Simple behavioral criteria for the diagnosis of disorders of arousal

Ariana Barros, Ginevra Uguccioni, Victoire Salkin-Goux, Smaranda Leu-Semenescu, Pauline Dodet, Isabelle Arnulf

▶ **To cite this version:**

HAL Id: hal-02946497

https://hal.sorbonne-universite.fr/hal-02946497

Submitted on 23 Sep 2020

Simple Behavioral Criteria for the Diagnosis of Disorders of Arousal

Ariana Barros, MD,[1,2] Ginevra Uguccioni, PhD,[1] Victoire Salkin-Goux, MS,[1] Smaranda Leu-Semenescu, MD, [1,3] Pauline Dodet, MD, [1,3] Isabelle Arnulf MD, PhD[1,3]

1 AP-HP, Pitié -Salpêtrière Hospital, Sleep Disorders Unit, Paris, France*

2 Centro Hospitalar de Lisboa Central, Lisboa, Portugal

3 IHU@ICM, Sorbonne University, Paris, France

* Institution where the work was performed.

Article submitted to: JCSM **; Version:** #2

Correspondance to: Isabelle Arnulf, Service des Pathologies du Sommeil, Hôpital Pitié-Salpêtrière, 47-83 Boulevard de l'Hôpital, 75651 Paris Cedex 13, France

Email: isabelle.arnulf@aphp.fr; phone: 33 1 42 16 77 02; fax: 33 1 42 16 77 00.

Statements

- All authors have seen and approved the manuscript

- None of the authors has any conflict of interest relative to this study

- This study is a surrogate study of NCT02648568 and NCT03074578 on clinicaltrial.gov

- Number of tables: 4 ; Number of figures: 2

- Abstract count words: 234; Manuscript count words: 4253

ABSTRACT

Study Objectives: This case-control study aimed to identify and validate behavioral markers supporting the diagnosis of disorders of arousal (DOA) with video-polysomnography.

Methods: All behaviors associated with 1,335 episodes of N3 interruptions were compared in 52 adult patients with DOA *vs.* 52 participants without DOA (healthy controls and patients with insomnia, hypersomnia or sleep apnea syndrome).

Results: Patients with DOA had more frequent (5.1 ± 2.4 *vs.* 3.4 ± 1.9 interruptions/N3 time) and longer (35.8 ± 33 *vs.* 23.1 ± 21.4 sec) arousals and awakenings from N3 than controls. In the DOA group, the onset of behaviors was more abrupt, and behaviors including eye opening (69 *vs.* 16%), head raising (41 *vs.* 9%), visually exploring the environment (27 *vs.* 1%), expression of fear/surprise (21 *vs.* 0%), speaking (18 *vs.* 0.3%), trunk raising (13 *vs.* 0.3%), and interacting with the environment (13 *vs.* 0.5%), were (unlike quiet, comfort behaviors) more frequent than in controls. A cut-off of 2 or more N3 interruptions containing eye opening yielded a sensitivity of 94.2% and a specificity of 76.9% for a DOA diagnosis. This accuracy was confirmed in a second set of data (second night of monitoring). Behaviors including an expression of fear/surprise (67.3%), sitting (32.7%), screaming, and standing up were specific to patients with DOA.

Conclusion: A simple, behavioral video marker of behavioral reactions during N3 interruption (i.e., opening the eyes at least two times in the same night) was sensitive, specific and reproducible for discriminating patients with DOA from sleep laboratory controls.

Keywords: Disorders of arousal; sleepwalking; sleep terror; diagnosis; video-polysomnography, behaviors, eye opening

Brief summary: No behavioral criterion supporting the diagnosis of sleepwalking and/or sleep terrors has been validated with video-polysomnography. Consequently, we compared the behaviors upon awakening/arousing from N3 sleep in 52 adults with sleepwalking with those in 52 adults with other sleep disorders or healthy participants. A cut-off of two or more interruptions of N3 sleep with eye opening accurately discriminated adults with sleepwalking from other patients and healthy participants, providing a simple and sensitive criterion to support diagnosis in clinical practice. Only sleepwalkers displayed expressions of fear/surprise, sat, screamed, and stood up during N3 interruptions, suggesting that these specific criteria could be used for forensic purposes.

INTRODUCTION

Disorders of arousal (DOA, including sleepwalking, sleep terrors and confusional arousals) are characterized by inappropriate, mostly amnestic behaviors occurring during partial arousals from N3 sleep, reflecting various degrees of motor, emotional and autonomic activations. The disorders expose people to risks of trauma, violence and uncontrolled behaviors that may be life threatening. In addition, patients with DOA have more difficulties initiating and maintaining sleep, a lower quality of life, higher levels of anxiety, and more daytime fatigue and sleepiness than healthy controls.[1,2,3] Disorders of arousal are common in children but may persist (and sometimes appear) in up to 2-4% of adults.[4]

The diagnosis is mostly based on interviews of the patients and their family, as indicated in the international classification of sleep disorders.[5] In contrast to many other sleep disorders (e.g., sleep apnea syndrome, REM sleep behavior disorder, periodic leg movements disorder, and central hypersomnias), in which sleep and sometimes video monitoring are mandatory to reach the diagnosis, the place of video-polysomnography in the diagnosis process of DOA is debated and not required based on the present criteria.[5] Ambulation and complex behaviors during N3 arousals are rare in the sleep laboratory, and some authors suggest that "normal" polysomnographies are sometimes observed in participants with typical symptoms of DOA. However, a sleep study may document simple behaviors typical of confusional arousals, rule out differential diagnoses (REM sleep behavior disorder, epilepsy, sleep-related motor disorders) and identify potential triggers of parasomniac episodes, such as noises, sleep-related breathing disorders or periodic leg movements. In addition, there is an unmet need (for clinical and forensic purposes) to determine the sensitivity and specificity of the signs observed in a sleep study. Some EEG markers (hypersynchronous delta sleep activities

and spectral analysis) lack specificity, but recently, an index of arousals plus awakenings (greater than 6.8 per N3 time) has proven sensitive (79%) and specific (82%) for distinguishing patients referred for DOA from healthy participants.[6] A slow/mixed arousal index (greater than 2.5 per N3 time) was even more sensitive (94%) and totally specific (100%). However, this analysis requires highly trained sleep scorers and 8 EEG channels.

We aimed to identify easy video-based criteria to distinguish patients with DOA from healthy controls and from patients with other sleep disorders. For this purpose, we analyzed the behaviors during N3 arousals/awakenings in patients with DOA vs. those in controls (healthy controls and patients with insomnia, hypersomnia and sleep apnea syndrome) to find simple, sensitive and specific markers of DOA.

METHODS

Study design and population

This is a retrospective, case–control study. The participants were selected from the patient and control populations of the sleep disorder center at the Pitié-Salpêtrière University Hospital. The patients with DOA were recruited between January 2016 and July 2018 and diagnosed with sleepwalking or sleep terror by ~~one of~~ the sleep physicians using the international criteria.[5] They met criteria for disorders of arousal, including (i) recurrent episodes of incomplete awakening from sleep, (ii) inappropriate or absent responsiveness to efforts of others to intervene or redirect the person during the episode, (iii) limited or no associated cognition or dream imagery, (iv) partial or complete amnesia for the episode and (v) no better explanation for the disturbance, such as another sleep disorder, a mental disorder, a medical condition, or medication or substance use. In addition to the above criteria, patients were classified as suffering

from sleepwalking if there was (i) a history of ambulation during sleep or (ii) the persistence of sleep or impaired judgment during ambulation. In addition to the above criteria, patients were classified as having sleep terror if there was (i) a history of sudden episodes of terror occurring during sleep, usually initiated by a cry or loud scream with sympathetic and behavioral manifestations of intense fear, and (ii) difficulty in arousing the person, mental confusion when awakened from an episode, complete or partial amnesia from the episode, or dangerous or potentially dangerous behaviors. Because the comparison of behaviors arising during a confusional arousal in N3 and during a sleep-associated (whatever the sleep stage) epileptic seizure has already been studied,[9] the patients with positive clinical history of epilepsy were excluded in both groups (DOA and control).

The control group comprised healthy participants and participants with sleep disorders but no DOA. The healthy participants had been recruited by word of mouth for another trial (COREV) that included a video and sleep study and had no sleep complaints, especially present parasomnia, as checked by an interview with the physician. The control population with sleep disorders had to be representative of the source population from which the DOA cases were extracted. For this purpose, it consisted of: i) all consecutive participants who were referred to our sleep unit from May 2018 to July 2018; ii) who had no DOA; iii) who had no positive clinical history of epilepsy; iv) who did not take any psychotropic treatment for at least several months; iv) who were matched for age and sex with the DOA patient; and v) who had at least one awakening episode during N3 sleep (in order to be able to observe at least one behavior associated with one N3 interruption). This control group was matched (as a group and not 1 to 1) for age and sex with the DOA group.

Ethics

The patients with DOA and the healthy controls gave their written consent to participate in the study, which was approved by the ethics committee Ile de France-06 as part of a larger study on hypnosis in sleepwalking (NCT02648568 on clinicaltrial.gov) and on emotions (NCT03074578 on clinicaltrial.gov). Following the French regulation on clinical research, the other part of the group (control patients) received before, during and after their stay the oral and in writing information that their deidentified measures could be later reused by investigators for research purposes (unless they refused this use) and did not oppose to that. The center had received the agreement MR003 (declaration 1999732) from the National Commission for Rights and Freedom related to Informatics (CNIL) since October 2016 for storing and reusing for research purpose deidentified clinical, polysomnographical and video/audio measures performed in clinical routine in the sleep laboratory.

Video-polysomnography

All participants completed at least one night with video-polysomnography at the sleep unit, but the 52 patients with DOA and 16 healthy controls underwent two consecutive night studies. All rooms were equipped with a single infrared camera (synchronized with the audio recording and the polysomnography), placed in front of the bed. Lights were switched off by the patients by 11 PM at the latest and remained off until they were turned on by staff the next morning. The recordings included at least three channels for electroencephalography (EEG; Fp1/C3, C3/A2 and C3/O1; extensive 8-electroencephalophic channel montage in patients with DOA), two channels for electrooculography (inferior left epicanthus/A2 and superior right epicanthus/A2) and three channels for surface electromyography (chin [mentalis] muscle, and right and left

tibialis anterior muscles) as well as an electrocardiogram, thoracic and abdominal belts to measure the respiratory efforts via inductive plethysmography, a measure of naso-oral airflow through nasal pressure and oral thermistors, and a measure of transcutaneous oxygen saturation. The sleep stages, arousals, apnea and hypopnea events, and periodic leg movements were scored by experienced sleep neurologists according to international criteria.[7]

Analysis of behaviors

All interruptions of stage N3 sleep were collected. They included classical awakenings (alpha or higher EEG background rhythm lasting longer than 15 s), as well as arousals (lasting 3 to 15 s) with an acceleration of the background EEG frequency, which was mostly in the alpha rhythm but also (especially in patients with DOA) in a theta rhythm or a mixed theta-alpha rhythm. Additionally, we included the "slow-wave arousals" characterized by sudden events with persistent delta waves (0.5-3 Hz) on central derivations associated with increased muscle activity and lasting more than 3 s, most of which had a high amplitude (named hypersynchronous delta activity[8]). The behaviors concomitant with these N3 arousals/awakenings were observed on the video/audio, and the following behaviors were scored by two scorers: (1) eye opening, (2) staring or exploring the environment, (3) head raising, (4) trunk raising, (5) sitting, (6) standing up/walking, (7) speaking, (8) screaming, (9) showing any facial expression of fear or surprise, (10) interacting with the environment and (11) making usual comfort arousal movements (i.e., rubbing, yawning, repositioning oneself). The additional features collected for each episode included the triggering event (present/absent; the nature of the trigger), the onset (abrupt or slow) and duration of the episode (defined from onset to the end of any motor activity). These categories were chosen according to local

experience of their high frequency in DOA and to previous studies of behaviors in patients with DOA vs. epilepsy[9] or controls[10]. The scorer was not blinded to the DOA/non DOA diagnosis of each participant.

Statistical analysis

Descriptive statistics were used to describe continuous variables, and qualitative measures were presented as frequencies and percentages. The comparisons between patients were performed using Student's t-test (quantitative measures) and chi-squared test or Fisher's exact test (qualitative measures). The levels of agreement between raters were tested in 20 randomly selected, representative cases using Cohen's kappa for nominal (qualitative) variables and an interrater correlation coefficient for quantitative variables. Values of Cohen's kappa and interrater correlations greater than 0.8 indicate excellent agreement (almost perfect agreement). The association between quantitative measures was tested using Pearson correlation coefficient or Spearman correlation coefficient (in cases where the normality assumption was not verified). All tests were two-sided, and a *p*-value below 0.05 was considered statistically significant. Receiver operating characteristic (ROC) analysis was applied to evaluate the best cut-off value for diagnostic performance of the criteria to provide the best pair of values for sensitivity and specificity. To compare the diagnostic accuracy of the results, the area under the curve (AUC) with a 95% confidence interval (CI) was calculated. Likelihood ratios (LRs) for positive test results and accuracy (number of correct results divided by the total number of tests) were provided as a measure of diagnostic accuracy. The best possible threshold was defined as the highest Youden index [(specificity + sensibility) − 1]. Diagnostic criteria based on cut-offs identified by the ROC analysis were evaluated for each test individually and for each test combined with the number of N3

interruptions/awakenings (considered positive for a participant who was positive on each test). The association between test results and the presence of the disease was evaluated by Pearson's test. The Rasch model was conducted with behavioral features of N3 interruptions to examine the model fit statistics, namely, the information-weighted goodness-of-fit statistics (INFIT), and the ordering of the behavior categories. All analyses were performed using IBM SPSS Statistics for Macintosh (Version 23.0. Armonk, NY).

RESULTS

Clinical and Sleep Characteristics of the Participants

Fifty-two patients with DOA and 52 controls (23 healthy participants, 10 patients with sleep apnea syndrome [range of apnea/hypopnea index, 15.2-90.1/hour], 11 patients with idiopathic hypersomnia and 8 patients with insomnia, as randomly determined) were enrolled in the study. As expected by matching, age and gender were not different between the DOA and control group (Table 1). In the DOA group, most patients (N = 40, 76.9%) had combined sleepwalking and sleep terrors, 7 (13.5%) had isolated sleepwalking and 5 (9.6%) had isolated sleep terrors. Associated confusional arousal was diagnosed in 12 of these patients (23.1%). Sleep talking was reported by 49 (94.2%) patients. The mean age at DOA onset was 10.8 ± 5.4 years (mean ± SD), and the mean disease duration was 19.9 ± 10.6 y. Twenty-nine (55.8%) patients reported having had violent or dangerous sleep-related behaviors at home. A positive family history of DOA was reported in 69.2% of patients. The total score on the Paris Arousal Disorders Severity Scale[11] was 16.9 ± 5.0, the behavior subscore A was 8.0 ± 3.8, the subscore B (frequency of the behaviors) was 4.2 ± 1.1 and the subscore C (impact of the DOA) was 4.7 ± 1.4. Neurological and psychiatric comorbidities were equally rare in

the DOA (0 and 1.9%, respectively) and control (1.9 and 3.8%, respectively) groups. These comorbidities included migraine, anxiety and depression. The EEG tracings during relaxed wakefulness and sleep stages did not disclose any paroxysmal activity in patients with DOA or in controls. Compared with controls, patients with DOA had longer total sleep time, higher sleep efficiency, higher REM sleep percentage, and a lower apnea/hypopnea index, as well as higher stage shifts and awakening indices (Table 1). Patients with DOA had, on average, a higher number and index of N3 interruptions, awakenings, and arousals than controls (Table 1). The between-scorer agreement for the quantification of N3 interruptions was 0.99 (95% CI = 0.966–0.995).

Behavioral Characteristics of N3 Interruptions

The interruptions from N3 sleep were analyzed in 172 night recordings and totaled 1,335 interruptions (953 in DOA patients and 382 in controls). The N3 interruptions lasted longer in the DOA than in the control group (35.8 ± 33 *vs.* 23.1 ± 21.4 sec, $P < 0.0001$). As many as 913 (95.8%) N3 interruptions in the DOA group and 365 (95.5%) in the control group were accompanied by behaviors. In 6% of the cases, the behaviors "eye opening" "exploring the environment" and "expression of fear/surprise" could not be defined from the video images, because the patients were sleeping in the prone position. An abrupt onset of motor behavior, defined as a behavior starting at the very moment of the sleep interruption, was observed in 622 (65.3%) interruptions in the DOA group and in none of the control group ($P < 0.0001$), whereas the motor behavior started several seconds after the N3 interruption in the controls. In all groups, 170 (12.7%) episodes of N3 interruptions were triggered by a precipitating event (with a similar frequency between groups), including ambient noises (65.7%), limb movement or myoclonus (14.5%), nurse intervention (7.6%), deep inspiration/cough (4.7%),

snoring (4.1%) and apnea/hypopnea/flow limitation (3.5%). Patients with DOA opened their eyes during 653 (68.9%) episodes of N3 interruption, whereas controls opened their eyes during 59 (15.5%) episodes (Figure 1). The N3 interruptions triggered by internal stimuli (apnea, hypopnea, cough, leg movement/myoclonus) led to less frequent (14/48; 29.2%) eye openings than those triggered by external stimuli (noises, contacts), which led to eye openings in 77/124 interruptions (62.1%, $P < 0.0001$). The patients with DOA reacted more to external triggers by opening the eyes (64/90, 71.1%) than the controls (13/34, 38.2%, $P = 0.0008$) did. All types of behaviors described in Figure 1, which accompanied the N3 interruptions, were more frequent in the DOA than in the control group ($P < 0.0001$ for each difference), including in decreasing order of frequency: head raising, visually exploring the environment, speaking, trunk raising and interacting with the environment. Some behaviors were exclusively observed in the DOA group, including facial expressions of fear or surprise (20.7%), sitting (8.6%), screaming (3.4%) and standing up (0.3%). In contrast, usual, quiet movements (repositioning, yawning, stretching, rubbing nose or some other body part) were more frequent in the control group (N = 355, 92.9% of N3 interruptions) than in the DOA group (N = 790, 81.9%, $P < 0.0001$). Notably, the rapid arousing motor pattern (eye opening and head and trunk raising) described in Figure 1, when present, preceded the quiet usual movements in the DOA group. According to the Rasch Model, there was a hierarchical structure of behaviors predicting the "propensity" for DOA during the N3 interruptions: eye opening was the "easiest" criteria (lowest propensity), and standing up/walking was the "hardest" (highest propensity), in the same order as the frequency of these behaviors (Supplemental Table). The factor analysis (INFIT values between 0.7 - 1.4) of the behaviors supports their unidimensionality.

Quantification of Behaviors

Although these N3 interruptions were similarly spontaneous or triggered by an event in the two groups, the frequency of the corresponding behaviors was markedly different between groups (Table 2), including a higher number and index of eye opening, visually exploring the environment, head raising, trunk raising, sitting, speaking, screaming, expressing fear/surprise, interacting with the environment and engaging in usual arousal movements in the DOA group. The interrater agreement for identifying the episodes of eye openings was 97.7% ($\kappa = 0.94$, 95% CI = 0.84 – 0.96). The other κ coefficients for individual behavior classification were close to 0.9 or above for all other components. The number of eye openings positively correlated with the number of other behaviors, except for the number of screaming and sitting behaviors.

Cut-off Identification

The results of the ROC analysis for markers are shown in Table 3 and Figure 2. The highest AUC (measure of discriminative ability of the test) was obtained for the number of interruptions with eye opening (AUC = 0.93 [95% CI, 0.89-0.98]), the number of N3 awakenings (0.86 [95% CI, 0.78-0.93]), and the number of N3 interruptions (AUC = 0.77 [95% CI, 0.685-0.864]). The best pair of values for sensitivity and specificity was found at a cut-off value of 6 or more N3 interruptions per night (LR+ = 2.27, accuracy = 77.1%) and 6 or more N3 awakenings per night (LR+ = 3.62, accuracy = 77.9%). For the number of episodes with eye opening, the optimal cut-off point was 2 per night (LR+ = 4.08, accuracy = 85.6%). Of the behaviors that were 100% specific to the DOA group during the first night, 35/52 (67.3%) of patients displayed a facial expression of

fear/surprise at least once, 17/52 (32.7%) sat in the bed, 10 (19.2%) screamed, and 2/52 (3.8%) stood up and walked.

Cut-off Application and Test Combination

With the application of the identified cut-off values (Table 4), two or more episodes of N3 interruptions with eye opening yielded a sensitivity of 94.2% and specificity of 76.9%. When this number of eye openings was combined with at least six N3 interruptions (whether arousal and/or awakenings), the sensitivity was 75%, and the specificity was 88.5%. When the number of eye openings was associated with at least six N3 awakenings, the sensitivity was 67.3%, but the specificity reached 94.2%. To retest the cut-offs in a second set of data, their accuracy was tested in the second night (52 nights of DOA patients and 16 nights of controls). The cut-off of two N3 interruptions with eye opening yielded a sensitivity of 94.2% and a specificity of 85.7%, which was better than during the first night.

DISCUSSION

In this controlled study, patients with DOA had more frequent and longer arousals and awakenings from N3 than controls had. Notably, the behaviors observed during the 1335 episodes of N3 interruptions differed between groups. In the DOA group, the onset of behaviors was more abrupt (starting at the exact time of N3 interruption), and eye opening, head raising, visually exploring the environment, speaking, trunk raising, and interacting with the environment (in descending order of frequency) were more frequent than in controls, whereas the frequency of quiet behaviors usually associated with comfort did not differ between groups. Using receiver operator curves, a cut-off of two or more N3 interruptions containing eye opening yielded a sensitivity of 94.2% and

a specificity of 76.9% for a diagnosis of DOA. A similarly excellent sensitivity (94.2%) and specificity (88.7%) was confirmed in a second set of data. Combined markers (i.e., six or more N3 awakenings in the whole night, combined with two or more eye openings during these interruptions) yielded a higher specificity (94%).

Motor phenomena during sleep have been the topic of intensive studies since the introduction of video-polysomnography. The pattern of behaviors associated with N3 interruptions is routinely observed in sleep laboratories for more than 50 years, but it has been more rarely quantified and compared to that in controls. It has been previously studied in 30 adult sleepwalkers and 10 controls.[10] The behaviors had an abrupt onset in 55% of the episodes, which is close to the 65.3% found in our series.

Eye opening is the most frequent behavior associated with N3 interruptions in patients with DOA, as it was observed in 84% of 57 episodes in an Australian series,[9] in 76% of 184 episodes in an Italian series,[10] and in 68.9% of 953 episodes in our series. Of interest, eye opening also occurred during N3 interruptions in controls (15%) here and in one of the eight head flexion/extension movements in the Italian series,[10] but opening the eyes more than one time per night during N3 interruption was exceptional, to the point of providing a marker with a high sensitivity and specificity for a DOA diagnosis. Additionally, this marker was reproduced with a higher specificity/sensitivity in a second set of data, was simple to recognize and had excellent interrater agreement. Because our control group included not only healthy controls but also patients with sleep apnea, hypersomnia and insomnia, this marker could be a useful indicator (due to its excellent sensitivity) for the diagnosis of DOA in a sleep clinic context. However, when nocturnal behaviors were compared between patients with DOA and patients with

nocturnal frontal lobe epilepsy in an Australian series, episodes with eye opening did not discriminate the groups, as 97% of frontal seizures was associated with an opening of the eyes.[9] Therefore, this criterion can be used for discriminating patients with DOA from usual sleep clinic patients and form healthy controls, but not from patients with sleep-related hypermotor epilepsy.

The specificity of this "two episodes of eye openings" marker did not reach 100% in our group and therefore cannot be used for forensic purposes. Patients with DOA more frequently opened the eyes (72%) in reaction to an external stimulus (noise, contact) triggering the N3 interruption than did controls (38%), suggesting that DOA patients are more reactive (at least if one considers eye opening as a motor, orientating and eventually fear behavior) than are controls. This observation of a higher reactivity to external stimuli in patients with DOA was obtained in a clinical setting. Similarly, it has been previously shown that forced arousals provoked by strong auditory stimuli trigger abnormal confusional reactions (clumsy movements, sleep talking) in patients with DOA but not in healthy controls.[12]

The pattern of behaviors accompanying N3 interruption often extended beyond the simple, immediate opening of the eyes. In patients with DOA (and much more exceptionally in controls), the behavioral pattern was more typical of a general arousal/orientation behavior, which includes the behavior of raising the head and trunk, as well as staring, looking around, speaking and touching things around them. Because these events follow a hierarchical order (i.e., the probability that the eye is open when the head or trunk is raised is high), as shown here by the Rasch model, these behaviors may result from the activation of a central pattern generator of arousal behavior. When

present, this abrupt arousing motor pattern strongly contrasts with the quiet movements associated with comfort (stretching, rubbing, repositioning oneself) performed in both groups but appear after the abrupt initial arousal pattern in the DOA group.

A facial expression of fear or surprise was observed in 20% of N3 interruptions, further reinforcing the feeling of an explosive, surprised behavior, which seems specific to patients with DOA, as expressions of fear/surprise were not observed in the control group. A similar pattern of arousal behaviors has been described at the onset of nocturnal frontal lobe seizures, making paroxysmal arousals indistinguishable from confusional arousals.[9] However, we show here that these arousal behaviors are markedly distinguishable from the usual, quiet behaviors associated with arousals and awakenings from N3 in normal participants and in patients with non parasomniac sleep disorders (without epilepsy).

For forensic purposes, one may look for highly specific rather than highly sensitive cut-offs, to be able to determine with a high level of certainty that a participant has a DOA. In this case, several behaviors were exclusively observed in patients with DOA here (100% specific) and never observed in patients with other sleep disorders. These behaviors included showing a facial expression of fear/surprise upon N3 interruption (displayed by more than two-thirds of patients with DOA), sitting (one-third of patients), and screaming (19%). As often reported, ambulation is exceptional in the sleep laboratory, as here only 2/52 patients stood up and walked during the first night of recordings.

A recent study proposed the first video-polysomnography-based criteria for the diagnosis of DOA in adults, based on quantification of the N3 interruptions (N3 fragmentation index) and the postarousal EEG pattern (fast and slow/mixed arousal indices).[6] In this study, the N3 fragmentation index cut-off value of 6.8/h reached a sensitivity of 79% and a specificity of 82%. In our study, this cut-off of 6.8/h was too high, as it resulted in a lower sensitivity (19.2%), whereas the "best" cut-off of 3.6/h had a sensitivity of 69.2% and specificity of 59.6%. Eventually, a number of 6 or more N3 interruptions (a number rather than an index, which is possibly easier to collect) yielded a high level of accuracy in our series. Sample fluctuations (the sample size of Lopez and colleagues was larger than ours)[6] and differences in inclusion criteria (our controls were required to have at least one N3 interruption to describe the associated behaviors, whereas this was not an inclusion criteria in the other sample) may explain these different results. Notably, the duration of N3 interruptions was longer in patients with DOA than in controls, suggesting that the reactivity during N3 interruption is higher not only in terms of the complexity of the behavioral response in these patients but also in the time it takes to resume sleeping.

Our study presents some methodological limitations, such as a reasonable but limited number of participants in each group. However, more than 1300 episodes of N3 interruption were carefully examined on video, which constitutes one of the largest behavioral studies in this field. The eyes are not visible when the participants sleep in the front position, which sometimes (here in 6% of N3 interruptions) limits the ability to use the two eye opening marker. In these instances, the surrogate markers can be the number of awakenings from N3. Patients with nocturnal epilepsy were not included in this study. They are known to open the eyes when starting a seizure at night. However,

seizures mainly take place in N1 and N2 stages, and rarely in N3,[8] which may help to distinguish DOA from seizures more easily than behavioral criteria. Eventually, the scorer was not blinded to the diagnosis of DOA in participants when analyzing the videoclips.

In conclusion, we found that a simple, video-polysomnography-based marker of behavior reaction during N3 interruption (i.e., eye opening at least two times in the same night) was sensitive, specific and reproducible in the discrimination of patients with DOA from sleep laboratory controls (more than half of the control group were patients with sleep disorders), but not from sleep-related hypermotor epilepsy. This cut-off value allowed the correct classification of 85.6% of patients. Other simple markers (6 or more N3 awakenings or N3 interruptions per night) isolated or in combination with eye openings were also accurate. These markers could be used in clinical settings to support the DOA diagnosis. Eventually, this behavioral work could be used in the future to develop methods based on automatic video analysis (even at home) to classify patients.

ABBREVIATIONS LIST: AUC: area under the curve; CI: confidence interval; DOA: disorders of arousal; INFIT: information-weighted goodness-of-fit statistics; LR: likelihood ratio; REM sleep: rapid eye movement sleep; ROC: receiver operating characteristic; SE: standard error.

REFERENCES

1.   Lopez R, Jaussent I, Scholz S, Bayard S, Montplaisir J, Dauvilliers Y.
       Functional impairment in adult sleepwalkers: a case-control study. *Sleep.*
       2013;36(3):345-351.

2.   Carrillo-Solano M, Leu-Semenescu S, Golmard JL, Groos E, Arnulf I.
       Sleepiness in sleepwalking and sleep terrors: a higher sleep pressure? *Sleep
       Med.* 2016;26:54-59.

3.   Oudiette D, Leu S, Pottier M, Buzare MA, Brion A, Arnulf I. Dreamlike
       mentations during sleepwalking and sleep terrors in adults. *Sleep.*
       2009;32(12):1621-1627.

4.   Ohayon MM, Mahowald MW, Dauvilliers Y, Krystal AD, Leger D. Prevalence
       and comorbidity of nocturnal wandering in the U.S. adult general population.
       Neurology. 2012;78(20):1583-1589.

5.   Ito E, Inoue Y. [The International Classification of Sleep Disorders, third
       edition. American Academy of Sleep Medicine. Includes bibliographies and
       index]. *Nihon Rinsho.* 2015;73(6):916-923.

6.   Lopez R, Shen Y, Chenini S, et al. Diagnostic criteria for disorders of arousal: A
       video-polysomnographic assessment. *Ann Neurol.* 2018;83(2):341-351.

7.   Iber C, Ancoli-Israel S, Chesson A, Quan S. *The AASM Manual for the Scoring
       of Sleep and Associated Events: Rules, Terminology and Technical*

Specifications, 1rst Ed. Westchester, IL: American Academy of Sleep Medecine; 2007.

8. Guilleminault C. Hypersynchronous slow delta, cyclic alternating pattern and sleepwalking. *Sleep.* 2006;29(1):14-15.

9. Derry CP, Harvey AS, Walker MC, Duncan JS, Berkovic SF. NREM arousal parasomnias and their distinction from nocturnal frontal lobe epilepsy: a video EEG analysis. *Sleep.* 2009;32(12):1637-1644.

10. Loddo G, Sessagesimi E, Mignani F, et al. Specific motor patterns of arousal disorders in adults: a video-polysomnographic analysis of 184 episodes. *Sleep Med.* 2018;41:102-109.

11. Arnulf I, Zhang B, Uguccioni G, et al. A scale for assessing the severity of arousal disorders. *Sleep.* 2014;37(1):127-136.

12. Pilon M, Montplaisir J, Zadra A. Precipitating factors of somnambulism: impact of sleep deprivation and forced arousals. *Neurology.* 2008;70(24):2284-2290.

FIGURE CAPTIONS

Figure 1 – Type and Frequency of Behavior Occurring during N3 Interruptions (N = 1,335) in

Patients with Disorders of Arousal (plain bars) and Controls (empty bars).

Figure 2 – Receiver Operating Characteristics Analyses for Various Markers During N3

Interruptions.

Table 1 – Sleep measures in patients with disorders of arousal vs. controls (healthy participants and participants with sleep disorders but no DOA)

	DOA	Controls	*p*-value
Number of participants	52	52	-
Female gender (N, %)	23 (44.2%)	29 (55.8%)	0.62
Age, years	30.4 ± 8.2	32.2 ± 10.4	0.33
Sleep measures (Night 1)			
Total sleep time, min	491.6 ± 98.2	428.7 ± 69.9	0.000
Sleep efficiency, %	88.2 ± 7.1	81.1 ± 10.4	0.000
Wakefulness after sleep onset, min	79.5 ± 45.7	66.9 ± 46.6	0.17
Latency to, min			
Sleep onset	21.9 ± 35.9	23.1 ± 28.0	0.85
N3	17.9 ± 14.9	25.7 ± 32.5	0.12
REM	112.2 ± 50.5	139.2 ± 96.4	0.08
Sleep stages, % of total sleep time			
N1	3.1 ± 2.9	4.4 ± 4.9	0.10
N2	52.5 ± 6.8	55.1 ± 10.9	0.15
N3	21.9 ± 6.7	23.1 ± 9.2	0.45
REM	22.1 ± 5.4	18.1 ± 6.7	0.001
Sleep fragmentation, number of events/hour			
Stage shifts	14.9 ± 4.6	12.4 ± 3.8	0.003
Arousals	7.5 ± 3.6	9.7 ± 9.6	0.12
Awakenings	4.1 ± 1.4	3.3 ± 1.4	0.005
Apnea-hypopnea	1.9 ± 3.4	6.9 ± 14.6	0.017
Periodic limb movements	5.6 ± 7.2	5.9 ± 11.0	0.87

Fragmentation of N3 sleep

Total interruptions, n/night	8.8 ± 4.4	5.1 ± 2.8	0.000
Total interruption index, n/hour of N3 sleep	5.1 ± 2.4	3.4 ± 1.9	0.003
Awakening, n/night	7.0 ± 3.4	3.0 ± 1.9	0.000
Awakening index, n/hour of N3 sleep	4.1 ± 2.1	2.0 ± 1.4	0.000

Measures are mean ± SD or %. DOA: Disorders of arousal.

Table 2 – Behavioral Characteristics during N3 Interruptions (Night 1) in Patients with Disorders of Arousal vs. Controls

Participants	DOA	Controls	*p*-value
Number of patients	52	52	-
Eye opening, n/night	5.8 ± 3.6	0.8 ± 1.3	0.000
Eye opening index, n/hour of N3 sleep	3.4 ± 2.2	0.5 ± 1.5	0.000
Head raising, n/night	3.1 ±2.9	0.3 ± 0.7	0.000
Head raising index, n/hour of N3 sleep	1.8 ± 1.5	0.2 ± 0.5	0.000
Exploring the environment, n/night	2.2 ± 2.0	0.0 ± 0.0	0.000
Showing an expression of fear/surprise	1.6 ± 1.7	0.0 ± 0.0	0.000
Speaking, n/night	1.3 ± 1.8	0.0 ± 0.0	0.000
Trunk raising, n/night	1.1 ± 1.4	0.0 ± 0.0	0.000
Interacting with the environment	1.1 ± 1.8	0.4 ± 0.1	0.000
Sitting, n/night	0.6 ± 1.1	0.0 ± 0.0	0.000
Screaming, n/night	0.3 ± 1.3	0.0 ± 0.0	0.000
Standing up/walking, n/night	0.02 ± 0.1	0.0 ± 0.0	0.32
Comfort movements, n/night	6.9 ± 4.2	4.7 ± 2.8	0.002
Comfort movement index, n/hour of N3 sleep	3.9 ± 2.2	3.0 ± 1.6	0.03
Triggered events, n/night	1.1 ± 1.6	1.0 ± 1.1	0.62
Triggered event index, n/hour of N3 sleep	0.6 ± 0.9	0.9 ± 1.2	0.33

DOA: Disorders of arousal.

Table 3 - Optimal Cut-Off Points of N3 Interruptions for a Diagnosis of Disorders of Arousal, Evaluated with Receiver Operating Characteristics Curves.

Test Cut-off	Cut-off point	Sensitivity (%)	Specificity (%)	Youden index	Cut-off point for 95% sensitivity	Cut-off point for 95% specificity
No. of N3 interruptions	≥ 6	75.0	67.3	0.423	2.5	11.5
N3 interruption index	≥ 3.6	69.2	59.6	0.288	1.8	6.2
No. of N3 awakenings	≥ 6	67.3	92.3	0.596	2.5	6.5
No. of eye-opening N3 interruptions	≥ 2	94.2	76.9	0.711	0.5	4.5

Table 4 - Diagnostic Accuracy of Behavioral and EEG markers for Disorders of Arousal

Test Cut-off	Sensibility (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)	Positive Likelihood Ratio	Accuracy (%)	χ^2	p-value
Eye openings, N = 2/night	94.2	76.9	80.3	93.2	3.26	85.6	51.4	<0.000
Eye openings, N = 2 + N3 Interruptions, N ≥ 6/night	75.0	88.5	86.7	78.0	6.52	81.7	40.1	<0.000
Eye openings, N = 2 + N3 Awakenings, N ≥ 6/night	67.3	94.2	92.1	74.2	11.6	80.8	39.8	<0.000
N3 interruption index >6.8	19.2	96.2	83.3	54.3	5.05	57.7	4.62	0.028

Supplemental Table - Behaviors Associated with N3 Interruptions, Hierarchically Arranged According to the Rasch model.

Behavior	Measure (SE) (Logit*)	Mean Square	
		INFIT	t
Standing-up/walking	3.600 (0.275)	1.01	0.21
Screaming	2.472 (0.109)	1.20	1.34
Sitting	1.948 (0.080)	0.90	-1.20
Interacting with the environment	1.704 (0.087)	1.07	1.04
Trunk raising	1.687 (0.068)	0.89	-1.58
Speaking	1.431 (0.061)	1.17	2.80
Showing an expression of fear/surprise	1.339 (0.058)	0.97	-0.56
Exploring the environment	1.118 (0.054)	1.06	1.29
Head raising	0.628 (0.046)	0.92	-1.85
Eye opening	-0.108 (0.44)	0.90	-2.60

SE: standard error; INFIT: information-weighted goodness-of-fit statistics;

*Linear Rasch (rating scale model) - calibrated item logit.