Two-group classification of patients with obstructive sleep apnea based on analysis of brain recurrence

Paul Y. Kim a, David E. McCarty a, Lei Wang a, Clifton Frilot II b, Andrew L. Chesson Jr. a, Andrew A. Marino a,⇑

a Division of Sleep Medicine, Department of Neurology, LSU Health Sciences Center, Shreveport, LA, United States
b School of Allied Health Professions, LSU Health Sciences Center, Shreveport, LA, United States

Keywords:
Brain recurrence analysis
Obstructive sleep apnea
Sleep depth
Sleep fragmentation
Linear discrimination analysis
Biomarkers

Abstract

Objective: To demonstrate that the severity of obstructive sleep apnea (OSA) could be predicted algorithmically by means of recurrence analysis of the sleep-staged electroencephalogram (EEG).

Methods: A randomly selected cohort of 20 sleep-staged patients with OSA (apnea–hypopnea index (AHI) 5–30) was divided into mild and moderate sub-cohorts (AHI 5–15, 16–30, respectively), and the sleep EEG (C3) was analyzed using analysis of brain recurrence (ABR) (LSU cohort). Twenty distinct but related markers for sleep depth and fragmentation were computed from four ABR variables, and a marker function capable of classifying each patient into one of the two sub-cohorts was determined by linear discriminant analysis. Classification accuracy of individual patients was evaluated using area under the receiver operator characteristics curve (AUROC). As a control procedure, 20 additional sleep-staged patients with OSA whose polysomnographic data was obtained from an independent database were also evaluated (SHHS cohort).

Results: On average, markers for sleep depth were reduced and those for sleep fragmentation were increased in the patients with moderate OSA, as expected. All patients in both cohorts were correctly classified using as few as 5–6 markers.

Significance: The degree of severity of OSA was reflected in objective changes in the sleep EEG. Recurrence analysis of the EEG potentially has uses beyond identification of the degree of OSA.

1. Introduction

Obstructive sleep apnea (OSA) is a disorder characterized by periodic collapse of the upper airway during sleep, resulting in intermediate hypoxia, hypercarbia, and variable degrees of sleep disruption. Blood oxygen levels are affected almost immediately by apneic events, resulting in afferent signals to the brain that initiate appropriate compensatory responses. These autonomic processes are accompanied by changes in brain electrical activity (Lurie, 2011; Zhang et al., 2013).

The gold-standard for assessing OSA severity is the apnea–hypopnea index (AHI), a rule-based measure of the frequency of the sum of both phenomena (American Academy of Sleep Medicine, 2007). Obstructive respiratory events fragment sleep and impair its restorative qualities (Jackson et al., 2011). Substantial evidence showed that high AHI levels were associated with serious health consequences including symptoms of wake impairment and risk for cardiovascular disease (Golbidi et al., 2012). The AHI is determined from physiological signals obtained during laboratory or home-based polysomnography, but these signals do not traditionally include data obtained from the
electroencephalogram (EEG) (American Academy of Sleep Medicine, 2007). Recently, EEG arousals have been incorporated into the scoring of hypopneas for adults, but this is not yet universally accepted (American Academy of Sleep Medicine, 2012). Consequently, information contained in the EEG typically remains unused in the process of assessing the severity of the OSA. In principle, the degree of severity of OSA would be expected to create changes in brain activity, which would possibly be reflected in measurable changes in the EEG. If so, such changes might be useful in connection with clinical evaluation of OSA patients, for example, by helping to classify patients into classes defined by the AHI.

Analysis of brain recurrence (ABR) is a novel method for quantifying brain activity as reflected in the EEG (Carrubba et al., 2012). We recently showed that sleep depth and sleep fragmentation could be objectively characterized using ABR (Wang et al., 2013). Poor quality sleep generally involves dysregulation of one or both processes, and may result from different sleep disorders, including OSA (Fietze et al., 1997). We were interested in the possibility that ABR-defined markers for sleep depth and fragmentation could be used to characterize OSA severity in individual patients, at least to the extent that patients known to belong to one of two classes could be reliably classified.

Our first objective was to find evidence that sleep-staging and ABR analysis of the EEG could be combined to yield markers for sleep depth and fragmentation that changed (on average) in the direction consistent with the ABR model. Our second objective was to show that the markers could be used to distinguish between patients with mild or moderate OSA. We planned to conclude that a classification accuracy >90% would indicate that ABR was potentially useful for classifying OSA severity based on EEG measurements.

2. Methods

2.1. Patients

We reviewed consecutive records of patients seen in a sleep-medicine clinic who underwent attended overnight polysomnography (PSG) for suspected OSA. The study cohort selected consisted of the first 20 consecutive patients who were diagnosed with OSA (AHI > 5 events/h) (LSU cohort). Exclusion criteria included <30 min of REM sleep (less than 4% of the PSGs examined), significant medical co-morbidities, current use of sleep-altering medications, and prior treatment for OSA. The cohort was divided into subcohorts with mild (AHI 5–15 events/h) and moderate (AHI 16–30 events/h) OSA (Table 1). The PSGs were staged by two sleep-medicine physicians, using standard rules (American Academy of Sleep Medicine, 2007). They agreed on 92% of the epochs; the remaining 8% were staged by consensus. Every 30-s epoch was staged as N1, N2, N3, REM, or wake after sleep onset (WASO). The N1 and N2 epochs were combined for statistical analysis (see below).

As a control for the potential role of the particular demographic characteristics associated with the LSU cohort, we also analyzed PSG data from a comparable group of OSA subjects obtained from the Sleep Heart Health Study (SHHS) (SHHS cohort), a multi-center study of the cardiovascular and other consequences of sleep-disordered breathing (Quan et al., 1997). An SHHS dataset collected in 2001 and 2003 (3295 PSGs) was searched to identify the participants for whom the AHI had been determined based on standard rules (National Heart Lung & Blood Institute). Our exclusion criteria (searchable fields in the database) were participants with heart failure, emphysema, chronic bronchitis, or hypertension. From the resulting group (N = 390) we randomly selected 10 participants with AHI of 5–15 events/h, and 10 with AHI of 16–30 events/h (SHHS cohort, Table 1). In general, the cohort was younger, contained more females, and had higher body mass index (BMI) measurements compared with the LSU cohort.

All research-related procedures were approved by the institutional review board for human research.

2.2. EEG measurements

The PSGs of the LSU subjects were recorded with commercial equipment (Respironics, Alice 5, Murrysville, PA, USA), using standard digital specifications and electrode montage (O1, O2, C3, C4, F3, F4, International 10–20 system) (Fietze et al., 1997). The EEGs were digitized at 500 Hz, and exported as CSV files for analysis. The EEGs of the SHHS subjects (C3, C4) were obtained as 250-Hz EDF files and interpolated to 500 Hz for analysis. All EEGs were digitally filtered to pass 0.5–35 Hz and evaluated using custom codes in a standard numerical computing environment (Matlab, Mathworks, Natick, MA, USA).

2.3. Analysis of brain recurrence

Analysis of brain recurrence (ABR) is a nonlinear technique for extraction of physiologically relevant information from the EEG. The technique is based on the complexity conjecture regarding the nature and origin of the EEG, namely that brain function is mediated by electrical activity in localized neuronal networks and their inter-network electrical synchronization (Carrubba et al., 2012). In this perspective, a given EEG signal is regarded as a dynamical measure of the instantaneous state of brain electrical activity, and the complexity of the EEG is quantified by means of ABR. The basic signal-processing techniques (and their special applicability to signals having the statistical properties of EEGs) were previously described in detail (Zbilut and Webber, 2006). Briefly, 5-component vectors were formed that consisted of the EEG amplitude at t and four earlier times identified by four successive lags of five points (10 ms). The sequence of all such vectors obtainable from one second of the EEG (480 vectors, given our choices of sampling rate, vector dimension, and number of lag points) was interpreted to be a result of law-governed (non-random) activity in the EEG. The amount of law-governed activity was quantified using the variables percent recurrence (r), defined as the percent of the 480 vectors in the path that were near other vectors (and hence were recurrent), and percent determinism (d), defined as the percent of the recurrent points that were adjacent to at least one other point. Detailed analysis of these variables provides a theoretical rationale for why they quantify the amount of law-governed activity in the EEG (Zbilut and Webber, 2006). The Euclidean norm was used for measuring distance, and vectors were identified as near if they were within 15% of the distance between the two vectors that were further apart. These choices as well as those for the other parameters used in the calculation were...
Both variables were computed for each second of the EEG, resulting in approximately $60 \times 60 \times 8 = 28,800$ values for a typical eight-hour overnight EEG. The sequence of values constituted the time series $r(t)$ and $d(t)$, which were interpreted as independent measures of sleep depth wherein higher values corresponded to deeper sleep. The time-series macroarchitectures exhibit ultradian cycles, with minimal values during wake after sleep onset (WASO) and maximal values during N3 sleep (Carrubba et al., 2012; Wang et al., 2013). Both variables were averaged over the entire overnight sleep study and separately over N1/N2, N3, REM, and WASO, and the values were used as markers of sleep depth in a statistical procedure to predict OSA severity (see below) (Fig. 1). We expected that the markers would be lower in patients with higher AHI values, indicating less restorative sleep.

Markers for fragmentation in sleep depth were created by generalizing the conventional definition of EEG arousals (American Academy of Sleep Medicine, 2007). For both $r(t)$ and $d(t)$, the ratio of the mean for 3 s (one value per second) to the mean for the preceding 10 s (ten values) was determined, and the process was repeated using successive steps of 3 s, resulting in a time series of approximately 9000 ratios for an overnight EEG. Whenever the ratio increased by more than 100% for $r(t)$ or 50% for $d(t)$ the change was counted as an arousal, and the hourly rate of arousals, termed the generalized arousal index (GAI) was determined for WASO, N1/N2, N3, and REM sleep separately, and for total sleep (TS) (all epochs in the PSG between sleep onset and lights on except for WASO epochs). The values during sleep were normalized by the value for WASO and expressed as percent change. The procedure was performed using $r(t)$ and $d(t)$, resulting in 5 GAI markers obtained from $r(t)$ and 5 from $d(t)$ (Fig. 1, line 5). We expected that the arousal rates would be higher in patients with more severe OSA.

Preliminary studies showed that none of the results presented here depended on whether the EEG signal analyzed was obtained from O1, O2, C3, C4, F3, or F4. Consequently, only results from C3 were presented here.

2.4. Statistics

Fisher’s linear discriminant analysis was used to determine the coefficients of $B$, a marker function that combined the sleep
markers in the way that best separated the mild (AHI 5–15) and moderate (AHI 16–30) patients (Theodoridis and Koutroumbas, 2008). The accuracy of the predictions was assessed using area under the receiver operating characteristics analysis (AUROC) (Theodoridis and Koutroumbas, 2008). The reliability of B was assessed by means of a 10-fold cross validation process in which B was determined using 90% of the sub-cohorts, evaluated by AUROC against the remaining 10%, and the process was repeated ten times with differing choices for the composition of the training and evaluation sets. The AUROC results of the ten sub-analyses were averaged.

Tolerances shown for means were standard errors. For clarity of presentation of the overnight \( r(t) \) and \( d(t) \), the curves were smoothed using a cubic 119-point Savitzky–Golay filter (Sgolayfilt, Matlab).

3. Results

Both percent recurrence \( (r(t)) \) and percent determinism \( (d(t)) \) computed from the EEGs of patients with mild or moderate OSA (AHI 5–30 events/h) exhibited ultradian rhythms consisting of 2–5 cycles, depending on the patient (Figs. 2 and 3). Across the entire patient cohort, both variables were lowest during wake and N1 and exhibited progressively higher values during deeper sleep, with the REM values located between the N1 and N2 stages (Fig. 4), as expected.

The rate of generalized arousals (an objective measure of EEG fragmentation) was greatest in wake, least in N3 sleep, and intermediate in the other sleep stages, regardless of whether the indices were computed from \( r(t) \) or \( d(t) \) (Figs. 1, 5 and 6). The stage-specific probability of a generalized arousal (for each sleep stage, the number of arousals divided by the maximum possible number) decreased continuously with increasing sleep depth (Fig. 7).

To examine the relationship of specific ABR depth markers to OSA severity, the study cohort was divided into mild and moderate sub-cohorts, based on the AHI values. Because N1 sleep was rare (2.8% of the epochs in the cohort), N1 and N2 epochs were combined for sub-cohort analyses. Sleep depth was greater in all sleep stages in the mild-OSA sub-cohort (Fig. 8), indicating that more severe disease was associated with lighter sleep. The arousal indices similarly depended on OSA severity: they exhibited greater decreases during sleep in the mild-OSA sub-cohort, indicating that the more severe disease was associated with more fragmented sleep (Fig. 9).

Linear discriminant analysis was performed to identify combinations of the markers that reliably classified individual patients with respect to OSA severity (Fig. 1). When all combinations of 2 markers (out of 20) were evaluated, 10 combinations yielded...
an AUROC of 0.75 (15 of 20 patients classified correctly). For example, when only percent recurrence and percent determinism in N3 sleep \(r_{N3}\) and \(d_{N3}\), respectively) were used, 15 patients were correctly classified (Fig. 10, left panel). When 5 markers were included, 18 patients were correctly classified (Fig. 10, middle panel). Inclusion of an additional marker resulted in correct classification of all patients (Fig. 10, right panel); other combinations of six markers also yield 100% correct classification. Similar results were found in the cross-validation studies (data not presented).

To assess the generalizability of the use of ABR markers for accurately predicting OSA severity (Fig. 10, right panel), the same markers were used to classify the SHHS sub-cohorts, and a classification accuracy of 90% was achieved (Fig. 11a). When other combinations of markers were considered, 100% CA was achieved for the SHHS sub-cohorts (Fig. 11b) (see Appendix for an example of a biomarker function). This result was obtained consistently when other random choices of the SHHS sub-cohorts were made (from 103 and 77 patients with mild and moderate OSA, respectively) (data not presented).

4. Discussion

We hypothesized that the severity of OSA was reflected in brain-activity changes, whether as cause or effect, and that such changes could be characterized using algorithmically-determined variables for depth and fragmentation of sleep extracted from the EEG. ABR was used to quantify both phenomena in terms of distinct but related markers, 10 related to sleep depth, and 10 related to sleep fragmentation. The hypothesis was tested at the level of individual patients by asking whether the markers reliably facilitated prediction of whether patients in an OSA cohort (AHI 5–30) had mild (AHI 5–15) or moderate (AHI 16–30) OSA. Various combinations of the markers were systematically evaluated using discriminant analysis to produce marker functions (one marker function per combination), and the ability of each function to...
discriminate between patients with mild or moderate OSA was evaluated using AUROC.

Only a few markers were needed to achieve good classification accuracy (CA) (Fig. 10, left panel). Inclusion of more markers in the discriminant analysis increased CA, and a combination of 6 markers yielded complete CA (Fig. 10, center and right panels).

Several lines of evidence reinforced the validity of the classification success achieved using the 6-marker function. First, the dynamical behavior of the ABR variable (Figs. 2, 3, 5 and 6) was consistent with how they were interpreted. Depth variables were expected to be highest in N3, lowest in wake, and to exhibit intermediate values in the other sleep states. Similarly, the GAI — a measure of the dynamic variance or fragmentation of brain states reflected in the EEG — was expected to exhibit a particular relationship to sleep state (greatest variation in wake, least in N3). These expected correlations were seen at the level of individual patients (Figs. 2, 3, 5 and 6), and as an average property of the cohort (Figs. 4 and 7). Thus the ABR variables (the source of the discriminatory markers) had a sound physiological basis.

Second, the averages of the markers consistently differed between the subcohorts in the expected directions (Figs. 8 and 9). When the contributions to discrimination between the subcohorts provided by individual markers were combined, inclusion of additional markers produced progressively better results, which is the expected observation if the markers were related to disease severity, as we assumed. Thus the discrimination of the markers can be seen to arise from the summation of small differences in individual markers that were jointly but not independently significant at P < 0.05.

Third, the marker function computed from the LSU cohort successfully discriminated patients with OSA who were randomly selected from the SHHS database (Fig. 11). This result was obtained consistently when other random choices of the SHHS subcohorts were made. Because the demographics of the two cohorts differed (Table 1), the results obtained from the SHHS subjects supported the overall conclusion that recurrence in the EEG extracted by ABR reliably coded for OSA severity.

Fourth, links between hypoxia and brain metabolic processes that could explain changes in brain electrical activity related to OSA severity have been identified and could provide a mechanistic understanding of how OSA affects the brain. One possibility is that intermittent hypoxia selectively activates an upregulation of intracellular NFκB, which is responsible for triggering multiple immunologic messengers, including TNFα, IL-6, and IL-1 (Ryan et al.,...
GAIr in N1/N2, and GAIr values are generalized arousal indices computed from only EEG signals reflected in the confluence of global brain activity. The redundancy of this view, all outputs of the brain (any EEG signal) reflect an integrative property of the brain that reflects the global regulatory function of the brain, which includes the response to hypoxia and other physiologic stressors. ABR is based on the assumption that the response to hypoxia and other stressors produces a pattern of brain electrical activity during sleep that reflects the brain’s capacity to maintain normal function in response to stress. The pattern of brain electrical activity during sleep arises from the complexity of brain electrical activity during sleep, and the pattern of brain electrical activity during sleep is a measure of the brain’s capacity to maintain normal function in response to stress.

In addition, intermittent hypoxia in patients with OSA creates an ischemia–reperfusion intracellular environment that promotes the development of reactive oxygen species and oxidative free radicals, all of which can promote cellular destruction (McNichols, 2009). Functional imaging has shown evidence of subcortical damage in patients with OSA (Torelli et al., 2011). All of these processes impact brain metabolism and consequently would be expected to leave footprints in the EEG.

Finally, complete CA was achieved using any of the 6 EEG derivations recorded in a standard overnight sleep study. This observation supports our basic concept that the EEG is an emergent state property of the brain that reflects the global regulatory function of the brain, which includes the response to hypoxia and other physiologic stressors. ABR is based on the assumption that the complexity of brain electrical activity during sleep arises from brain-wide interactions among localized neuronal networks. In this view, all outputs of the brain (any EEG signal) reflect an integrative confluence of global brain activity. The redundancy of the observations from different derivations was not evidence that EEG signals reflected only non-localized information about brain activity, but the redundancy was evidence that a method designed to detect non-localized information in the EEG actually did so.

Although 20 markers were initially defined and available for incorporation in the marker function, fewer were actually needed to achieve complete CA. Considering all analyses performed, including those involving the SHHS subjects, the actual number of markers needed for 100% accuracy was 5–10, depending on the particular cohort. Future efforts to classify OSA patients into more than two classes might require more of the available markers to optimize classification accuracy. Such improvements would be part of a recursive process of development regarding which we have only taken the first step. The potential clinical utility of this approach, if any, is presently unknown.

In summary, patients with OSA (AHI 5–30) could be individually classified as having mild or moderate disease (AHI 5–15, 16–30, respectively) based solely on an algorithmic analysis of the sleep-staged EEG from only one derivation.

Appendix. Biomarker function

Let B be the biomarker function determined by linear discriminant analysis that separates mild and moderate OSA patients in the SHHS cohort (Fig. 11b). B = a1V1 + a2V2 + a3V3 + a4V4 + a5V5, where V1 = rN1/N2, V2 = dWASO, V3 = dREM, V4 = GAIr in N1/N2, and V5 = GAIr in REM. The respective coefficients a1–a5 are 0.333, −0.006, −0.010, −0.014, and −0.020. The threshold for discrimination between the groups (chosen using AUROC) was 0.38.

References


