The Arousal Disorders Questionnaire: a new and effective screening tool for confusional arousals, Sleepwalking and Sleep Terrors in epilepsy and sleep disorders units

The Arousal Disorders Questionnaire: a new and effective screening tool for Confusional Arousals, Sleepwalking and Sleep Terrors in epilepsy and sleep disorders units.

Giuseppe Loddo[1]&Giusy La Fauci[1], Luca Vignatelli[2], Corrado Zenesini[2], Rosalia Cilea[2], Francesco Mignani[2], Annagrazia Cecere[2], Susanna Mondini[2], Luca Baldelli[1], Francesca Bisulli[1,3], Laura Licchetta[1,3], Barbara Mostacci[3], Pietro Guaraldi[2], Giulia Giannini[1], Paolo Tinuper[1,3], Federica Provini[1,2]*

[1]Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italia

[2]IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italia

[3]IRCCS Istituto delle Scienze Neurologiche di Bologna, Reference Center for Rare and Complex Epilepsies – EpiCARE, Bologna, Italy

Giuseppe Loddo[1]&Giusy La Fauci[1]contributed equally to the manuscript.

***Corresponding author**: Federica Provini; IRCCS, Istituto delle Scienze Neurologiche di Bologna, Bologna, Italia; Bellaria Hospital; Via Altura, 3; 40139 Bologna (ITALY); Phone: ++ 39 051 4966829; e-mail: federica.provini@unibo.it

Abbreviations: ADQ, Arousal Disorders Questionnaire; DoA, Disorders of arousals; CA, Confusional Arousals; FLEP, Frontal Lobe Epilepsy and Parasomnias; ICSD-3, International Classification of Sleep Disorders-Third Edition; MUPS, Munich Parasomnia Screening; NES, Night Eating Syndrome ; PADSS, Paris Arousal Disorders Severity Scale; RBD, REM Sleep Behavior Disorder; SHE, Sleep-related Hypermotor Epilepsy; ST, Sleep Terrors; SW, Sleepwalking; VPSG, Video-polysomnography.

Abstract

Background: Arousal Disorders (DoA) include Confusional Arousals, Sleepwalking and Sleep Terrors. DoA diagnosis is mainly clinical but no validated questionnaires exist for DoA screening according to the criteria of the International Classification of Sleep Disorders, Third Edition. Recently our group proposed the Arousal Disorders Questionnaire (ADQ) as a new diagnostic tool for DoA diagnosis. The objective of this study was to evaluate the diagnostic accuracy of the ADQ in a sleep and epilepsy center.

Methods: One interviewer blinded to clinical and video-polysomnographic (VPSG) data administered the ADQ to 150 patients consecutively admitted to our Sleep and Epilepsy Centers for a follow-up visit. The final diagnosis, according to VPSG recordings of at least one major episode, classified patients either with DoA (DoA group) or with other sleep-related motor behaviors confounding for DoA (nDoA group).

Results: 47 patients (31%) composed the DoA group; 56 patients with REM sleep behavior disorder, 39 with sleep-hypermotor epilepsy, 6 with night eating syndrome, and 2 with drug-induced DoA composed the nDoA group. The ADQ had a sensitivity of 72% (95% CI: 60-82) and a specificity of 96% (95% CI: 89-98) for DoA diagnosis; excluding the items regarding consciousness and episode recall, sensitivity was 83% (95% CI: 71-90) and specificity 93% (95% CI: 86-97).

Conclusions: The ADQ showed good accuracy in screening patients with DoA in a sleep and epilepsy center setting. Diagnostic criteria related to cognition and episode recall reduced ADQ sensitivity, therefore a better definition of these criteria is required, especially in adults.

1. Introduction

According to the International Classification of Sleep Disorders-Third Edition (ICSD-3), Disorders of arousals (DoA) are NREM parasomnias including confusional arousals (CA), sleepwalking (SW) and sleep terrors (ST). Along with REM Sleep Behavior Disorder (RBD) also DoA are characterized by abnormal sleep related complex behaviors sometimes resulting in sleep injuries, sleep disruption and adverse health and psychosocial effects. These consequences can affect both patients and bedpartners [1]. When particularly frequent and violent, DoA may be mistaken for seizures, especially Sleep-related Hypermotor Epilepsy (SHE), previously called Nocturnal Frontal Lobe Epilepsy, which is a rare form of focal epilepsy characterized by brief and stereotyped seizures occurring predominantly during sleep [2].

Furthermore, almost 34% of SHE patients report DoA in their childhood [3] raising issues of differential diagnosis in adults. A correct diagnosis of sleep-related motor behaviors in adults is essential not only to set up a correct therapy and to prevent sleep-related injuries and/or relevant daytime consequences [4–7] but also for the prognosis [8].

The Frontal Lobe Epilepsy and Parasomnias (FLEP) scale has been proposed as a screening tool in order to differentiate SHE from DoA [9] but a risk of misdiagnosing DoA with SHE patients with ambulatory seizures has been reported. Furthermore, FLEP gave a misleading seizure diagnosis in around one third of RBD patients who underwent it [10].Contrary to RBD, whose diagnosis requires video-polysomnography (VPSG) recording [11], DoA diagnosis is based solely on clinical criteria although promising EEG markers have been recently proposed (e.g. slow-wave sleep fragmentation and slow/mixed arousal indexes) [12].

VPSG is currently indicated only for the evaluation of atypical, complicated, and injurious DoA or to exclude other associated sleep disorders such as obstructive sleep apnea syndrome or periodic limb movements [1,13,14], but it is an expensive procedure, not universally available and likely with low sensitivity. Indeed, in patients with less frequent events, it is rare to capture an episode during a single night. Screening tools are therefore essential to identify subjects with DoA.

The Munich Parasomnia Screening (MUPS) is a self-rating instrument with 21 items assessing the lifetime prevalence and current frequency of parasomnias and nocturnal behaviors in adults including SW, ST, CA. This questionnaire has a good sensitivity (83 to 100%) and specificity (89 to 100%) for DoA diagnosis but it

has not been validated against VPSG being the presence or absence of the sleep-related behaviors assessed only by the information obtained in a clinical interview with a sleep medicine expert [15].

The Paris Arousal Disorders Severity Scale (PADSS) assesses the severity of DoA with a self-rated scale evaluating episodes' frequency and consequences and listing behaviors observed during the episodes.

The scale identifies patients with SW and/or ST from healthy controls with 83% sensitivity and 98% specificity and from patients with RBD with a specificity of 89% but has not been designed according to the ICSD-3 criteria and it has not been validated in patients with SHE [16].

Recently, we proposed a standardized questionnaire (Arousal Disorders Questionnaire, ADQ) created by applying the ICSD-3 criteria with an almost perfect inter-rater reliability for DoA criteria and the final diagnosis among the raters [17]. The aim of our study was to evaluate the diagnostic accuracy of the ADQ in a clinical setting such as a sleep medicine and epilepsy unit.

2. Material and methods

For the publication of this study, we followed the recommendations of the STARD 2015 guidelines for reporting diagnostic accuracy studies [18]. The study was performed in accordance with the Declaration of Helsinki and was approved by the Bologna-Imola Ethics Committee in 15 March 2018 (no. 17175).

2.1 Patients

The study involved 150 patients >15 years reporting nocturnal behaviours consecutively admitted to the Sleep and Epilepsy Centers of the Institute of Neurological Sciences of Bologna between April 2018 and January 2020 for a follow-up visit. Non-native Italian-speaking subjects were excluded from the study. For each patient the diagnosis was previously established according to these VPSG findings:

1) DoA diagnosis was confirmed if VPSG documented at least one SW or ST episode;

2) SHE diagnosis was confirmed only if VPSG documented at least one hyperkinetic seizure associated with a clear-cut epileptic discharge or with interictal epileptiform abnormalities according to the diagnostic criteria established in the Consensus Conference held in Bologna in September 2014 [2];

3) RBD diagnosis was confirmed if VPSG documented episodes of sleep-related vocalization and/or complex motor behaviors during REM sleep associated with the presence of REM sleep without atonia [1];

4) Night Eating Syndrome (NES) diagnosis was confirmed if VPSG documented at least one night eating episode emerging before sleep onset or after an awakening from sleep.

VPSG included standard bipolar EEG (according to the International10-20 system), ECG, electro-oculogram, chin and limb EMG, and chest and abdominal respirograms. VPSG recordings were performed with synchronized, closed-circuit audio-video monitoring. The sleep stages were scored in accordance with the AASM version 2.4 criteria [11].

2.2 ADQ

The ADQ is a questionnaire assessing CA, SW and ST according to the ICSD-3 criteria [1,17]. For every single DoA the questionnaire includes two parts, both of which must be completed and obtain positive answers to satisfy diagnostic criteria.

The first part is specific for the different DoA, assessing CA and ST with two items and SW with one item. Each item is in the form of a question: "During your life, have you ever…?" (**Table 1-S1).** If the subject provides a "yes" response to all the items she/he is invited to complete the second part of the questionnaire. This latter part is composed of five common items assessing general criteria for DoA. All the items need a positive answer in order to satisfy the criteria for DoA diagnosis (**Table 1-S1**) [17].

The details of the ADQ assessment and its interobserver reliability have been published previously [17].

Informed consent was obtained from all participants included in the study. After signing the informed consent, each patient underwent the ADQ interview in order to diagnose the occurrence of DoA at any time in the subject's life. The interview was administered after training by a general practitioner (GLF) blinded to any clinical and VPSG data and on the final diagnosis. Parental or another observer's input was collected for 93 subjects to improve the accuracy of lifelong history. Please, insert Table 1 here.

2.3 Statistical analysis

The main outcome measures of this study were sensitivity, specificity, likelihood ratio (LR+, LR-) and diagnostic odds ratio (DOR) of the ADQ. The variability of estimates was expressed for each measure with 95% confidence interval (95% CI).

Outcome measures were obtained according to the reference standard considering two possible scenarios to better understand the possible application of the questionnaire in ordinary clinical practice.

In the worst-case scenario (worst reference standard), patients with SHE and RBD, with a clinical history suggestive for DoA only in childhood, were considered false positive patients according to the ADQ (index test).

In the best-case scenario (best reference standard), patients with SHE and RBD, with a clinical history suggestive for DoA only in childhood, were considered true positive patients according to the ADQ.

To understand the influence of ADQ items on its accuracy, outcome measures were obtained excluding individually one of the five common items assessing general criteria for DoA (e.g. A, B, C or D criteria) or two items (e.g. A and B, A and C, A and D, B and C, B and D or C and D criteria).

Two subgroup analyses were further conducted: 1) outcome measures were calculated according to the presence of any witnesses during the interview; patients' age (< 65 vs. ≥ 65 years); education (≤ 8 vs. > 8 years) and comorbidity (absence vs. presence of neurological and psychiatric comorbidity)". The Delong test was used to compare the diagnostic accuracy between the groups in the best and worst scenario; 2) outcome measures were individually calculated comparing the DoA and RBD group and the DoA and SHE group in the best- and worst-case scenario.

Statistical analysis was performed using the Stata SE, 14.2 statistical package (Stata Corp., Tulsa, OK, USA).

3. Results

3.1 Patients

The sample comprised subjects (90 males, 60 females) aged 17 – 82 years (mean age ± standard deviation, 52 ± 17 years). Duration of education was 5–22 years (mean age ± standard deviation, 13 ± 4 years).

The DoA group was composed of 47 patients (31%).

The nDoA group (103 patients) was composed of 56 RBD (38%), 39 SHE (26%), 6 NES (4%) and 2 patients with drug-induced DoA (1%). Patients' clinical features are described in **Table 2**.

Two RBD and 9 SHE patients had suffered in infancy from a sleep disorder suggestive for DoA. In these patients the clinical interview and the VPSG recording excluded the presence of current active DoA.

Please, insert Table 2 here.

3.2 ADQ Accuracy

Thirty-three DoA (70%) and 13 nDoA (12%) patients were positive on the ADQ, among the latter 11 were SHE patients and two RBD patients. Among them, 7 SHE and both RBD patients presented with a clinical history suggestive for DoA only during childhood.

The 11 SHE cases misdiagnosed by the ADQ were mainly patients with SW (4 patients) or CA (4 patients) (**Table 3**). The 2 RBD cases were both positive for SW.

No patients with NES or with drug-induced DoA were positive on the ADQ. The number and percentage of positive observations for each different DoA is reported in **Table 3**.

Please, insert Table 3 here.

In the worst-case scenario the ADQ had a sensitivity of 72% (95% CI: 57-83) and a specificity of 86% (95% CI: 78-91) (LR + 5.1, 95%, CI: 3.1-8.5; LR- 0.32, 95% CI: 0.20-0.53) (**Table 4**). In the best-case scenario it had the same sensitivity (95% CI: 60-82) but a higher specificity, i.e. 96% (95% CI: 89-98) (LR + 16.6, 95% CI: 6.3-44; LR- 0.28, 95% CI: 0.19-0.44) (**Table 5**).

Modifying ADQ items the best sensitivity values were found only excluding C and D criteria in both the best (83%, 95% CI: 67-90) and the worst-case scenario (81%, 95% CI: 71-90). Modification of other items worsened or did not change the ADQ specificity in both scenarios (**Table 4-5**).

Please, insert Table 4-5 here.

3.3 ADQ accuracy in relation to episodes' witnesses, patients' age, education and comorbidity *(sub-analysis 1)*

The group of patients with more than 8 years of education (n = 100) showed a better diagnostic accuracy (p values < 0.05) than the patients with education of less than or equal to 8 years (n = 27) (for 23 patients the education was missing) also excluding individually one of the five common items assessing general criteria for DoA (e.g. A, B, C or D criteria) or two items (e.g. A and B, A and C, A and D, B and C, B and D or C and D criteria). It was verified both in the best (ROC area: 0.86 vs 0.60; p=0.0135) and in the worst scenario (ROC area: 0.77 vs 0.48; p<0.0001). No significant differences in diagnostic accuracy were found in either scenario, between the groups with or without witnesses (respectively 93 and 57 patients); between the groups with age < 65 vs. ≥ 65 years (respectively 104 and 46 patients); between the groups with absence or presence of neurological and psychiatric comorbidity (respectively 116 and 34 patients).

3.4 ADQ accuracy comparing DoA and RBD patients (sub-analysis 2)

Comparing individually DoA and RBD patients, in the worst-case scenario the ADQ had a sensitivity of 70% (95% CI: 56-81) and a specificity of 96% (95% CI: 88-99) (LR+ 19.7, 95% CI: 5-77; LR- 0.31, 95% CI: 0.20-0.48) (**Table S2**). In the best-case scenario it had a similar sensitivity (71%, 95% CI: 58-82) and 100% specificity (95% CI: 93-100) (LR + not calculable, perfect specificity; LR- 0.28, 95% CI: 0.18-0.45) (**Table S3**).

3.5 ADQ accuracy comparing DoA and SHE patients (sub-analysis 2)

Comparing singularly DoA and SHE patients, in the worst-case scenario the ADQ had a sensitivity of 70% (95% CI: 56-81) and a specificity of 72% (95% CI:56-83) (LR+ 2.5, 95% CI:1.5-4.2; LR- 0.41, 95% CI: 0.26-0.67) (**Table S4**). In the best-case scenario it had a sensitivity of 71% (95% CI:59-82) and a specificity of 87% (95% CI:70-95) (LR+ 5.4, 95% CI:2.1-13.5; LR- 0.33, 95% CI:0.21-0.51) (**Table S5**).

4. Discussion

ADQ has good accuracy in screening DoA in our sleep and epilepsy centers evaluating, for the first time, with the support of VPSG, the value of the ICSD-3 single and general clinical criteria for DoA diagnosis. Although DoA diagnosis is mainly clinical, no validated screening instruments exist for DoA assessment according to the ICSD-3 criteria which represent the reference standard. Nevertheless, a correct diagnosis in adults is essential to prevent sleep-related injuries or relevant psychosocial consequences [4–7] and to differentiate DoA from other sleep-related behaviors characterized by a different prognosis and requiring specific therapies [1,19].

The clinical impact of our study is exemplified in the scenarios in Table 6 where the diagnostic value of the ADQ is listed according to different baseline pre-test probabilities of DoA according to Fagan's nomogram [20].

When a patient with nocturnal motor-behavioral episodes is negative for the ADQ, in the context of an epilepsy center, DoA diagnosis can be reasonably ruled out; in the context of a sleep center, we suggest a clinical follow-up or home-video monitoring.

In the case of a patient who is positive for the ADQ, both in a sleep or an epilepsy center, in typical cases (onset during childhood, less than one episode per month, episodes in the first third of the night with long duration or non-injurious or non-stereotyped motor pattern, no suspicion of other sleep disorders triggering DoA [21]) we suggest a clinical follow-up or home-video monitoring in order to confirm the diagnosis. In atypical cases (adult onset, high frequency of episodes per night, events occurring in any part of the night with short duration or stereotyped or injurious pattern, suspicion of other sleep disorders triggering DoA [21]) a VPSG recording is mandatory. ADQ revealed a good specificity especially when administered to patients with RBD (96% specificity after sub-analysis in the worst-case scenario and 100% in the best-case scenario) (**Table S2-S3**). RBD shares with DoA common behaviors (vocalization and intense motor activity) and consequences (patient and bedpartner injuries, such as bruising, lacerations and fractures) [1,7,22,23]. Usually, in RBD, the memory of dream mentation appropriate to the observed behavior, the complete alertness and orientation on awakening, and the absence of ambulatory movements are clinically helpful features distinguishing RBD from DoA [1]. All these features are systematically well explored with the ADQ and this may explain the accuracy of the questionnaire in discriminating DoA from RBD. In patients with

RBD who have no memory of their episodes, especially if they live alone, the ADQ showed a good specificity also if the recall of the episodes (D criterion) is not included (**Table S2-S3**). The ADQ questionnaire will be therefore useful to recognize DoA in patients > 50 years with complex motor behaviors during sleep (in which RBD is more often suspected [1]) or in patients with Parkinson's disease in which RBD is frequent but also SW can be often observed [24].The association between DoA and RBD refers to a condition described as parasomnia overlap disorder [25–27]. The prevalence of this disorder is unknown [26] but, if suspected, the ADQ could support an early diagnosis before VPSG recording.

Among SHE patients the ADQ revealed the worst specificity. Other instruments used to screen DOA from SHE were not satisfactory in differentiating them [10,28].

SHE seizures may have a bizarre semeiology, with vocalization, complex automatisms, and ambulation. Conversely, some DoA episodes may be violent and can be confused with SHE [29]. This probably justifies the positivity of some SHE patients on SW or CA after being administered the ADQ (**Table 3**). In addition, the ADQ explores lifetime DoA prevalence and at least one third of SHE patients [3] presented DoA during their childhood. For this reason, we chose to evaluate ADQ accuracy in two scenarios to better represent diagnostic problems often encountered in ordinary clinical practice. In the worst-case scenario, patients with SHE (but also RBD) were considered false positive patients according to the ADQ (index test) if they had a clinical history suggestive for DoA only in childhood, according to a sleep/epilepsy expert interview. In the best-case scenario they were considered true positive patients and, in this perspective, the specificity of the ADQ grows to 86%. We are confident that this scenario fits the context of a sleep or epilepsy center where sleep and epilepsy experts should guarantee a thorough clinical interview exploring DoA in current or past medical history. Considering that SHE is a rare form of focal epilepsy, with an estimated minimum prevalence of 1.8/100,000 individuals [30], we argue that ADQ specificity can be acceptable also in primary care practice or neurology clinics in which SHE prevalence is reasonably rare as described in literature.

In relation to sensitivity, the study showed that not all the ICSD-3 criteria used to screen DoA have satisfactory values. The best ADQ sensitivity values in both scenarios were indeed obtained only without C and D criteria (81-83%) (**Table 4-5**), achieving values similar to those of the MUPS and PADSS [15,16]. C and D ADQ criteria investigate consciousness and recall of the episodes, respectively. Their influence on

ADQ sensitivity can be explained by considering the intrinsic ambiguity or the difficulties in correctly interpreting cognition during the episodes and the instability of patient responses linked to past events difficult to recall. In addition, the complete amnesia for the DoA events is not the rule in adults, while it might be more common in children, possibly because of higher arousal thresholds [4,7,31]. It is therefore possible that C and D criteria in adults, although important to discriminate DoA from SHE patients (**Table S4-S5**), require a better definition and discussion in future versions of the ICSD.

None of the patients with NES were positive on the ADQ. This result is encouraging, although only 6 patients were enrolled, it is justified by the fact that in all patients eating episodes were recorded during full wakefulness (as explored by the A criterion of the ADQ). Further studies including other parasomnias (e.g., sleep-related eating disorder, SRED) should be performed to confirm the accuracy of the ADQ. A similar consideration can be made when taking into account the 2 patients with drug-induced DoA. They were both negative on the ADQ only because of the E criterion ("can you rule out that these episodes were due to other sleep problems, medical conditions, use of drugs or other substances?"). Criterion E, as reported in Tables S3-S4, is essential to guide the clinician to consider the full clinical context in which parasomnias appear. In particular, it could be also relevant in the differential diagnosis between DoA and SHE because, at least in some cases, as SHE patients seem capable of discriminating active seizures from DoA episodes of their childhood.

There are, however, some limitations to this study. First of all, the study was monocentric involving only our centre, the validation was performed retrospectively, and a recall bias may have influenced the patients' responses. Furthermore, ADQ was not directly performed by a sleep or epilepsy expert and this probably reduced the sensitivity of the questionnaire in particular in patients with low education in which the questionnaire had a lesser accuracy.

In conclusion, the ADQ showed good accuracy in screening patients with DoA in the setting of a sleep medicine and epilepsy center.

In case of a clinical history typical for DoA, we argue that the ADQ positivity can reasonably avoid VPSG to confirm diagnosis. In case of atypical clinical histories or in case of clinical DoA suspicion, with ADQ negativity, VPSG remains essential to confirm DoA diagnosis. Probably, a better definition of ICSD-3

general DoA diagnostic criteria, in particular those regarding cognition (C criterion) and recall of the episodes (D criterion), is required especially for adults.

Please, insert Table 6 here.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

This study is not industry-sponsored. Dr Giuseppe Loddo, Dr Giusy La Fauci, Dr Luca Vignatelli, Dr Corrado Zenesini, Dr Rosalia Cilea, Dr Francesco Mignani, Dr Annagrazia Cecere, Dr Susanna Mondini, Dr Luca Baldelli, Dr Francesca Bisulli, Dr Laura Licchetta, Dr Barbara Mostacci, Dr Pietro Guaraldi, Dr Giulia Giannini declare no conflicts of interest.

Prof. Tinuper reports non-financial support and other from EISAI, non-financial support and other from Livanova, nonfinancial support and other from UCB, non-financial support and other from GW, outside the submitted work.

Prof. Provini reports personal fees from Vanda Pharmaceutical, personal fees from Sanofy, personal fees from Zambon, personal fees from Fidia, personal fees from Bial, personal fees from Eisai Japan, personal fees from Italfarmaco, outside the submitted work.

References

[1] American Academy of Sleep Medicine. International Classification of Sleep Disorders, 3rd ed. Darien, IL: American Academy of Sleep Medicine. 2014.

[2] Tinuper P, Bisulli F, Cross JH, Hesdorffer D, Kahane P, Nobili L, et al. Definition and diagnostic criteria of sleep-related hypermotor epilepsy. Neurology 2016;86:1834.

[3] Provini F, Plazzi G, Tinuper P, Vandi S, Lugaresi E, Montagna P. Nocturnal frontal lobe epilepsy: A clinical and polygraphic overview of 100 consecutive cases. Brain 1999;122:1017–1031.

[4] Zadra A, Desautels A, Petit D, Montplaisir J. Somnambulism: clinical aspects and pathophysiological hypotheses. Lancet Neurol 2013;12:285–294.

[5] Cartwright R. Sleepwalking violence: a sleep disorder, a legal dilemma, and a psychological challenge. Am J Psychiatry 2004;161:1149–58.

[6] Siclari F, Khatami R, Urbaniok F, Nobili L, Mahowald MW, Schenck CH, et al. Violence in sleep. Brain J Neurol 2010;133:3494–509.

[7] Baldini T, Loddo G, Sessagesimi E, Mignani F, Cirignotta F, Mondini S, et al. Clinical features and pathophysiology of disorders of arousal in adults: a window into the sleeping brain. Front Neurol 2019;10:526.

[8] Licchetta L, Bisulli F, Vignatelli L, Zenesini C, Di Vito L, Mostacci B, et al. Sleep-related hypermotor epilepsy: Long-term outcome in a large cohort. Neurology 2017;88:70–7.

[9] Derry CP, Davey M, Johns M, Kron K, Glencross D, Marini C, et al. Distinguishing sleep disorders from seizures: diagnosing bumps in the night. Arch Neurol 2006;63:705–9.

[10] Manni R, Terzaghi M, Repetto A. The FLEP scale in diagnosing nocturnal frontal lobe epilepsy, NREM and REM parasomnias: Data from a tertiary sleep and epilepsy unit. Epilepsia 2008;49:1581–1585.

[11] Berry RB, Brooks R, Gamaldo C, Harding SM, Lloyd RM, Quan SF, et al. AASM scoring manual updates for 2017 (version 2.4). J Clin Sleep Med 2017;13:665–666.

[12] Lopez R, Shen Y, Chenini S, Rassu AL, Evangelista E, Barateau L, et al. Diagnostic criteria for disorders of arousal: A video-polysomnographic assessment: Disorders of Arousal. Ann Neurol 2018;83:341–51.

[13] Drakatos P, Marples L, Muza R, Higgins S, Nesbitt A, Dongol EM, et al. Video polysomnographic findings in non-rapid eye movement parasomnia. J Sleep Res 2019;28:e12772.

[14] Fois C, Wright M-AS, Sechi G, Walker MC, Eriksson SH. The utility of polysomnography for the diagnosis of NREM parasomnias: an observational study over 4 years of clinical practice. J Neurol 2015;262:385–93.

[15] Fulda S, Hornyak M, Müller K, Cerny L, Beitinger PA, Wetter TC. Development and validation of the Munich Parasomnia Screening (MUPS): A questionnaire for parasomnias and nocturnal behaviors. Somnologie 2008;12:56–65.

[16] Arnulf I, Zhang B, Uguccioni G, Flamand M, de Fontréaux AN, Leu-Semenescu S, et al. A Scale for Assessing the Severity of Arousal Disorders. Sleep 2014;37:127–136.

[17] Loddo G, Vignatelli L, Zenesini C, Lusa F, Sambati L, Baldelli L, et al. Interobserver reliability of ICSD-3 diagnostic criteria for disorders of arousal in adults. Sleep Breath 2019;23:1309–14.

[18] Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open 2016;6:e012799.

[19] Zhang X, Sun X, Wang J, Tang L, Xie A. Prevalence of rapid eye movement sleep behavior disorder (RBD) in Parkinson's disease: a meta and meta-regression analysis. Neurol Sci 2017;38:163–70.

[20] Fagan TJ. Letter: Nomogram for Bayes theorem. N Engl J Med 1975;293:257.

[21] Loddo G, Lopez R, Cilea R, Dauvilliers Y, Provini F. Disorders of Arousal in adults: new diagnostic tools for clinical practice. Sleep Sci Pract 2019;3:5.

[22] Uguccioni G, Golmard J-LL, de Fontréaux AN, Leu-Semenescu S, Brion A, Arnulf I, et al. Fight or flight? Dream content during sleepwalking/sleep terrors vs rapid eye movement sleep behavior disorder. Sleep Med 2013;14:391–398.

[23] Schenck CH, Bundlie SR, Ettinger MG, Mahowald MW. Chronic behavioral disorders of human REM sleep: a new category of parasomnia. Sleep 1986;9:293–308.

[24] Oberholzer M, Poryazova R, Bassetti CL. Sleepwalking in Parkinson's disease: a questionnaire-based survey. J Neurol 2011;258:1261–7.

[25] Schenck CH, Howell MJ. Spectrum of rapid eye movement sleep behavior disorder (overlap between rapid eye movement sleep behavior disorder and other parasomnias). Sleep Biol Rhythms 2013;11:27–34.

[26] Dumitrascu O, Schenck CH, Applebee G, Attarian H. Parasomnia overlap disorder: A distinct pathophysiologic entity or a variant of rapid eye movement sleep behavior disorder? A case series. Sleep Med 2013;14:1217–1220.

[27] Kohler WC, Kurz PJ, Kohler EA. A Case of Successful Use of Hypnosis in the Treatment of Parasomnia Overlap Disorder. Behav Sleep Med 2015;13:349–358.

[28] Bisulli F, Vignatelli L, Naldi I, Pittau F, Provini F, Plazzi G, et al. Diagnostic accuracy of a structured interview for nocturnal frontal lobe epilepsy (SINFLE): A proposal for developing diagnostic criteria. Sleep Med 2012;13:81–87.

[29] Bisulli F, Vignatelli L, Provini F, Leta C, Lugaresi E, Tinuper P. Parasomnias and nocturnal frontal lobe epilepsy (NFLE): lights and shadows–controversial points in the differential diagnosis. Sleep Med 2011;12 Suppl 2:S27–32.

[30] Vignatelli L, Bisulli F, Giovannini G, Licchetta L, Naldi I, Mostacci B, et al. Prevalence of nocturnal frontal lobe epilepsy in the adult population of Bologna and Modena, Emilia-Romagna region, Italy. Sleep 2015;38:479–85.

[31] Pilon M, Desautels A, Montplaisir J, Zadra A. Auditory arousal responses and thresholds during REM and NREM sleep of sleepwalkers and controls. Sleep Med 2012;13:490–495.

Table 1. Arousal Disorders Questionnaire (ADQ), adapted from Loddo et al 2019 [17].

CONFUSIONAL AROUSALS

	YES	NO
A) During your life, have you ever or have you ever been told that while sleeping you suddenly woke up and were confused or behaved in a confused manner?	□	□
B) If yes, were night terrors and sleepwalking excluded during these episodes?	□	□

IF THE ANSWER WAS YES, ASK:

	YES	NO
A. During these episodes were you partially awake?	□	□
B. During these episodes did you ever not respond or respond inappropriately to someone's attempts to intervene or guide you?	□	□
C. During these episodes were you partially conscious (for example did you see a single visual scene)? Were you unconscious? Did you see images as if you were dreaming? (Tick Yes if one or more of the answers are positive)	□	□
D. Did you have little or no recall of the episode?	□	□
E. Can you rule out that these episodes were due to other sleep problems, medical conditions, use of drugs or other substances?	□	□

SLEEPWALKING

	YES	NO
A) During your life, have you ever or have you ever been told that you got out of bed and started walking or presenting other complex behaviors out of bed?	□	□

IF THE ANSWER WAS YES, ASK:

	YES	NO
A. During these episodes were you partially awake?	□	□
B. During these episodes did you ever not respond or respond inappropriately to someone's attempts to intervene or guide you?	□	□
C. During these episodes were you partially conscious (for example did you see a single visual scene)? Were you unconscious? Did you see images as if you were dreaming? (Tick Yes if one or more of the answers are positive)	□	□
D. Did you have little or no recall of the episode?	□	□
E. Can you rule out that these episodes were due to other sleep problems, medical conditions, use of drugs or other substances?	□	□

SLEEP TERRORS

	YES	NO
A) During your life, have you ever had or have you ever been told that during sleep you presented episodes of sudden fear typically starting with alarmed vocalizations like a shout of fear?	□	□
B) If yes, was this episode accompanied by dilation of pupils, tachycardia, rapid breathing, and sweating?	□	□

IF THE ANSWER WAS YES, ASK:

	YES	NO
A. During these episodes were you partially awake?	□	□
B. During these episodes did you ever not respond or respond inappropriately to someone's attempts to intervene or guide you?	□	□
C. During these episodes were you partially conscious (for example did you see a single visual scene)? Were you unconscious? Did you see images as if you were dreaming? (Tick Yes if one or more of the answers are positive)	□	□
D. Did you have little or no recall of the episode?	□	□
E. Can you rule out that these episodes were due to other sleep problems, medical conditions, use of drugs or other substances?	□	□

Table 2. Clinical features of the patients who were administered the Arousal Disorders Questionnaire (ADQ).

	DoA	**SHE**	**RBD**	**NES**	**DID §**
Total number of patients	47	39	56	6	2
Age, yrs ± SD (range)	40 ± 15 (17-74)	42 ± 13 (21-82)	67 ± 9 (43-81)	63 ± 13 (47-78)	66 ± 12 (58-75)
Male/Female, n (%)	26 (55)/21 (45)	17 (43)/22 (57)	45 (80)/11 (20)	1 (16)/5 (84)	2 (100)/0 (0)
CONCOMITANT MEDICAL DISEASE, n (%)					
Allergic diseases	5 (10)	0 (0)	0 (0)	0 (0)	0 (0)
Thyroid diseases	4 (8)	3 (7)	6 (10)	1 (16)	0 (0)
Heart diseases	0 (0)	1 (2)	9 (16)	1 (16)	0 (0)
Hypertension	3 (6)	0 (0)	15 (26)	3 (50)	1 (50)
Diabetes	1 (2)	1 (2)	7 (12)	2 (33)	1 (50)
CONCOMITANT NEUROLOGIC DISEASE, n (%)					
Migraine	3 (6)	1 (2)	0 (0)	0 (0)	0 (0)
Tension-type headache	2 (4)	1 (2)	0 (0)	1 (16)	1 (50)
Parkinson's Disease	0 (0)	0 (0)	18 (32)	0 (0)	0 (0)
Multiple System Atrophy	0 (0)	0 (0)	3 (5)	0 (0)	0 (0)
Pure Autonomic Failure	0 (0)	0 (0)	6 (10)	0 (0)	0 (0)
CONCOMITANT PSYCHIATRIC DISEASES, n (%)					
Depression	5 (10)	2 (5)	8 (14)	2 (33)	1 (50)
Anxiety	5 (10)	0 (0)	7 (12)	1 (16)	0 (0)
CONCOMITANT SLEEP DISORDERS, n (%)					
Insomnia	3 (6)	0 (0)	4 (7)	4 (66)	1 (50)
Obstructive sleep apnea (AHI <15)	2 (4)	2	9 (16)	2 (33)	0 (0)

SD: Standard Deviation; DoA: Disorders of Arousal; SHE: Sleep-related Hypermotor Epilepsy; RBD: REM Behaviour Disorder; NES: Night Eating Syndrome; DID: Drug-induced Disorders of Arousal.

Table 3. Number and percentage of positive observations on Arousal Disorders Questionnaire ADQ) for Confusional Arousals (CA), Sleepwalking (SW) and Sleep Terrors (ST) in the group of patients with Disorders of Arousal (DoA) and without DoA (nDoA).

	DoA group	**nDoA group**	**nDoA (SHE)**	**nDoA (RBD)**	**nDoA (NES)**	**nDoA (DID)**
	(47)	(103)	(39)	(56)	(6)	(2)
ADQ Diagnosis	**33 (70)**	**13 (12)**	**11 (28)**	**2 (3)**	**0 (0)**	**0 (0)**
CA, n (%)	5 (11)	5 (5)	4 (10)	0 (0)	0 (0)	0 (0)
SW, n (%)	11 (23)	5 (5)	4 (10)	2 (4)	0 (0)	0 (0)
ST, n (%)	1 (2)	1 (1)	1 (3)	0 (0)	0 (0)	0 (0)
CA + SW, n (%)	2 (4)	1 (1)	1 (3)	0 (0)	0 (0)	0 (0)
CA + ST, n (%)	3 (6)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
SW + ST, n (%)	6 (13)	1 (1)	1 (3)	0 (0)	0 (0)	0 (0)
CA + SW + ST n (%)	5 (11)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

DoA: Disorders of Arousal; SHE: Sleep-related Hypermotor Epilepsy; RBD: REM Behaviour Disorder; NES: Night Eating Syndrome; § Drug-induced Disorders of Arousal.

Table 4. Diagnostic validity of the Arousal Disorders Questionnaire (ADQ) in the worst-case scenario according to its different criteria.

	Sensitivity (95% CI)	**Specificity (95% CI)**	**LR+ (95% CI)**	**LR- (95% CI)**	**DOR (95% CI)**
All ADQ criteria	72% (57-83)	86% (78-91)	5.1 (3.1-8.5)	0.32 (0.20-0.53)	15.8 (6.6-37.4)
ADQ without A criterion	74% (60-85)	80% (72-87)	3.8 (2.5-5.8)	0.31 (0.19-0.53)	11.9 (5.1-27.4)
ADQ without B criterion	72% (57-83)	78% (70-85)	3.4 (2.2-5.0)	0.35 (0.22-0.58)	9.4 (4.1-21.2)
ADQ without C criterion	77% (62-87)	85% (77-91)	5.1 (3.2-8.3)	0.27 (0.16-0.47)	18.7 (7.7-45.4)
ADQ without D criterion	77% (62-87)	84% (76-90)	4.8 (3.0-7.7)	0.27 (0.16-0.48)	17.4 (7.2-41.9)
ADQ without E criterion	74% (60-85)	79% (71-86)	3.6 (2.4-5.5)	0.32 (0.19-0.54)	11.2 (4.9-25.7)
ADQ without A and B criteria	74% (60-85)	72% (63-80)	2.7 (1.9-3.8)	0.35 0.21-0.60)	7.4 (3.3-16.6)
ADQ without A and C criteria	79% (65-89)	78% (70-85)	3.7 (2.5-5.5)	0.26 (0.15-0.48)	13.7 (5.7-32.8)
ADQ without A and D criteria	76% (62-87)	78% (70-85)	3.57 (2.40-5.31)	0.29 (0.17-0.51)	12.05 (5.1-28.0)
ADQ without B and C criteria	76% (62-87)	77% (69-84)	3.4 (2.3-5.0)	0.30 (0.17-0.52)	11.4 (4.9-26.4)
ADQ without B and D criteria	79% (65-89)	74% (65-81)	3.0 (2.1-4.3)	0.28 (0.16-0.51)	10.6 (4.5-24.9)
ADQ without C and D criteria	81% (67-90)	82% (74-88)	4.6 (3.0-7.1)	0.22 (0.12-0.43)	20.2 (8.1-50.5)

CI: Confidence Interval; LR: likelihood ratio; DOR: diagnostic odds ratio

Table 5. Diagnostic validity of the Arousal Disorders Questionnaire (ADQ) in the best-case scenario according to its different criteria.

	Sensitivity (95% CI)	**Specificity (95% CI)**	**LR+ (95% CI)**	**LR- (95% CI)**	**DOR (95% CI)**
All ADQ criteria	72% (60-82)	96% (89-98)	16.6 (6.3-44.0)	0.28 (0.19-0.44)	57.7 (18.1-183.4)
ADQ without A criterion	76% (63-85)	90% (82-95)	7.7 (4.1-14.7)	0.27 (0.17-0.42)	28.9 (11.6-72.2)
ADQ without B criterion	72% (60-82)	87% (79-92)	5.5 (3.2-9.6)	0.31 (0.20-0.48)	17.5 (7.5-40.3)
ADQ without C criterion	78% (65-86)	96% (89-98)	17.8 (6.8-47.0)	0.23 (0.15-0.38)	76.1 (23.4-247.0)
ADQ without D criterion	76% (63-85)	93% (86-97)	11.6 (5.3-25.6)	0.26 (0.16-0.41)	45.04 (16.1-125.3)
ADQ without E criterion	74% (61-84)	88% (80-93)	6.2 (3.5-11.0)	0.29 (0.19-0.46)	21.1 (8.9-49.9)
ADQ without A and B criteria	76% (63-85)	80% (72-88)	3.9 (2.5-6.0)	0.30 (0.19-0.48)	12.9 (5.8-28.5)
ADQ without A and C criteria	83% (71-90)	90% (82-95)	8.5 (4.5-15.9)	0.19 (0.11-0.34)	44.2 (16.8-116.5)
ADQ without A and D criteria	78% (65-86)	88% (80-93)	6.5 (3.7-11.5)	0.26 (0.16-0.41)	25.4 (10.5-61.5)
ADQ without B and C criteria	78% (65-86)	87% (79-92)	5.9 (3.4-10.2)	0.25 (0.16-0.42)	23.07 (9.7-54.8)
ADQ without B and D criteria	78% (65-86)	81% (72-88)	4.2 (2.7-6.6)	0.27 (0.17-0.45)	15.2 (6.7-34.3)
ADQ without C and D criteria	83% (71-90)	93% (86-97)	12.7 (5.8-27.8)	0.18 (0.11-0.33)	68.8 (23.5-200.9)

CI: Confidence Interval; LR: likelihood ratio; DOR: diagnostic odds ratio.

Table 6. Diagnostic validity of the Arousal Disorders Questionnaire (ADQ) in different settings in the best-case scenario excluding C and D criteria.

DoA pre-test probability in different settings	**Negative for ADQ LR 0.18**		**Positive for ADQ LR 12.7**	
	DoA post-test probability (%)	*Further diagnostic step*	*DoA post-test probability (%)*	*Further diagnostic step*
Epilepsy center visit (12%)	2	Possible clinical follow-up	67	**Typical cases §:** Possible clinical follow-up/home video monitoring **Atypical cases §:** VPSG
Sleep center visit (31%)	7	Possible clinical follow-up	85	**Typical cases §:** Possible clinical follow-up in typical cases/home video monitoring **Atypical cases §:** VPSG

DoA: Arousal Disorders; LR: likelihood ratio; §: definition of typical and atypical DoA is reported in the discussion.

Applying the likelihood ratios corresponding to the ADQ in the best scenario excluding C and D criteria (derived from specificity and sensitivity) to Fagan's nomogram [20], we calculated the post-test probability (positive predictive value) of DoA in different settings. The pre-test probability of NFLE was calculated according to the following data. In our epilepsy center 40 out of 370 patients per year have a history of nocturnal motor-behavioral episodes, five of them (12%) received a DoA diagnosis. In our Sleep Medicine Centre 70 out of 480 patients per year come to our attention complaining nocturnal motor-behavioral episodes and 21 of them (31%) received a DoA diagnosis.

Giuseppe Loddo: Conceptualization, Writing - original draft, Writing - review & editing, Investigation, Resources. **Giusy La Fauci:** Investigation. **Luca Vignatelli:** Conceptualization, Methodology, Writing - review & editing, Visualization, Validation. **Corrado Zenesini:** Conceptualization, Methodology, Formal analysis, Writing - review & editing. **Rosalia Cilea:** Investigation**. Francesco Mignani:** Investigation**. Annagrazia Cecere:** Investigation. **Susanna Mondini:** Investigation, Resources. **Luca Baldelli:** Investigation, Resources. **Francesca Bisulli:** Investigation, Resources**. Laura Licchetta:** Investigation, Resources. **Barbara Mostacci:** Investigation, Resources. **Pietro Guaraldi:** Investigation, Resources. **Giulia Giannini:** Investigation, Resources. **Paolo Tinuper:** Investigation, Resources, Supervision. **Federica Provini:** Conceptualization, Investigation, Resources, Writing - review & editing, Supervision, Project administration.