https://doi.org/10.5664/jcsm.9830

JCSM | Journal of Clinical Sleep Medicine

SCIENTIFIC INVESTIGATIONS

Validation of the Dutch translation of the Paris Arousal Disorders Severity Scale for non-REM parasomnias in a 1-year and 1-month version

Petra van Mierlo, MSc[1,*]; Lieke W.A. Hermans, PhD[2,*]; Isabelle Arnulf, MD, PhD[3,4]; Angelique Pijpers, MD, PhD[1]; Sebastiaan Overeem, MD, PhD[1,2]; Merel M. van Gilst, PhD[1,2]

[1]*Sleep Medicine Center Kempenhaeghe, Heeze, The Netherlands;* [2]*Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands;* [3]*Sleep Disorder Unit, Pitié-Salpêtrière Hospital, AP-HP, Paris, France;* [4]*Sorbonne University, and Paris Brain Institute (ICM), Paris, France; *Joint first authors*

Study Objectives: We created a Dutch version of the Paris Arousal Disorders Severity Scale (PADSS), which assesses non–rapid eye movement (NREM) parasomnia symptoms over the past year (PADSS-year). This questionnaire was previously validated in patients with sleep walking and/or sleep terrors (SW/ST). We validated the questionnaire in SW/ST patients, and in a broader population, including patients with confusional arousals, comorbidities, and medication users ("other NREM parasomnias"). Furthermore, we introduced a version covering the past month (PADSS-month), with the potential purpose of evaluating symptom evolution and treatment response.

Methods: We compared PADSS scores among 54 SW/ST patients, 34 age-matched controls, and 23 patients with other NREM parasomnias. We evaluated discriminative capacity, internal consistency, and construct validity. Furthermore, we assessed the test–retest reliability and treatment response of PADSS-month.

Results: Healthy controls scored significantly lower than both patient groups. We found an excellent diagnostic accuracy (area under the curve PADSS-year 0.990, PADSS-month 0.987) and an acceptable internal consistency. Exploratory factor analysis identified 3 components: "behaviors outside the bed," "behaviors in/around the bed," and "violent behaviors," with the former 2 factors reflecting the distinction between SW and ST. PADSS-month showed an acceptable test–retest reliability (0.75). Additionally, PADSS-month significantly decreased after pharmaceutical and/or behavioral treatment. This change was correlated with the clinical impression of the caregiver, implying that PADSS-month is sensitive to treatment effects.

Conclusions: The Dutch PADSS questionnaire can be used as a screening tool in a broad population of patients with NREM parasomnia, not only SW/ST. Furthermore, we validated a PADSS-month version to assess the evolution of symptoms and treatment effect.

Keywords: NREM parasomnia, sleep terror, sleepwalking

Citation: van Mierlo P, Hermans L, Arnulf I, Pijpers A, Overeem S, van Gilst M. Validation of the Dutch translation of the Paris Arousal Disorders Severity Scale for non-REM parasomnias in a 1-year and 1-month version. *J Clin Sleep Med*. 2022;18(4):1135–1143.

BRIEF SUMMARY

Current Knowledge/Study Rationale: The Paris Arousal Disorders Severity Scale (PADSS) is one of the few questionnaires for assessing parasomnia-related symptoms. The PADSS was developed for and tested in a selected group of participants with sleepwalking and sleep terrors, and it evaluates symptoms over a relatively long period of one year.

Study Impact: We created and validated a Dutch version of the PADSS in patients with sleep walking and/or sleep terrors and showed that the PADSS can also be successfully used as a screening tool in other parasomnias. Furthermore, we validated a version of the PADSS covering the past month, which could be used to assess the evolution of symptoms, as well as treatment effect.

INTRODUCTION

Non–rapid eye movement (NREM) parasomnias (also called arousal disorders) are abnormal behaviors that are assumed to arise upon incomplete arousal from N3 sleep.[1,2] Patients may show disoriented behavior, sit up in bed, talk, scream, walk, or handle objects. These inappropriate behaviors are accompanied by impaired consciousness, and typically there is no or poor recollection of the event the following morning.[1,3] NREM parasomnias are quite common during childhood and adolescence and usually do not continue in adulthood. However, sometimes they remain present or appear in adulthood. The prevalence of NREM parasomnias in adults was estimated to be 6.9% for confusional arousals, 1.7% for sleep walking (SW), and 2.7% for sleep terrors (ST).[4,5]

NREM parasomnias are often considered semi-normal behaviors during sleep, with no need for treatment or physician consultation. Still, NREM parasomnias can sometimes have a high impact on daily life, causing anxiety, daytime sleepiness, social embarrassment, or disturbance of the sleep of the patient, bed partner, or other house members. Additionally, NREM parasomnias, particularly SW and ST can lead to dangerous situations—for example, patients may injure themselves or others. These can all be reasons to visit a sleep specialist.

A diagnosis for NREM parasomnias is currently made based on the *International Classification of Sleep Disorders*, third edition, criteria.[6] The diagnosis requires an extensive interview with the patient, usually followed by a video-polysomnography or post-deprivation PSG, to support the clinical diagnosis and

to exclude other sleep disorders. A quick screening tool, allowing patients to accurately indicate the occurrence, frequency, and severity of their clinical symptoms for arousal disorders, could be very helpful for the diagnostic process. Furthermore, there is need for a quantitative tool to measure the changes in the severity of the symptoms over time—for example, during and after treatment.

The Paris Arousal Disorders Severity Scale (PADSS) was developed in 2014 to evaluate the clinical symptoms and severity of arousal disorders.[7] The PADSS is one of the few questionnaires for assessing parasomnia-related symptoms that are currently available. The PADSS showed a high sensitivity, specificity, internal consistency, and test–retest reliability.[7] Additionally, it is easy to complete within 5 minutes. However, the PADSS was developed for and tested in a selected group of participants with SW and ST, then tested in 158 patients with SW/ST,[8] in patients with sexsomnia,[9] and in patients with rapid eye movement (REM) sleep behavior disorder (a REM sleep parasomnia)[7,10] in France. However, it would be useful to find out if the PADSS is also suitable in another country, and in patients with other types of parasomnias related to NREM sleep, patients using medication, and patients with comorbid sleep disorders. Additionally, the PADSS addresses symptoms over the past year, which is a fairly long period if one would want to evaluate the response to treatment. In a clinical setting, it would be preferable to evaluate the effect of treatment over a shorter time span. Given the night-to-night variability in NREM parasomnias, an evaluation window of 1 month seems to be appropriate.

This study has 4 goals: (1) translating the French PADSS questionnaire into Dutch, (2) validating the Dutch version of the questionnaire in a group of SW/ST patients, (3) validating the Dutch version of the questionnaire in other patients with NREM parasomnia, and (4) validating a short-term version of the questionnaire, assessing the severity of the arousal disorder during the past month.

METHODS

Study design

The PADSS was translated from French to Dutch. Forward and backward translation of the questionnaire was performed by 2 independent translators. We first validated the Dutch version of the translated PADSS questionnaire (year-version) in a selected group of patients with SW and/or ST, which matched the inclusion and exclusion criteria of Arnulf et al.[7] ("SW/ST patients"). These patients were compared with a group of 34 heathy controls, who were matched group-wise for age and sex. Additionally, we also used the questionnaire in a broader group of patients with less strict exclusion criteria ("other NREM parasomnia patients"). Finally, we validated a short-term version of the questionnaire (month-version) using the same approach.

PADSS scale

The PADDS scale is a self-reported scale that is completed by the patient, consisting of 3 parts (A = parasomnia behaviors, B = frequency, C = consequences).[7]

Year-version

A 5-point Likert scale is used to rate the items (for parts A and C: 0 = never, 1 = sometimes, 2 = often; for part B: 6 = 2 episodes or more per night, 5 = 1 episode per night, 4 = at least 1 episode per week, 3 = at least 1 episode per month, 2 = at least 1 episode per year, 1 = less than 1 episode per year, 0 = never any motor episode).

Month-version

For the month-version of the scale, a few adjustments were made to the original scale (**Figure S1** in the supplemental material). In part A, patients were asked if they exhibited the listed behaviors during the past month instead of the past year. In part B, the frequency of the abnormal episodes over the past month was addressed. Thus, the answer option "at least 1 episode per year" was removed and the answer option "less than 1 episode per year" became "less than 1 episode per month" (5 = 2 episodes or more per night, 4 = 1 per night, 3 = at least 1 episode per week, 2 = at least 1 episode per month, 1 = less than 1 episode per month, 0 = never any motor episode). Part C (effects of the abnormal behaviors) remained unchanged.

Participants

Healthy controls

Healthy controls were recruited from the general population. Healthy controls were included to the study if they were 18 years or older and were not under treatment for a sleep disorder, including parasomnia. Healthy controls were excluded if they had a history of a sleep disorder and in case of psychiatric or neurological disorders.

Patients with NREM parasomnia

The participating patients were selected from all consultive outpatients visiting the sleep medicine center Kempenhaeghe (Heeze, The Netherlands) with unwanted behaviors during sleep and a suspicion of a NREM parasomnia between April 2016 and June 2018. Patients were included if they were older than 18 and diagnosed with parasomnia related to NREM sleep. We included patients with confusional arousals, SW, ST, parasomnia due to a medical condition, and unspecified parasomnia. We excluded patients with REM-related parasomnias, nightmare disorder, and those with suspicion of nocturnal frontal lobe epilepsy. Each diagnosis was made according to the *International Classification of Sleep Disorders*, third edition, criteria,[6] based on an extensive patient interview by an European Sleep Research Society - accredited (ESRSaccredited) experienced somnologist. Almost all participants underwent a diagnostic study: overnight video-PSG (n = 47), video-PSG after 1 night of sleep deprivation (n = 37), both overnight and sleep deprivation PSG (n = 1), or PSG at another hospital (n = 1). PSG was used to exclude the presence of other sleep disorders that could explain the clinical symptomatology. In addition, polysomnographic outcomes supporting the NREM parasomnia were the presence of direct awakenings from stage N3 and/or associated observed nighttime movements/behaviors. Two patients chose to not undergo a diagnostic study, and

therefore the diagnosis was based on the patient interview alone.

SW/ST group

The SW/ST group was a subset of the total group of patients with NREM parasomnia, containing patients who matched additional stricter inclusion and exclusion criteria. Patients in this group all had a diagnosis of SW and/or ST. For this group, we only selected patients who were not using medication that potentially influences parasomnia complaints (such as benzodiazepine receptor agonists, antiepileptic drugs, antidepressants, antipsychotic medications, or beta-blockers). We also did not select participants with an apnea-hypopnea index (AHI) ≥15 events/h.

Other NREM parasomnia patients

The other NREM parasomnia patients group refers to the group of patients with NREM parasomnia who did match the broader inclusion criteria of the study but did not meet the additional criteria for the SW/ST group. This included patients with NREM parasomnias other than SW and/or ST, patients with an AHI ≥15 events/h, and patients who used medication potentially influencing parasomnia complaints.

The study was approved by the institutional ethics committee of sleep medicine center Kempenhaeghe (protocol number 15.19). All participants provided informed consent.

Study procedures

A schematic representation of the study procedures is shown in Figure 1. All medical assessments and treatments in this study were part of standard medical care. Within this usual procedure, patients were asked to complete the PADSS at different time points. Patients who were referred to the sleep center with a suspicion of an arousal disorder were asked to complete both the year- and month-version of the PADSS upon their first visit to the sleep center (PADSS-year and PADSS-month-1). The order of the versions was counterbalanced across participants. A variable number of weeks after their first visit, patients underwent either a sleep-deprivation PSG or a standard 1-night PSG as part of the standard diagnostic process.

During the visit for evaluation of the PSG results with the physician (usually approximately 6–8 weeks after the diagnostic PSG), the patients were asked to complete the month-version of the PADSS again (PADSS-month-2). Finally, the month-version of the PADSS was completed for a third time at the end of the treatment period (PADSS-month-3). The physicians or psychologists who were treating the patients were asked to complete the treatment evaluation question (question 2, Change) of the Clinical Global Impression (CGI) Scale at the end of the treatment (CGI-C). The healthy controls completed the year-version and the month-version of the PADSS at only 1 time point. Again, the order of the versions was counterbalanced across participants.

In some situations, a NREM parasomnia was observed during clinical PSG, while this was not directly suspected beforehand. In those cases, PADSS-year and PADSS-month-1 were completed during the second appointment and PADSS-month-2 was omitted.

Treatment

Because there are no large, controlled studies evaluating treatment options in people with NREM parasomnia,[2,11] interventions are usually selected based on the local protocol of the sleep center. At sleep center Kempenhaeghe, the treatment of choice for NREM parasomnia is usually cognitive behavioral therapy for parasomnia (CBT-p), pharmacological treatment using gabapentin or clonazepam, or a combination of both. Furthermore, in some cases, the intervention involves treatment of factors that can trigger NREM parasomnia, such as obstructive breathing events in comorbid obstructive sleep apnea. The choice of a specific type of intervention is based on the assessment of the treating physician and the preferences of the patient.

Clinical Global Impression

The CGI is an established self-reported scale that is used to evaluate a clinician's view of the severity of a patient's illness before and after treatment.[12] For this study, we used the second question of the CGI to measure therapeutic treatment response (CGI-C).[12] A 7-point Likert scale is used to rate the improvement of the patient after treatment: 1 = very much improved, 2 = much improved, 3 = minimally improved,4 = no change, 5 = minimally worse, 6 = much worse, 7 = very much worse.[12]

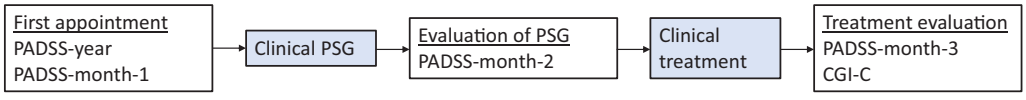
Analyses

Evaluation and validation of the Dutch PADSS scale (before treatment; year- and month-versions)

First, we validated the Dutch version of the PADSS scale in the SW/ST patients. We performed the analysis in this subgroup to facilitate comparison with the French cohort of Arnulf et al, which was subject to the same inclusion and exclusion criteria.

The total scores and subscores of PADSS-year and PADSS-month-1 were compared between SW/ST patients, healthy age-matched controls, and other NREM parasomnia patients. Additionally, the receiver operating characteristics (ROC)

Figure 1—Schematic representation of the study procedures.



Blue rectangles represent procedures that were part of clinical treatment; white rectangles represent time instances during which data were collected for the study. GGI-C = treatment evaluation question (reflecting change between before and after treatment) of the Clinical Global Impression Scale, PADSS = Paris Arousal Disorders Severity Scale, PSG = polysomnography,

curve was used to evaluate the capacity of the scale to discriminate between healthy sleepers and SW/ST patients (area under the curve [AUC]). From this ROC curve, we determined the optimal cutoff value, which was defined as the value with the lowest distance to the 100/100% point in the graph. Subsequently, we calculated the sensitivity of distinguishing other NREM parasomnia patients from healthy sleepers, based on this previously determined cutoff value.

Furthermore, we assessed the internal consistency of PADSS-year and PADSS-month in the SW/ST patients using the Cronbach's alpha coefficient. We evaluated the internal consistency of both the total PADSS and subpart A. We also performed exploratory factor analysis in the SW/ST patients to assess construct validity. We a priori limited the analysis to 2 or 3 factors to match the analysis performed by Arnulf et al. Finally, to evaluate the stability of the month-version of the PADSS, we assessed the correlation between PADSS-month-1 and PADSS month-2 in the SW/ST patients. We also assessed if there was a statistically significant difference between PADSS-month-1 and PADSS-month-2 in this group.

Evaluation of treatment effect (month-version)

For this part of the analysis, we combined the SW/ST patients and the other NREM parasomnia patients into 1 group. To evaluate the effect of treatment, we selected all participants who completed both PADSS-1 and PADSS-3. We assessed the difference between PADSS-1 and PADSS-3 for this subset. We also analyzed the difference between PADSS-1 and PADSS-3 for subgroups stratified for type of treatment. Finally, we assessed the correlation between CGI-C and the difference between PADSS-1 and PADSS-3, for both the total scale and subpart C.

Statistics

Descriptive values are shown as mean ± standard deviation (SD) in case of normally distributed variables and as median ± interquartile range (IQR) in case of non–normally distributed variables. For the statistical tests, we used analysis of variance (ANOVA) and Pearson correlation tests in case of normally distributed variables and Kruskal-Wallis tests and Spearman correlation tests in case of a nonnormal distribution. The statistical analysis was done using SPSS (IBM Corporation, Armonk, NY). In case of missing answers to sub-questions, we performed listwise deletion of both the sub-part that was not answered and the total PADSS score.

RESULTS

Participants

For the initial validation of the scale, 54 participants with SW and/or ST were included (32 with SW, 17 with ST, and 5 mixed). These participants had a median AHI of 1.5 events/h and IQR of 4.6 (0–14.4) events/h and a median periodic limb movement (PLM) index of 4.2 and IQR of 11.0 (0–26). Seven SW/ST patients had a comorbid insomnia disorder. Additionally, 34 age-matched healthy controls participated in the study.

The other NREM parasomnia group consisted of 23 additional patients. Of these, 14 participants had another type of NREM parasomnia than SW and/or ST, including confusional arousals (n = 11), parasomnia due to a medical condition (n = 2), and unspecified parasomnia (n = 1). Additionally, 7 patients had an AHI ≥ 15 events/h and 5 patients used medication that can affect sleep and NREM parasomnias. The other NREM parasomnia patients had a median AHI of 5.8 events/h and IQR of 14.9 (0.5–27.6) event/h and a median PLM index of 6.7 and IQR of 32.6 (0–89.5). The total group of patients (SW/ST and other NREM parasomnia patients together; n = 78) had a median AHI of 2.2 events/h and IQR of 6.4 (0–27.6) events/h and a median PLM index of 4.6 and IQR of 15.8 (0–89.5). Demographic characteristics of the participants are shown in **Table 1**.

In 4 SW/ST patients and 2 other NREM parasomnia patients, there was no suspicion of a parasomnia at the first visit (diagnosis was made after PSG) or first-visit questionnaires were missing, and therefore, PADSS-year and PADSS-month-1 were completed during the second appointment.

PADSS scores at the first visit

In the SW/ST group, 1 patient (∼1.9%) did not answer category B of the PADSS-year and was therefore omitted from analysis for PADSS-year-B and PADSS-year-total. In the other parasomnia group, 1 patient (∼4.3%) did not completely answer category A of the PADSS-year. All healthy participants answered all questions of the PADSS-year. The month-version (PADSS-month-1) was completed by all patients from the SW/ST group and all healthy participants. PADSS-month-1 was not completed by 1 participant (∼4.3%) from the other parasomnia group, and 1 other patient from the same group again did not completely answer category A. From the SW/ST group, 5 patients (∼9.3%) answered positively to question 16 ("I unwillingly performed a sexual act") and 6 patients (∼11.1%) answered positively to question 17 ("I prepared or ate some food or a drink"). In the other parasomnia group, 2 patients (∼5.9%) positively answered question 16 and 2 patients (∼5.9%) positively answered question 17.

PADSS scores of the different groups for the year- and the month-versions are shown in **Figure 2**. Overview statistics of the PADSS questionnaires and their components, as well as statistical comparisons between the SW/ST group and the healthy controls, are listed in **Table 1**. A significant difference between the total score of PADSS-year between the 3 groups was found. Post hoc analysis indicated that the healthy participants significantly differed from both patient groups. The same statistical differences were found for subparts A, B, and C of PADSS-year; the total score of PADSS-month-1; and subparts A, B, and C of PADSS-month-1. The Dutch SW/ST patients had lower total scores on the PADSS-year version than the previously published French cohort (mean ± SD: 16.8 ± 4.5 vs 19.4 ± 6.3, $t = -2.7$, $P = .008$).

Discriminative capacity

The ROC analyses indicated that the total score of PADSS-year had an excellent diagnostic accuracy with an AUC of 0.990 (**Figure 3A**). The optimal cutoff was between 8 and 10

Table 1—Demographics and PADSS scores.

	Patients		Controls (n = 34)	Statistical Test
	SW/ST (n = 54)	Other NREM Parasomnias (n = 23)		
Age, median ± IQR (years)	26 ± 7 [18–52]	36 ± 15 [22–78]*	27 ± 8 [19–39]	–
Sex (% males)	37.0	46.0	38.2	–
PADSS-year PADSS total	17 ± 7 [7–26]	15 ± 7 [9–24]	0 ± 2 [0–12]*	Kruskal-Wallis H = 68.2, $P < .001$
PADSS-A	8 ± 5 [2–16]	7 ± 5 [1–14]	0 ± 0 [0–5]*	Kruskal-Wallis H = 66.3, $P < .001$
PADSS-B	4 ± 1 [3–6]	4 ± 1 [0–6]	0 ± 0 [0–4]*	Kruskal-Wallis H = 67.3, $P < .001$
PADSS-C	4 ± 2 [1–8]	5 ± 2 [2–7]	0 ± 0 [0–4]*	Kruskal-Wallis H = 66.6, $P < .001$
PADSS-month-1 PADSS total	14 ± 6 [5–30]	12 ± 9 [7–20]	0 ± 0 [0–11]*	Kruskal-Wallis H = 66.2, $P < .001$
PADSS-A	6 ± 4 [1–19]	6 ± 5 [1–11]	0 ± 0 [0–5]*	Kruskal-Wallis H = 64.8, $P < .001$
PADSS-B	3 ± 1 [2–5]	4 ± 1 [1–5]	0 ± 0 [0–3]*	Kruskal-Wallis H = 57.3, $P < .001$
PADSS-C	4 ± 2 [1–7]	5 ± 1 [2–6]	0 ± 0 [0–3]*	Kruskal-Wallis H = 69.2, $P < .001$

All PADSS scores are shown as median ± IQR because the scores of the healthy participants did not have a normal distribution. *Post hoc testing indicated a significant difference between this group and the 2 other groups. IQR = interquartile range, NREM = non–rapid eye movement, PADSS = Paris Arousal Disorders Severity Scale, SW/ST = sleep walking/sleep terrors.

(specificity 91.2%, sensitivity 98.1%; none of the participants had a total score of 9). Applying the threshold of between 8 and 9 in the total PADSS-year of the other NREM parasomnia patients resulted in a sensitivity of 100% and applying the threshold of between 9 and 10 resulted in a sensitivity of 95.5%. Thus, we would recommend using a threshold of 8/9, implying that a score < 9 is not indicative of NREM parasomnia and a score ≥ 9 is indicative of NREM parasomnia.

The total score of PADSS-month-1 also had an excellent diagnostic accuracy with an AUC of 0.987 (**Figure 3B**). The cutoff value with the lowest distance to the 100/100% point in the graph was between 5 and 7 (specificity 91.2%, sensitivity 98.1%; none of the participant had a total score of 6). Because all other NREM parasomnia had a PADSS-month-1 score larger than 6, both thresholds 5/6 and 6/7 resulted in a sensitivity of 100%. We would recommend using a threshold of 5/6, implying that a score < 6 is not indicative of NREM parasomnia and a score ≥ 6 is indicative of NREM parasomnia.

Internal consistency

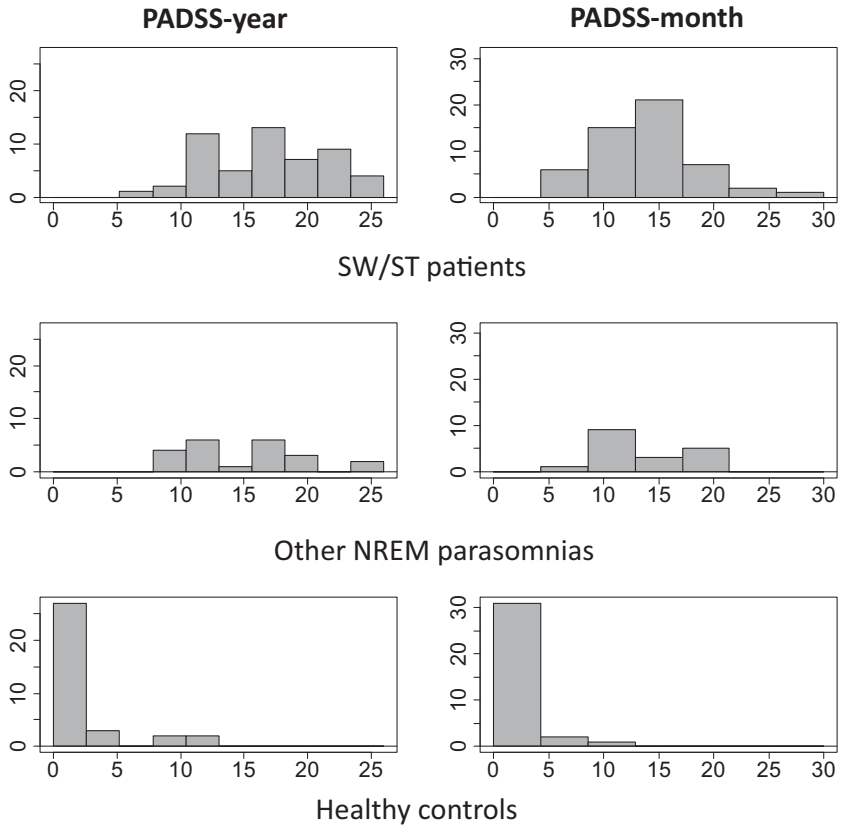
In the SW/ST patients, the internal consistency of the year-version was Cronbach alpha = 0.705 for the total PADSS and Cronbach alpha = 0.702 for PADSS-A. The internal consistency of the month-version was Cronbach alpha = 0.733 for the total PADSS and Cronbach alpha = 0.763 for PADSS-A.

Construct validity

For the SW/ST patient group, the value of the Kaiser-Meyer-Olkin measure of sample adequacy for the year-version was 0.602 and the Bartlett test of sphericity was significant ($P < .001$), indicating that the data were suitable for structure detection. Question 9 ("I climbed out a window") did not exhibit any variance, because all patients answered it with "never." We applied a 2-factor model, in the same way as Arnulf et al. The 2-factor model accounted for 38.1% of the scale variance. The factor loadings can be found in **Table S1** in the supplemental material. A component plot can be found in **Figure S2**. Factor 1 consisted of 10 questions (Q5–8, 10, 11, 13–16), which were mostly related to things that people usually do outside the bed (eg, climbing the stairs, handling objects, preparing food or drinks). Factor 2 consisted of 4 questions (Q1–4), which consisted of acts that may take place in or around the bed (screaming, sitting up in bed, hitting or kicking someone, falling out of bed). One question (Q15: touching things around windows and openings) was related to both components, but correlated most with factor 1. Two questions (Q12: breaking objects; Q17: unwillingly performing a sexual act) were not correlated with any category. These 2 questions had a low average score (Q12 0.11 and Q17 0.09), indicating that these acts were only reported by a few participants.

Because the 2-factor model only accounted for a relatively small part of the scale variance and our factors were different

Figure 2—Total scores for first-visit PADSS questionnaires for SW/ST patients, healthy controls, and other parasomnia patients.



NREM = non–rapid eye movement, PADSS = Paris Arousal Disorders Severity Scale, SW/ST = sleep walking/sleep terrors.

than the factors found by Arnulf et al, we reran the analysis for a 3-factor model. The 3 factors accounted for 48.6% of the variance. The factor loadings can be found in **Table S2**. In the 3-factor model, factor 1 contained the same factors as in the 2-factor model with similar factor loadings. Factor 2 (Q1, Q2, Q4) also largely remained the same, except that question 3 moved to factor 3. Factor 3 consisted of Q3 (hitting or kicking someone) and Q4 (unwillingly performed a sexual act). Question 12 (breaking objects) was still not correlated with any of the categories. Two- and 3-factor models for the month questionnaire (PADSS-month-1) yielded similar factors.

Correlation between first and second month-versions

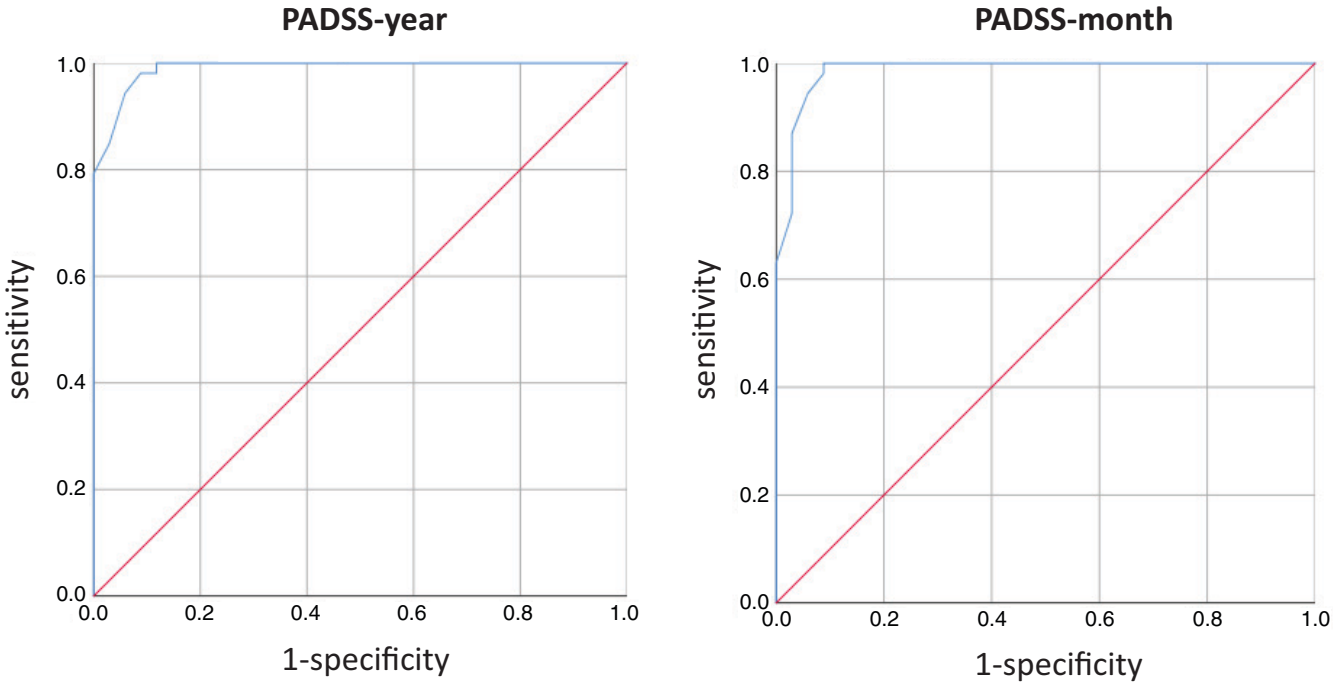
A total of 41 patients from the SW/ST group completed both PADSS-month-1 and PADSS-month-2. The correlation between PADSS-month-1 and PADSS-month-2 in these patients was 0.75 (Pearson $P < .001$) for PADSS-total, 0.67 (Spearman $P < .0001$) for PADSS-A, 0.55 (Spearman $P < .001$) for PADSS-B, and 0.56 (Pearson $P < .001$) for PADSS-C. A paired t test indicated that there was no statistical difference between total PADSS scores for the first and the second month-versions in the SW/ST group ($t = 1.61$, $P = .116$).

Treatment effect

For the statistical evaluation of the diagnostic effect, we selected all SW/ST patients and all other NREM parasomnia patients who completed PADSS-month-1 (baseline; before treatment) and PADSS-month-3 (after treatment).

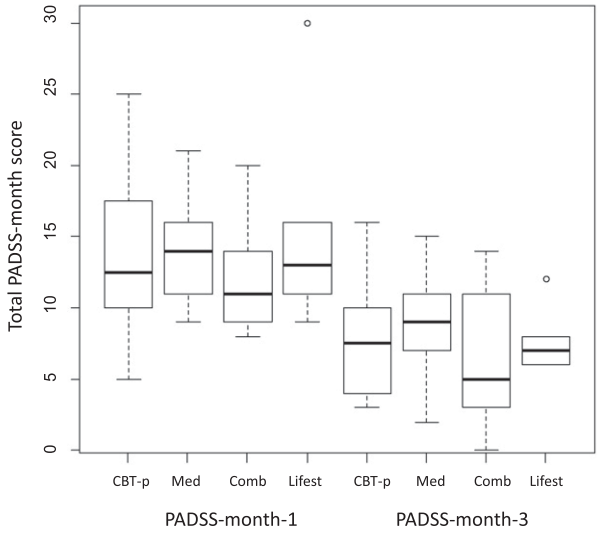
This yielded 45 participants, of whom 21 participants were treated with CBT-p, 9 participants received medication (gabapentin or clonazepam), 5 participants received both medication and CBT-p, and 8 participants did not receive any treatment apart from the standard lifestyle advice that was provided to all patients. Two participants received various other types of treatment: posture training for comorbid obstructive sleep apnea complaints, CBT-p, and continuous positive airway pressure therapy for obstructive sleep apnea complaints. **Figure 4** shows total scores of PADSS-month-1 and PADSS-month-3 for subgroups of participants based on type of treatment. When grouping all subgroups together and performing a pairwise comparison of the total scores of PADSS-1-month to PADSS-3-month, we found a statistically significant difference (Wilcoxon $V = 983$, $t < 0.001$), indicating that the total PADSS score decreased between the first and the last appointment. When performing the same analysis for the subgroups, significant differences were found for all groups: participants who received CBT-p ($t = 6.14$, $P < .001$), participants who received

Figure 3—ROC curve for discriminating between SW/ST patients and age-matched healthy participants.



The red diagonal lines are reference lines; the blue lines represent the ROC curves. PADSS = Paris Arousal Disorders Severity Scale, ROC = receiver operating characteristic, SW/ST = sleep walking/sleep terrors.

Figure 4—Total scores for PADSS-month-1 (before treatment) and PADSS-month-3 (after treatment) for subgroups of participants stratified for type of treatment.



The 2 participants who received various other types of treatment and the 2 participants for whom the treatment was not known are not shown in this graph. CBT-p = standard lifestyle advice + cognitive behavioral therapy (n = 21), Comb = standard lifestyle advice + cognitive behavioral therapy + medication (n = 5), Lifest = only standard lifestyle advice (n = 8), Med = standard lifestyle advice + medication (n = 9), PADSS = Paris Arousal Disorders Severity Scale.

medication ($t = 3.42$, $P = .009$), participants who received CBT-p and medication ($t = 2.83$, $P = .047$), and participants who only received lifestyle advice ($t = 2.49$, $P = .042$).

CGI-C was completed in 40 out of 45 participants. The CGI-C indicated "1 = very much improved" in 8 patients, "2 = much improved" in 22 patients, "3 = minimally improved" in 6 patients, and "4 = no change" in 4 patients. None of the participants had a CGI-C score > 4, indicating that the symptoms of all patients improved or stayed the same. A similar conclusion could be drawn for the difference between PADSS-month-1 and PADSS-month-3, where only 4 patients had a higher score at the end compared with the first visit (of these 4 patients, 2 had a CGI-C score of 4, 1 had a CGI-C score of 2, and 1 did not have a CGI-C score) and 1 patient had the same score before and after treatment. A moderately strong correlation was found between CGI-C and the difference between the total scores of PADSS-month-1 and PADSS-month-3 (Spearman rho = 0.312, $P = .0498$). We also found a moderately strong, statistically significant correlation between CGI-C and the difference between part C (consequences) of PADSS-month-1 and PADSS-month-3 (Spearman rho = 0.360, $P = .023$). No correlation was found between CGI-C and the difference between part B (frequency) of PADSS-month-1 and PADSS-month-3 (Spearman rho = 0.290, $P = .070$).

DISCUSSION

The aim of this study was to validate a translated Dutch version of the PADSS in patients with SW and/or ST ("SW/ST

patients"), as well as in a broader spectrum of people with NREM parasomnias with less strict exclusion criteria ("other NREM parasomnia patients"). Additionally, we introduced a shorter-term version of the PADSS, assessing the severity of the arousal disorder during the past month.

The Dutch version of the PADSS questionnaire (year-version) showed an acceptable internal consistency and an excellent diagnostic accuracy when distinguishing participants with SW/ST from age-matched healthy controls. An optimal cutoff value of ≥ 9 was determined (ie, a score < 9 is not indicative of NREM parasomnia and a score ≥ 9 is indicative of NREM parasomnia). This cutoff value resulted in a very high sensitivity in the other NREM parasomnia patients. From visual evaluation, the PADSS scores of SW/ST and other patients seemed comparable. The distribution was the same, without ceiling and floor effects. We found statistical differences between healthy participants and both SW/ST patients and other NREM parasomnia patients (but not between the 2 patient groups), for both the year- and the month-version of the questionnaire. These results together indicate that the PADSS questionnaire can be used as a diagnostic tool for both SW/ST patients and other patients with other NREM parasomnias, patients using medication for various reasons other than parasomnia, and/or patients with comorbid sleep disorders.

Our results on the Dutch year-version of the questionnaire showed a number of differences compared with the previously published French version of the questionnaire. Most importantly, the Dutch SW/ST patients had lower total scores on the PADSS year-version than the previously published French cohort, and the optimal cutoff value from the ROC was also lower. This could point toward a culturally defined different way of thinking about NREM parasomnia. Alternatively, it could mean that people with mild complaints of NREM parasomnia are more easily referred to a sleep center in the Netherlands. The identified factors also differed between the French version and the Dutch version. Arnulf et al identified factors that they interpreted as "wandering" and "violence," and we found 2 other factors that we interpreted as "outside the bed" and "in/around the bed." The 2-factor model of Arnulf et al accounted for a larger percentage of the scale variance than our 2-factor model. We expect that these differences were partly caused by the fact that the Dutch participants reported low incidences of violent and dangerous behaviors, which probably coincides with their overall lower PADSS scores compared with the French participants. Additionally, we could speculate that our factor "in/around the bed" includes questions that would correspond to symptoms of ST. Of the French participants, only 5% of the participants had isolated ST, while in our study sample, 33% of the participants had isolated ST. This difference could possibly explain the difference between the identified factors. Indeed, when we repeated the factor analysis with a 3-factor model, we found factors corresponding to "outside the bed," "in/around the bed," and "violent behavior."

Similar to the year-version, the month-version of the PADSS questionnaire showed an acceptable internal consistency and an excellent diagnostic accuracy. Additionally, moderate to strong positive correlations between the total scores of PADSS-month-1 and PADSS-month-2 were found, indicating that the outcome of the questionnaire did not change dramatically over time. We would hesitate to interpret these correlations to infer test–retest reliability because of 2 reasons. First and most importantly, the time between the 2 questionnaires was often longer than 1 month, and therefore the time period under consideration did not overlap. A more precise evaluation of test–retest reliability would require 2 assessments with less time in between—for example, 1 or 2 weeks. Furthermore, the patients already received lifestyle advice during their first visit, which could possibly have led to a lower PADSS score at the second appointment. However, since sleep-related lifestyle advice is widely available and already known to many people, especially patients visiting a tertiary sleep clinic, we do not expect large differences in PADSS scores.

Importantly, the total score of the month-version at the final appointment (after treatment) significantly decreased compared with the total score of the month-version during the first visit. These results imply that the month-version of the PADSS is sensitive to treatment effects. When subgroups stratified for type of treatment were analyzed, we found significant decreases in the PADSS-month score for all subgroups, including participants who only received lifestyle advice. The improvement in the PADSS scale in the latter subgroup is possibly a selection effect, because the choice of treatment was a shared decision between the patient and the caregiver. We could expect that the main reason to not provide additional treatment would be a spontaneous remission of the complaints between the first visit and the start of the treatment. Some of the participants received gabapentin or clonazepam, which currently are off-label treatments for NREM parasomnia. Visually, it seemed that the decrease in the PADSS score was similar in patients receiving medication, compared with patients receiving CBT-p. The effectiveness of medication in people with NREM parasomnia should be further evaluated in a separate randomized controlled study. For this type of research, the month-version of the questionnaire could be a very useful evaluation instrument. We additionally found a statistically significant moderate correlation between the clinical improvement as measured with the CGI scale during the last appointment and the change in the PADSS month-version. Furthermore, the change in the C part (consequences) of the PADSS was correlated with the CGI score, but the B part (frequency) was not. Therefore, we could speculate that physicians are more focused on the consequences of the behavior instead of the frequency of that behavior.

Several limitations can be identified for this study. First, we did not test the diagnostic accuracy of the questionnaires for NREM parasomnia with respect to other sleep disorders that can cause similar complaints, such as REM sleep behavioral disorder and movement disorders. Previous research showed a good discriminative capacity of the year-version of the scale between SW/ST patients and REM sleep behavioral disorder patients.[7] The discriminative capacity between other NREM parasomnia patients and REM sleep behavioral disorder patients and the discriminative capacity of the month version with respect to patients with REM sleep behavioral disorder remain to be evaluated in further research. Second, because of the lack of a gold standard for evaluating treatment results in people with parasomnia, it is difficult to draw a definite

conclusion on the ability of the PADSS month-version to evaluate treatment effect. This requires further research—for example, specifically asking the patients how satisfied they were with their treatment.

In conclusion, we showed that the translated PADSS questionnaire could be successfully used as a screening tool in Dutch patients with SW/ST. We also showed that the PADSS questionnaire can be used to identify NREM parasomnias in people with other types of NREM parasomnia (for example, confusional arousals), people with comorbid sleep disorders, and people who are using medication for reasons other than NREM parasomnia. Finally, we developed a month-version of the PADSS questionnaire that can be used to assess NREM parasomnia complaints over a shorter time span of 1 month. We showed that this month-version had a similar distinguishing capacity as the year-version, that it was reasonably stable over time, and that the total score changed significantly after treatment. These findings suggest that this newly developed month-version can be used to assess short time changes in PADSS, opening new avenues for assessing treatment effects in people with NREM parasomnia.

ABBREVIATIONS

CBT-p, cognitive behavioral therapy for parasomnia
CGI, clinical global impression
IQR, interquartile range
NREM, non–rapid eye movement
PADSS, Paris Arousal Disorders Severity Scale
PSG, polysomnography
REM, rapid eye movement
ROC, receiver operating characteristic
ST, sleep terrors
SW, sleep walking

REFERENCES

1. Howell MJ. Parasomnias: an updated review. *Neurotherapeutics.* 2012;9(4):753–775.
2. Castelnovo A, Lopez R, Proserpio P, Nobili L, Dauvilliers Y. NREM sleep parasomnias as disorders of sleep-state dissociation. *Nat Rev Neurol.* 2018;14(8):470–481.
3. Zadra A, Desautels A, Petit D, Montplaisir J. Somnambulism: clinical aspects and pathophysiological hypotheses. *Lancet Neurol.* 2013;12(3):285–294.
4. Hrozanova M, Morrison I, Riha RL. Adult NREM parasomnias. *Clocks Sleep.* 2018;1(1):87–104.
5. Lundetræ RS, Saxvig IW, Pallesen S, Aurlien H, Lehmann S, Bjorvatn B. Prevalence of parasomnias in patients with obstructive sleep apnea. A registry-based cross-sectional study. *Front Psychol.* 2018;9:1140 .
6. American Academy of Sleep Medicine. *International Classification of Sleep Disorders.* 3rd ed. Darien, IL: American Academy of Sleep Medicine; 2014.
7. Arnulf I, Zhang B, Uguccioni G, et al. A scale for assessing the severity of arousal disorders. *Sleep.* 2014;37(1):127–136.
8. Carrillo-Solano M, Leu-Semenescu S, Golmard J-L, Groos E, Arnulf I. Sleepiness in sleepwalking and sleep terrors: a higher sleep pressure? *Sleep Med.* 2016;26:54–59.
9. Dubessy A-L, Leu-Semenescu S, Attali V, Maranci J-B, Arnulf I. Sexsomnia: a specialized non-REM parasomnia? *Sleep.* 2017;40(2):zsw043.
10. Haridi M, Weyn Banningh S, Clé M, Leu-Semenescu S, Vidailhet M, Arnulf I. Is there a common motor dysregulation in sleepwalking and REM sleep behaviour disorder? *J Sleep Res.* 2017;26(5):614–622.
11. Harris M, Grunstein RR. Treatments for somnambulism in adults: assessing the evidence. *Sleep Med Rev.* 2009;13(4):295–297.
12. Guy, W. *ECDEU Assessment Manual for Psychopharmacology*, Revised. Publisher: US Dept. of Health, Education and Welfare. Rockville, MD, 1976.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication August 20, 2021
Submitted in final revised form December 6, 2021
Accepted for publication December 6, 2021

Address correspondence to: Prof. S. Overeem, MD, PhD, Sleep Medicine Center Kempenhaeghe; Email: OvereemS@kempenhaeghe.nl

DISCLOSURE STATEMENT

All authors have seen and approved the manuscript. S. Overeem consulted for Bioprojet, Jazz Pharmaceuticals, UCB Pharma, and Takeda, all paid to the institution, and not related to the present work. The other authors report no conflicts of interest.