Correspondence
Willem S. Eikelboom, Department of Neurology and Alzheimer Center Erasmus MC, Erasmus MC University Medical Center, PO Box 2040, 3000 CA, Rotterdam, The Netherlands.
Email: w.eikelboom@erasmusmc.nl

Funding information
Alzheimer Nederland and Memorabel ZonMw, Grant/Award Number: 733050823

Abstract
Objectives: This study investigates the stability of neuropsychiatric symptoms (NPS) assessed biweekly using the Neuropsychiatric Inventory (NPI) in a memory clinic population during a 6 week period.

Methods: Twenty-three spousal caregivers (mean [SD] age = 69.7 [8.8], 82.6% female) of 23 patients (43.5% had dementia) completed all assessments. The NPI was assessed four times during 6 weeks. We examined whether NPI domains were present during all four assessments, studied within-person variation for each NPI domain, and calculated Spearman's correlations between subsequent time points. Furthermore, we associated repeated NPI assessments with repeated measures of caregiver burden to examine the clinical impact of changes in NPI scores over time.

Results: The course of NPS was highly irregular according to the NPI, with only 35.8% of the NPI domains that were present at baseline persisted during all 6 weeks. We observed large within-person variation in the presence of individual NPI domains (61.3%, range 37.5%–83.9%) and inconsistent correlations between NPI assessments (e.g., range r = 0.20–0.57 for agitation, range r = 0.29–0.59 for anxiety). Higher NPI total scores were related to higher caregiver burden (r = 0.60, p < 0.001), but changes in NPI total scores were unrelated to changes in caregiver burden (r = 0.16, p = 0.20).

Conclusions: We observed strong fluctuations in NPI scores within very short time windows raising the question whether this represents erratic symptoms and/or scores. Further studies are needed to investigate the origins of these fluctuations.

KEYWORDS
Alzheimer’s disease, behavioral and psychological symptoms of dementia (BPSD), behavioral symptoms, dementia, Neuropsychiatric Inventory, neuropsychiatric symptoms
INTRODUCTION

Neuropsychiatric symptoms (NPS) such as depression, apathy, agitation, and sleep disturbances are frequently observed in individuals who visit the memory clinic. These symptoms have a major impact on the lives of patients and their caregivers and are associated with increased caregiver burden. The Neuropsychiatric Inventory (NPI) is considered the gold standard to assess NPS in neurocognitive disorders.

Previous studies that examined the course of NPS using the NPI have shown large within-person variability in the progression of NPI scores when administering the NPI every 6–12 months. It remains unclear whether there is also such within-person heterogeneity in longitudinal NPI scores when measured during shorter time intervals, for example, within weeks instead of months. Although several studies have administered the NPI twice within a timeframe of 2–3 weeks to establish the test-retest reliability of the NPI, knowledge on short-term trajectories of NPS according to repeated NPI assessments is lacking.

Here, we describe the stability of NPI scores over a period of 6 weeks in a memory clinic population. During this 6 week period, the NPI was administered biweekly in order to compare our findings with previous test-retest studies that have assessed the NPI within a similar timeframe. Furthermore, we compared the trajectories of NPI scores with repeated measures of caregiver burden. NPS is a well-known contributor to caregiver burden. Therefore, we included a measure of caregiver burden to examine the clinical impact of short-term changes in NPI scores. Based on previous test-retest studies, we hypothesized stable NPI scores over time for apathy and psychotic symptoms, while we expected less stable NPI scores for affective symptoms, agitation-related behaviors, and sleep disturbances.

METHODS

2.1 Study design and participants

We invited all caregivers of patients who visited the memory clinic of the Erasmus MC in Rotterdam, the Netherlands, between June 2020 and July 2020, and between November 2020 and January 2021, to participate in this study. We included participants regardless of clinical diagnosis and presence/severity of NPS at baseline, with the only requirement that caregivers had to live with the patient. All patients underwent a standard diagnostic workup including medical history taking, neurological examination, neuropsychological assessment, and brain MRI. Clinical diagnoses were established using conventional diagnostic criteria during a multi-disciplinary meeting.

2.2 Measures

The Dutch NPI and Dutch Caregiver Strain Index (CSI)-Expanded were administered in person to caregivers during the initial the memory clinic visit. During the 6 weeks that followed, the NPI and CSI-Expanded assessments were repeated every 2 weeks by telephone. Caregivers evaluated the presence, frequency (0–4), severity (0–3), and distress (0–5) of NPS in the previous 2 weeks. NPI domain scores were calculated by multiplying the frequency and severity scores (0–12). The presence of specific NPI domains was defined as an NPI domain score of ≥1. We summed the 12 NPI domain scores to obtain the NPI total score (0–144). The Dutch CSI-Expanded was used to assess caregiver burden. This instrument covers aspects of caregiver strain (13 items) and aspects of caregiving that may decrease burden (5 items), resulting in a total score ranging between −5 and 13.

2.3 Data analysis

We examined the prevalence of specific NPI domains at baseline and its persistence. NPI domains were persistent if they were present on all four assessments. For each NPI domain, we described the between-person variation (i.e., how many individuals had an NPI domain score of ≥1 at least once) and the within-person variation (i.e., total number of assessments in which NPS were present in individuals who had an NPI domain score of ≥1 at least once, with both 0% and 100% indicating no variation). For each NPI domain, we conducted Spearman’s correlations to examine the relationship between NPI domain scores on subsequent time-points (baseline-t1, t1-t2, t2-t3). Individual trajectories of NPI domain scores over time were plotted for descriptive purposes, but not analyzed at group-level.
We correlated NPI total scores with CSI-Expanded total scores across all time-points. Next, we calculated delta scores for NPI total scores and CSI-Expanded total scores for each time-point and associated these delta scores using Spearman’s correlations.

To examine the effects of cognitive status, we conducted exploratory analysis in which prevalence, persistence, between-person variation, within-person variation, and Spearman’s correlations between NPI domain scores were performed for patients with dementia and patients with cognitive impairment no dementia (CIND) separately.

### 2.4 Ethics

This study received ethical approval from the Medical Ethics Committee of the Erasmus University Medical Center (MEC-2020-0341). All participants gave informed consent.

### 3 RESULTS

#### 3.1 Study participants

A total of 26 caregivers agreed to participate in this study. There were three drop-outs during the study due to perceived burden (n = 1), acute health problems of the caregiver (n = 1), and loss of contact (n = 1). All analyses were conducted in the 23 caregivers who completed all assessments (Table 1). Caregivers had a mean age of 69.7 (SD = 8.8), 82.6% were female, and all were spouses. The patients had a mean age of 72.8 (SD = 8.2) and 21.7% were women. Most patients were diagnosed with dementia (n = 10, 43.5%), eight individuals had CIND (34.8%), and four patients had no evidence of cognitive impairment (17.4%). A clinical diagnosis could not be determined in one individual (4.3%). Two patients (8.7%) were on a stable dose of psychotropic medications, while Escitalopram was prescribed during study period in only one patient (4.3%). A cognitive enhancer was prescribed during the 6 week period for two patients (8.7%).

#### 3.2 Prevalence and course of NPS according to the NPI

At baseline, all caregivers indicated the presence of at least one NPS [mean number of NPI domains was 3.0 (range 1–6)], with a mean NPI total score of 12.3 (SD = 9.5). Irritability (56.5%), sleep disturbances (47.8%), and depression (42.3%) were most common at baseline (Table 2). Across all NPI domains, the within-person variation was 61.3% (range 37.5%–83.9%), indicating that NPI domains that were present once during the course of the study were observed at 61.3% of the four time-points (Table 2). Only 35.8% (range 0.0%–100.0%) of the NPI domains that were present at baseline persisted over all three follow-up assessments (Table 1). There were no substantial differences between patients with dementia and CIND in within-person variation across NPI domains (dementia: 63.5% (range 25.0%–94.3%), CIND: 62.8% (range 25.0%–79.2%)) and persistence of NPI domains (dementia: 37.2%, CIND: 35.8%) (see Table S2).

Figure 1 shows considerable heterogeneity in course of NPI domain scores between individuals, but especially reveals substantial fluctuations within individuals. Spearman’s correlations between NPI domain scores at two subsequent time-points (baseline-t1, t1–t2, t2–t3) varied greatly within NPI domains (see Table S1). NPI total scores correlated significantly between time-points (range \(r_s = 0.55–0.67\),

### TABLE 1 Baseline characteristics of included sample

<table>
<thead>
<tr>
<th>Caregivers (n = 23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (IQR)</td>
</tr>
<tr>
<td>Sex, N female (%)</td>
</tr>
<tr>
<td>Education, median (IQR)</td>
</tr>
<tr>
<td>Relationship with patient, N (%)</td>
</tr>
<tr>
<td>Spouse</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Patients (n = 23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (IQR)</td>
</tr>
<tr>
<td>Sex, N female (%)</td>
</tr>
<tr>
<td>Education, median (IQR)</td>
</tr>
<tr>
<td>Clinical diagnosis, N (%)</td>
</tr>
<tr>
<td>Dementia</td>
</tr>
<tr>
<td>Alzheimer’s disease dementia</td>
</tr>
<tr>
<td>Primary Progressive Aphasia</td>
</tr>
<tr>
<td>Behavioral variant frontotemporal dementia</td>
</tr>
<tr>
<td>Corticobasal syndrome</td>
</tr>
<tr>
<td>Cognitive impairment no dementia</td>
</tr>
<tr>
<td>Mild cognitive impairment</td>
</tr>
<tr>
<td>Radiation-induced cognitive decline</td>
</tr>
<tr>
<td>Cognitive impairment due to epilepsy</td>
</tr>
<tr>
<td>No cognitive impairment</td>
</tr>
<tr>
<td>Subjective cognitive decline</td>
</tr>
<tr>
<td>Major depressive episode</td>
</tr>
<tr>
<td>Could not be determined</td>
</tr>
<tr>
<td>Months since clinical diagnosis, median (IQR)*</td>
</tr>
<tr>
<td>Mini-mental state examination score, median (IQR)*</td>
</tr>
<tr>
<td>Cognitive enhancer use at baseline, N (%)</td>
</tr>
<tr>
<td>Cognitive enhancer started during study, N (%)</td>
</tr>
<tr>
<td>Psychotropic drug use at baseline, N (%)</td>
</tr>
<tr>
<td>Psychotropic drug started during study, N (%)</td>
</tr>
</tbody>
</table>

Abbreviation: IQR, interquartile range.

*Not applicable for n = 1.

†Missing data for n = 3.
later. Our findings provide further evidence for the large variability of
sample are poorly related to NPI scores administered only 2 weeks
This study shows that NPI scores at one time
4
caregiver burden (4)
Across all time
3.3
points, higher NPI total scores were related to higher
caregiver burden (\( r = 0.60, p < 0.001 \)). However, changes in NPI total
scores were unrelated to changes in caregiver burden (\( r = 0.16, p = 0.20 \)).
TABLE 2
Prevalence, persistence, and between-person and within-person variation of the presence of specific NPI domains

<table>
<thead>
<tr>
<th>NPI domain</th>
<th>Presence at baseline</th>
<th>Persistence ( ^a )</th>
<th>Between-person variation ( ^b )</th>
<th>Within-person variation ( ^c )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irritability</td>
<td>13 (56.5%)</td>
<td>7 (53.8%)</td>
<td>18 (78.3%)</td>
<td>70.8%</td>
</tr>
<tr>
<td>Sleep disturbances</td>
<td>11 (47.8%)</td>
<td>8 (72.7%)</td>
<td>14 (60.9%)</td>
<td>83.9%</td>
</tr>
<tr>
<td>Depression</td>
<td>10 (42.3%)</td>
<td>3 (30.0%)</td>
<td>16 (69.6%)</td>
<td>65.5%</td>
</tr>
<tr>
<td>Apathy</td>
<td>8 (34.8%)</td>
<td>4 (50.0%)</td>
<td>12 (52.2%)</td>
<td>60.4%</td>
</tr>
<tr>
<td>Anxiety</td>
<td>8 (34.8%)</td>
<td>1 (12.5%)</td>
<td>9 (39.1%)</td>
<td>55.6%</td>
</tr>
<tr>
<td>Eating behavior</td>
<td>6 (26.1%)</td>
<td>3 (50.0%)</td>
<td>10 (43.5%)</td>
<td>62.5%</td>
</tr>
<tr>
<td>Agitation</td>
<td>5 (21.7%)</td>
<td>1 (20.0%)</td>
<td>9 (39.1%)</td>
<td>61.1%</td>
</tr>
<tr>
<td>Aberrant motor behavior</td>
<td>5 (21.7%)</td>
<td>2 (40.0%)</td>
<td>7 (30.4%)</td>
<td>71.4%</td>
</tr>
<tr>
<td>Disinhibition</td>
<td>1 (4.3%)</td>
<td>0 (0.0%)</td>
<td>6 (26.1%)</td>
<td>50.0%</td>
</tr>
<tr>
<td>Euphoria</td>
<td>1 (4.3%)</td>
<td>1 (100.0%)</td>
<td>3 (13.0%)</td>
<td>50.0%</td>
</tr>
<tr>
<td>Hallucinations</td>
<td>1 (4.3%)</td>
<td>0 (0.0%)</td>
<td>4 (17.4%)</td>
<td>37.5%</td>
</tr>
<tr>
<td>Delusions</td>
<td>1 (4.3%)</td>
<td>0 (0.0%)</td>
<td>3 (13.0%)</td>
<td>66.7%</td>
</tr>
</tbody>
</table>

Abbreviations: NPI, Neuropsychiatric Inventory; NPS, neuropsychiatric symptoms.
\( ^a \)N (%) of individuals which showed NPS during all follow-up assessments when present at baseline.
\( ^b \)N (%) of individuals with NPS present at least at one time-point.
\( ^c \)For those with NPS at one time-point, % of assessments NPS was present.

\( p < 0.01 \), while low and inconsistent correlation coefficients were observed for specific NPI domains such as agitation (range \( r_s = 0.20–0.57 \)), irritability (range \( r_s = 0.26–0.65 \)), aberrant motor behavior (range \( r_s = 0.55–0.90 \)), and anxiety (range \( r_s = 0.29–0.59 \)). Spearman’s correlations were slightly higher in patients with dementia compared to individuals with CIND for NPI total scores (dementia: range \( r_s = 0.70–0.87, \) all \( p < 0.05, \) CIND range \( r_s = 0.56–0.83, 2/3 \ p > 0.05 \)) and several NPI domain scores (see Table S3).

We considered the presence of sleep disturbances, irritability, and aberrant motor behavior to be most stable, while the presence of hallucinations, disinhibition, and anxiety were the least stable within persons (Table 2). When present, we considered the severity/frequency of apathy, sleep disturbances, and euphoria most stable, while depression, anxiety, and hallucinations were the least stable (Figure 1 and Table S1).

3.3 Data associations between NPI scores and caregiver burden

Across all time-points, higher NPI total scores were related to higher
caregiver burden (\( r = 0.60, p < 0.001 \)). However, changes in NPI total
scores were unrelated to changes in caregiver burden (\( r = 0.16, p = 0.20 \)).

4 DISCUSSION

This study shows that NPI scores at one time-point in a memory clinic
sample are poorly related to NPI scores administered only 2 weeks
later. Our findings provide further evidence for the large variability of
NPI scores within individuals with neurocognitive disorders.\(^1,7–9\)
When looking at specific NPI domains, we found lowest stability over
time for anxiety, hallucinations, depression, and disinhibition, which is
in line with prior test-retest studies.\(^10–13\) Our findings extend previous
studies by looking at trajectories over a period over several weeks
compared to the commonly examined (bi)annual NPI assessments.\(^1,7–9\)

The large within-person variation in NPI scores could reflect
substantial fluctuations in the manifestation of NPS in patients visiting
the memory clinic. Previous studies that have used diaries to
daily assess NPS in dementia suggested a rather erratic nature of
NPS.\(^14–18\) This is in line with the growing body of evidence empha-
sizing the role of proximal causes of NPS including psychosocial
factors (e.g., caregiver burden, caregiver communication style),
environmental factors (e.g., light, temperature), and somatic condi-
tions (e.g., pain, thirst).\(^17,19\)

Alternatively, the irregular course of NPI scores could also arise
from methodologic issues related to the NPI. Our finding that
changes in NPI scores were unrelated to changes in caregiver burden
could support this. Several factors could affect the NPI scores that
are unrelated to the actual manifestation of NPS in our sample. First,
caregivers tend to use different terminologies to describe NPS
in dementia compared to the commonly examined (bi)annual NPI assessments.\(^1,7–9\)

Consequently, caregivers may have endorsed different NPI domains
during follow-up assessments, although similar NPS were present
during the course of the study. Furthermore, although recall bias was
reduced because caregivers were asked to evaluate the presence of
NPS during the last 2 weeks instead of the commonly used 4 weeks,
the recollection of NPS remains challenging.\(^20\) Moreover, mood, fa-
tigue, and distress among caregivers can affect the NPI
administration.\textsuperscript{21,22} To overcome these challenges, future studies could pair repeated NPI assessments with daily NPS measurements using an Ecological Momentary Assessments approach.\textsuperscript{23}

Also, the variation in NPI scores could be an effect of unknown measurement error related to the NPI as little is known about what we should consider as actual change in NPI scores. Different statistical methods such as the standard error of measurement and the reliable change index have been developed to determine the minimal detectable change of clinical outcome scales.\textsuperscript{24,25} These methods have been used to establish minimal detectable change for the individual domains of the nursing home version of the NPI after 2 weeks and the NPI total score after 52 weeks.\textsuperscript{13,26} However, these psychometric indices establish minimal detectable change and do not determine minimally important change, that is, clinically meaningful change.\textsuperscript{24,25} Anchor-based approaches can be used to determine clinically meaningful change by which changes on an instrument are compared with minimally important changed defined by patients, caregiver, and/or clinicians. Future studies are needed that align NPI trajectories with anchor definitions of meaningful change in NPS to establish which changes in NPI scores we should consider as clinically meaningful.

4.1 | Strengths and limitations

Strengths of this study are the inclusion of a representative tertiary memory clinic sample consisting of various clinical diagnoses and the
low level of psychotropic medications used across patients. There are also some limitations to our study. First, the majority of the participants were still undergoing diagnostic workup and received a diagnosis at some point during the study. This may have affected the manifestation of NPS as receiving a diagnosis can have great psychological impact. Second, we included a small and clinically heterogeneous memory clinic population. The low sample size may affect the stability of correlation coefficients, especially the correlations below 0.50.27 Furthermore, the proportion of female patients (22%) in our sample was lower than expected based on previous studies in Dutch academic memory clinics (40%–55% females).1,28,29 Although our within-person analysis reduces the potential impact of the clinical heterogeneity and underrepresentation of female patients, our results need to be replicated in larger samples including a higher proportion of females, especially since NPS may manifest differently in females than males.30 Third, we found indications that NPI scores were somewhat more stable in individuals with dementia compared to individuals with CIND. This suggest that the NPI may be more appropriate to repeatedly assess NPS when used in individuals with dementia, which could be expected as the NPI was originally developed and validated to measure NPS in dementia.31 Yet, future studies with larger samples are needed to examine the effects of demographic characteristics and clinical characteristics such as dementia type on the short term trajectories of NPI scores. Finally, no clear cutoffs exist for measures used in this study (e.g., within-person variation) making the comparison between NPI domains somewhat subjective.

4.2 Conclusions

This study suggest highly unstable NPI scores when assessed at 2-week intervals. These findings question the reliability of NPI scores when administered at short-term intervals at the memory clinic, but also as outcome measure in trials that evaluate the effectiveness of (non)pharmacological interventions, especially for those who do not meet diagnostic criteria for dementia (i.e., CIND). Further studies are needed to investigate whether the large within-person variability of NPI scores reflect the erratic nature of NPS in neurocognitive disorders or arise from methodological issues. Although the origin of these fluctuations remains unclear, memory clinic clinicians should be aware that NPI scores at one time point are poorly related to future NPI scores within a timeframe of weeks.

AUTHOR CONTRIBUTIONS

Willem S. Eikelboom designed the study, analyzed the data, and wrote the paper. Amy den Teuling collected the data and assisted with writing the paper. Daphne E. Pol collected the data and assisted with writing the paper. Michiel Coesmans assisted with writing the paper. Sanne Franzen collected the data and assisted with writing the paper. Judy van Hemmen collected the data and assisted with writing the paper. Ellen Singleton assisted with writing the paper. Rik Ossenkoppele assisted with writing the paper. Frank Jan de Jong assisted with writing the paper. Esther van den Berg designed the study and assisted with writing the paper. Janne M. Papma designed the study, supervised the study, and assisted with writing the paper.

ACKNOWLEDGMENTS

Dr. Janne Papma and Dr. Rik Ossenkoppele were supported by an Alzheimer Nederland and Memorabel ZonMw Grant 733050823 (Deltaplan Dementie).

CONFLICT OF INTEREST

The authors declared that they have no conflicts of interest to this work.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ETHICS STATEMENT

This study received ethical approval from the Medical Ethics Committee of the Erasmus University Medical Center (MEC-2020-0341). All participants gave informed consent.

ORCID

Willem S. Eikelboom https://orcid.org/0000-0001-5004-5141
Sanne Franzen https://orcid.org/0000-0002-4461-5109
Rik Ossenkoppele https://orcid.org/0000-0003-1584-7477
Esther van den Berg https://orcid.org/0000-0002-8120-7366
Janne M. Papma https://orcid.org/0000-0002-2790-438X

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.

---

**How to cite this article**: Eikelboom WS, denTeuling A, Pol DE, et al. Biweekly fluctuations of neuropsychiatric symptoms according to the Neuropsychiatric Inventory: erratic symptoms or scores? *Int J Geriatr Psychiatry*. 2022;1-7. https://doi.org/10.1002/gps.5770