2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Mapping Sleep Spindle Characteristics to Vigilance Outcomes in Patients with Obstructive Sleep Apnea

K. McCloy, B. Duce, C. Hukins, and U. Abeyratne, *Member, IEEE*

Abstract— **Obstructive Sleep Apnea (OSA) is a sleep disorder associated with reduced vigilance. Vigilance status is often measured using the Psychomotor Vigilance Task (PVT). This paper investigates modelling strategies to map sleep spindle (Sp) characteristics to PVT metrics in patients with OSA. Sleep spindles (n=2305) were manually detected across blocks of sleep for 20 patients randomly selected from a cohort of 190 undergoing Polysomnography (PSG) for suspected OSA. Novel Sp metrics based on runs or "bursts" of Sps were used to model Sp characteristics to standardized (z) Lapse and Median Reaction Time (MdRT) scores, and to Groups based on zLapse and zMdRT scores. A model employing Sp Burst characteristics mapped to MdRT Group membership with an accuracy of 91.9%, (95% C.I. 90.8-93.0). The model had a sensitivity of 88.9%, (95% C.I. 87.5-89.0) and specificity of 89.1% (95% C.I. 87.3-90.5) for detecting patients with the lowest MdRTs in our cohort.**

***Clinical Relevance*— Based on these results it may be possible to use Sp data collected during overnight diagnostic PSG for OSA to detect patients at risk for attention deficits. This would improve triage for OSA therapy by identifying at risk patients at the time of OSA diagnosis and would remove the need to employ additional testing to assess vigilance status.**

I. Introduction

Sleep Spindles (Sp) are a defining element of NREM (Stage N2) sleep. Patients with untreated Obstructive Sleep Apnea (OSA) have abnormal Sp characteristics [1], chronic sleep restriction and a high risk for altered vigilance [2]. The Psychomotor Vigilance Task (PVT) measures changes in vigilance due to sleep loss [3]. Due to staffing and time constraints, the PVT is not conducted clinically, and normative population data for healthy and patient-populations is lacking. If Sp characteristics measured during diagnostic Polysomnography (PSG) for OSA can be mapped to PVT outcomes in an OSA cohort, then both OSA and attentional loss could be detected at a population level and during a single procedure in this at-risk group.

The PVT measures the reaction time (RT) to a visual or auditory stimulus. Lapses (a RT≥500msec) are the most commonly reported element of the PVT [3]. Lapse numbers increase with the time spent at a task, and after both acute and chronic sleep restriction (as in OSA) [3]. They are subject to circadian effects, and are highest in the morning and lowest in the evening [4]. Lapse numbers have been correlated to poor performance in diverse activities including working memory [5] and driving simulation tasks [6], but they occur interspersed with faster RTs. The Median RT (MdRT) may indicate the predominance of slowing of RTs.

A Sp occurs during bursts of Thalamocortical activity in Stage N sleep with a refractory period of 5-20seconds [7]. Sp density[8] and in particular the density of Fast Sp (FSp,13-16Hz)[9] is correlated to memory performance. Sleep restriction results in a decrease in memory performance and an increase in Lapses, but the two are not thought to be related[10]. Neither Sp characteristics [1], PVT outcomes or memory [11]are improved reliably after treatment of OSA with Continuous Positive Airway Pressure (CPAP).

A recent pilot study (n=7) found correlations between reduced Stage N Sp density, and increased lapses after acute sleep deprivation in patients with severe OSA[12]. Sp density is a frequently used metric of Sp performance, but does not directly measure the functional burst nature of Sps. Little is known about interactions between Sp Bursts, Sp density and lapses or slowing of RTs in chronic sleep restriction or across different times of the sleep period in patients with OSA.

This study seeks to investigate relationships between Sp metrics, Lapses and the MdRT in patients undergoing diagnostic PSG for suspected OSA, and to determine if Sp burst Characteristics can be mapped to RT outcomes using data acquired from Diagnostic PSG. If successful diagnosis of OSA and vigilance deficits could be performed with a single night of PSG, patients with vigilance deficits could be identified and triaged to receive sleep and lifestyle therapies.

II. Materials and Methods

A. Data Gathering

Data for this retrospective study was gathered from a cohort undergoing PSG for suspected OSA at a tertiary sleep center using the Compumedics Grael acquisition system (Abbotsford, Australia). This study was under the supervision of the Human Research Ethics Committees of the Metro South Hospital and Health Services, and the University of Queensland. Scoring of the PSG was according to the AASM 2012 rules [13] by qualified sleep technicians who regularly undergo reliability testing. The PSG recorded 12 data channels

*Research supported by the National Health and Medical Research Council (NHMRC) under grant APP1126600.

K. McCloy is with the Department of Human Centred Computing Queensland University, Australia 4072 (phone: 614-427-969-055; fax:617-956-012; e-mail: k.mccloy@uqconnect.edu.au. B. Duce is with the Department of Respiratory and Sleep medicine Princess Alexandra Hospital Brisbane; e-mail: brett.duce@health.qld.gov.au. C. Hukins is with the Department of Respiratory and Sleep medicine Princess Alexandra Hospital Brisbane; e-mail: craig.hukins@health.qld.gov.au U. Abeyratne is with the Department of Human Centred Computing Queensland University, Australia 4072; e-mail: udantha@itee.uq.edu.au

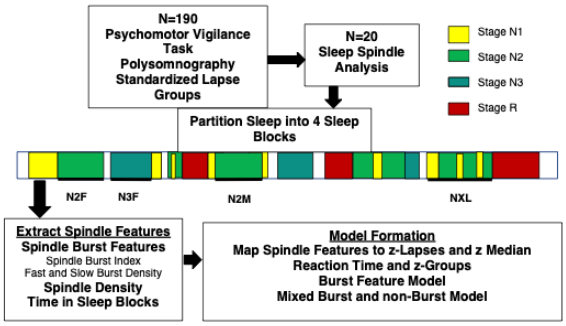

Figure 1. Study Methods

including electroencephalography (EEG). The recommended recording montage used was F4:M1, C4:M1, and 02:M1.

Only patients with complete data sets were included. Exclusions were made due to missing demographic data (n=7) and PVT results (n=5), leaving 190 patients with complete data sets. A Sp cohort (SC) was randomly selected from within the full 190 patients, leaving a Main Cohort (MC) of n=170. Sleep and demographic data were compared between the MC and the SC to monitor how closely the SC represented the larger sleep population. The methodology for this study is shown in Fig. 1.

B. *Psychomotor Vigilance Test*

During the ten-minute PVT, 120 red crosses are displayed on a computer screen with random interstimulus intervals between 2-10 seconds. Participants were asked to push a button when they saw a cross. To create Groups based on PVT results, Lapse and MdRT data were extracted from the PVT data and standardized to the mean so that:

$$Z(\propto) = \frac{\propto - \mu_{\propto}}{\sigma_{\propto}}$$

where $\propto$ = the number of lapses or the MdRT and $\mu_{\propto}$ and $\sigma_{\propto}$ represent the population mean and standard deviation of the number of lapses or the MdRT. Four PVT severity groups were also established so that: Group 1: $Z_{(\propto)} < \mu_{\propto} - \sigma_{\propto}$, Group 2: $\mu_{(\propto)} > Z_{\propto} \geq \mu_{\propto} - \sigma_{\propto}$, Group 3: $\mu_{\propto} < Z_{(\propto)} \leq \mu_{\propto} + \sigma_{\propto}$ and Group 4: $Z_{(\propto)} > \mu_{\propto} + \sigma_{\propto}$.

C. *Sleep Blocks*

The sleep period was divided into four Sleep Blocks (SB) of Stage N sleep (Fig.1) using the epoch scoring from the PSG. The time spent in each SB was recorded. Each SB consisted of a minimum of four epochs of sleep of a defined sleep stage. The SB was ended when more than four epochs with different sleep staging were scored. The N2F SB was the first SB of Stage N2 sleep after sleep onset. Similarly, N3F was the 1st SB of Stage N3 sleep. N2M was the Stage N2 SB after the first Stage R epoch, or from the middle of sleep if Stage R was absent or delayed past the first half of the sleep period. The last scored Stage N SB prior to the end of the sleep period (NXL) was a mixed Stage N (N1, N2 and N3) SB. All Sp feature extraction was done from within these four SBs for each subject.

D. *Sleep Spindle Feature Extraction*

The EEG was collected at 1024 samples/second and then band–pass filtered between 0.1-35Hz. Sps were detected manually across electrodes F4-M1, C3-M2, C4-M2 and O2-M1 during sleep that was visually identified as free of artifacts. A 20 second screen was used to score Sp in Profusion 4 (Build 456, Compumedics Ltd, Abbotsford Australia) software. Sp were classified with a minimum duration of 0.3sec, minimum amplitude of 10μV and frequencies of 11-16Hz [14]. Wave forms with a Frequency lower than12Hz were identified as Sp only if they were part of a train of at least 2 Sps [15].

We classified Sps as burst or non-burst Sps. Burst Sp had a maximum inter-Sp-Interval of 20 seconds [7]. A novel measure, the *Spindle Burst Index (SBI),* was defined as the number of burst Sp in a SB as a percentage of the total number of Sp in the SB. For each SB we also computed the number and density of burst fast SSp (FSP) and slow Sp (SSP), and the density of all Sp within the SB.

E. *Model Construction*

We constructed four models using the MATLAB Regression Learner and Classification Learner Applications (MATLAB R2018b, The Mathworks Inc., Natick MA) to map Sp features to i) individual z-Lapse and zMdRT scores, and ii) Lapse and MdRT Groups. For the z-score models, predictors were chosen based on reduction in the Root Mean Square Error (RMSE). For the Group models, predictor selection was based on accuracy. The number of predictors was limited to four to prevent over fitting. The RMSE, r-squared data, Mean Squared Error (MSE) and Mean Adjusted Error (MAE) for z-score models and the accuracy, sensitivity, specificity, and positive and negative predictive values (for Group data) were reported on validation folds after fifty repeats of randomised five-fold cross validation.

Models were constructed based on Sp Burst metrics (SPBM) and mixed burst and non-burst metrics (SBM). Sp characteristics were mapped to zLapse and zMdRT values, and to Lapse and MdRT Groups. Algorithms were chosen from the nonparametric selections in MATLAB Regression Learner Application (for zLapse and zMdRT scores) and the Classification Learner Application for Groups.

COMPARISON OF DEMOGRAPHIC AND SLEEP DATA FOR THE MAIN COHORT AND THE SPNDLE COHORT

Sleep and Demographic Variables	MC (n=170)	SC (n=20)	p
Age	54.6±13.7	57.0±14.8	0.431
Gender F:M	62:108	12:8	0.044*
BMI	36.6±9.2	36.5±7.5	0.847
ESS[a]	9.6±5.6	11.4±4.6	0.109
Sleep Latency	23.2±23.2	22.8±14.4	0.345
AHI[b]	32.3±30.3	24.9±23.9	0.276
Lapses	22.1±27.2	36.9±39.8	0.3
Median RT[c]	428±326.0	511.7±195.6	0.200

[a]=Epworth Sleepiness Scale,[b]= Apnea Hypopnea Index, c=Reaction Time, *=p<0.05

TABLE I. COMPARISON OF SPINDLE MODELLING VARIABES ACROSS THE Z-LAPSE AND Z-MEDIAN GROUPS IN THE SPINDLE COHORT

Spindle Variables	Lapse Groups(number)/mean±S.D. LG2(9)	LG3(5)	LG4(6)	p
Lapses	4.1∓3.0	36.2∓11.1	86.7∓31.8	<.001*
SBI N2F	90.6±11.2	74.0±25.5	47.6±42.6	0.110
Time in N3F	24.3±12.7	21.8±16.4	5.7±3.3	0.017*
B Den SSp N2F	0.6±0.3	1.2±1.1	1.3±1.6	0.851
nBFSp:nBSSp[a]	2.3±1.6	2.1±1.4	0.6±0.5	0.023*
nFSp:nSSp	2.2±1.4	2.0±1.4	0.5±0.2	0.045*
nBursts N2F	8.4±3.9	5.2±5.7	4.0±4.9	0.104
n all Sp	156.2±101.6	114.0±79.4	54.8±71.1	0.082
	MdG2(11)	**MdG3(6)**	**MdG4(3)**	
Median RT	352.5±452.5	475.5±23.2	1168.0±173	<.001*
SBI N2F	92.1±10.5	69.6±21.3	13.3±23.1	0.006*
Time in N3F	21.9±13.0	17.0±17.0	6.2±4.9	0.141
B Den SSp N2F	0.8±0.6	1.6±1.5	0.0±0.1	0.042*
nBFSp:nBSSp[a]	2.2±1.5	1.5±1.6	0.7±0.6	0.186
nFSp:nSSp	2.1±1.4	1.5±1.6	0.5±0.1	0.031*
Sp Density N3F	1.4±1.3	1.3±1.0	0.5±0.5	0.308
n all Sp	147.5±97.2	109.0±77.8	9.7±9.9	0.026*

[a]=number of Burst Fast Sp: Number of Burst slow Sp, *=p<0.05

TABLE III. MAPPING SPINDLE BURST CHARACCTERISTICS TO MEDIAN REACTION TIME GROUPS

SPBM Fine K-Nearest Neighbours	Median Reaction Time Groups %, (95%Confidence Interval) MdG2 (n=11)	MdG3 (n=6)	MdG4 (n=3)
Sensitivity	88.9 (87.3-90.5)	96.3 (93.7-98.9)	94.0 (90.3-97.7)
Specificity	98.2 99.6-99.8)	90.0 (88.6-91.4)	99.6 (99.2-100.0)
Positive Predictive Value	88.2 (86.8-89.5)	98.5 (97.4-99.5)	99.0 (98.4-99.6)
Negative Predictive Value	98.7 (97.4-99.5)	81.1Ω (79.2z82.z)	98.3 (96.4-100.0)
Accuracy	91.9 (90.8-93.0)		

Model: SBI N2F; time N3F; Burst Density SSp N2F; Ratio Burst FSp:SSp[a]

[a]=number of Burst Fast Sp: Number of Burst slow Sp.

F. Statistics

Statistical significance was set at p<0.05. The Kruskal Wallis One-way Analysis of Variance compared distributions for continuous variables using Dunn's pairwise tests and the Bonferroni correction. The Chi-Square test was used for categorical variables. Statistical analysis was done in SPSS version 26 (IBM corp., New York, USA).

III. RESULTS

After random selection of the SC, there were no significant differences for sleep or demographic characteristics between the MC and the SC except for gender (p=0.043) (Table 1). The SC had a higher number of females with less severe OSA, than the MC, but the OSA differences did not reach statistical significance. The SC contained 3 Lapse and MdRT Groups. There were no subjects in Group1 in either the SC or the MC. The SC had a higher number of Lapses and a higher MdRT than the MC, but this did not reach statistical significance.

A. Spindle Features

There were significant differences between Groups for the time spent in N3F, with Lapse Group 4 (LG4) having the lowest values. The SBI differed significantly between the MdRT Groups (MdRTG2-MdRTG4) and decreased from MdRTG2-MdRTG4. SBI values for each group were highest in N2F. The ratio of all fast: slow Sp decreased significantly from Lapse Group 2-4 (LG2-LG4) and MdRTG2-G4.

B. Modelling Outcomes

Modelling outcomes mapping Sp features to zLapse and zMdRT scores using Exponential Gaussian Process Regression are shown in Table 2. All models used the time spent in N3F, the N2F SBI and the burst density of Slow Sp in N2F. The zMdRT SBM contained the Sp density of all Sps in N3F and had the lowest RMSE (0.59, 95% C.I. 0.56-0.61). The highest r-squared value was in the zLapse SPBM which included the SBI NXL (0.75, 95% C.I. 0.73-0.77).

The K Nearest Neighbors algorithm was used to map Sp characteristics to the MdRT (Table 3) and Lapse Groups (Table 4). The SPBM mapped to the MdRT groups with an accuracy of 91.9% (95% C.I. 90.8-93.0%). It had a sensitivity of 88.9% (95%C.I. 87.3-90.5) for MdRT Group2 with the fastest RT, and sensitivity of 94.0% (95% C.I. 90.3-97.7) for mapping to MdRT Group4 with the slowest RTs, indicating that burst Sp metrics are useful for mapping to slowing of RTs in an OSA cohort. The best-performed Lapse Group model was the SBM (Table 4) with an

TABLE II. MAPPING SPINDLE FEATURES TO Z-LAPSE AND Z-MEDIAN REACTION TIMES

		RMSE[a]	R-Squared	MSE[b]	MAE[c]
		Number (95% Confidence Interval)			
SPBM[d]	zMdRT	0.61 (0.58-0.63)	0.67 (0.64-0.70)	0.38 (0.34-0.41)	0.53 (0.42-0.44)
	zL	0.73 (0.70-0.77)	0.75 (0.73-0.77)	0.55 (0.50-0.61)	0.55 (0.52-0.57)
SBM[e]	zMdRT	0.59 (0.56-0.61)	0.69 (0.66-0.72)	0.35 (0.32-0.39)	0.42 (0.41-0.43)
	zL	0.82 (0.79-0.85)	0.69 (0.67-0.72)	0.68 (0.63-0.74)	0.61 (0.59-0.63)

A=Root mean square error, b=Mean squared error, c=Mean absolute Error ,d=Spindle Burst Model, e=Mixed Burst/ non-burst model

TABLE IV. MAPPINZ MIXED SPNDLE BURST AND DENSITY METRICS TO LAPSE GROUPS

SBM Fine K-Nearest Neighbours	Lapse Groups %, (95%Confidence Interval) Group2 (n=9)	Group3 (n=5)	Group4 (n=6)
Sensitivity	88.9 (88.9-88.9)	57.1 (54.0-60.2)	62.3 (58.9-65.8)
Specificity	91.2 (89.8-92.6)	78.7 (77.2-80.1)	91.2 (90.2-92.2)
Positive Predictive Value	90.8 (90.7-91.0)	84.8 (83.8-85.9)	85.1 (84.0-86.3)
Negative Predictive Value	89.6 (88.2-90.9)	46.9 (44.8-49.1)	75.8 (73.8-77.7)
Accuracy		72.9 (71.8-74.0)	

SBI[f] N2F[g]; time N3F; Burst Density Slow Sp N2F; n All FSp:All SSp

accuracy of 72.9%, (95%C.I. 71.8-74.0). The Sp burst features from the N2F SB mapped accurately to changes in RT across the three MdRT Groups. Features from the N2F SB mapped accurately to changes in RT across the three MdRT Groups.

IV. Discussion

This study has shown that it is possible to use Sp characteristics obtained from overnight diagnostic PSG to map to z-scores and to Groups formed from PVT lapse and MdRT outcomes in a cohort of patients being tested for OSA. Population research has previously used standardized data as the basis for comparison of biomedical data across different clinical groups and different time periods [16]. The highest modelling accuracy in this study was obtained when Sp features were mapped to Groups with one or more standard deviations in the MdRT between groups, rather than to individual z-scores. These methods may form a basis within our clinical population to monitor changes between groups across time.

The functional impact of the Burst nature of Sps has been largely unexplored. Although runs of Sp with 3-6 sec inter-spindle intervals have been correlated to manipulations such as memory re-cuing during sleep [17], Sp with a 3-6 second periodicity were found to be largely absent in both healthy controls and patients with OSA across single electrode couplings in un-cued nocturnal sleep [1]. In contrast we have mapped Sp burst features with a maximal inter-spindle interval of 20sec across multiple electrodes to MdRT group data. There may be value in investigating the efficacy of modelling outcomes using different burst interval definitions, and different electrode arrays across different periods of sleep for predicting vigilance outcomes.

The time spent in the N3F SB was significantly lower in the L4 Lapse Group, and was incorporated in all of the constructed models. Previous work has found an increase in lapses in healthy controls after selective interruption of Slow-wave activity during recovery sleep using both visual [18] and auditory stimuli [19]. Reduced Stage N3 sleep is also associated with increased severity of OSA [20]. As there are complex interactions between OSA pathology, sleep quality and OSA metrics, there is likely to be a multifactorial association between altered Sp metrics and other sleep-related data.

The highest model accuracy was obtained using Sp burst features from the N2F SB to predict zMdRT Group membership. Future work could investigate the predictive value of different Burst and non-Burst Sp characteristics in larger populations that include healthy controls and incorporate automatic Sp detection. Extra data to be collected in future work could include information about educational level, dominant hand, medications and neurologic conditions.

Acknowledgments

We would like to thank the patients and the dedicated staff at the Princess Alexandra Hospital in Brisbane, Australia

References

[1] T. Saunamäki *et al*, "CPAP treatment partly normalizes sleep spindle features in obstructive sleep apnea," *Sleep Disorders,* vol. 2017, 2017.
[2] M. Olaithe *et al*, "Cognitive deficits in obstructive sleep apnea: insights from a meta-review and comparison with deficits observed in COPD, insomnia, and sleep deprivation," *Sleep Medicine Reviews,* vol. 38, pp. 39-49, 2018.
[3] M. Basner and D. F. Dinges, "Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss," *Sleep,* vol. 34, *(5),* pp. 581-591, 2011.
[4] D. J. Mollicone *et al*, "Time of day effects on neurobehavioral performance during chronic sleep restriction," *Aviat. Space Environ. Med.,* vol. 81, *(8),* pp. 735-744, 2010.
[5] N. Unsworth, A. L. Miller and M. K. Robison, "Individual differences in lapses of sustained attention: Ocolumetric indicators of intrinsic alertness." *Journal of Experimental Psychology: Human Perception and Performance,* vol. 46, *(6),* pp. 569, 2020.
[6] M. L. Jackson *et al*, "Cognitive components of simulated driving performance: Sleep loss effects and predictors," *Accident Analysis & Prevention,* vol. 50, pp. 438-444, 2013.
[7] A. Lüthi and D. A. McCormick, "Periodicity of thalamic synchronized oscillations: the role of Ca2 -mediated upregulation of Ih," *Neuron,* vol. 20, *(3),* pp. 553-563, 1998.
[8] S. Gais *et al*, "Learning-dependent increases in sleep spindle density," *J. Neurosci.,* vol. 22, *(15),* pp. 6830-6834, Aug 1, 2002.
[9] M. Mölle *et al*, "Fast and slow spindles during the sleep slow oscillation: disparate coalescence and engagement in memory processing," *Sleep,* vol. 34, *(10),* pp. 1411-1421, 2011.
[10] J. N. Cousins, K. Sasmita and M. W. Chee, "Memory encoding is impaired after multiple nights of partial sleep restriction," *J. Sleep Res.,* vol. 27, *(1),* pp. 138-145, 2018.
[11] S. P. Patil *et al*, "Treatment of adult obstructive sleep apnea with positive airway pressure: an American Academy of Sleep Medicine systematic review, meta-analysis, and GRADE assessment," *Journal of Clinical Sleep Medicine,* vol. 15, *(2),* pp. 301-334, 2019.
[12] A. E. Mullins *et al*, "Sleep EEG microstructure is associated with neurobehavioural impairment after extended wakefulness in obstructive sleep apnea," *Sleep and Breathing,* pp. 1-8, 2020.
[13] Rules *for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events: Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine*. Available:
[14] S. M. Purcell *et al*, "Characterizing sleep spindles in 11,630 individuals from the National Sleep Research Resource," *Nature Communications,* vol. 8, pp. ncomms15930, 2017.
[15] A. Rodenbeck *et al*, "A review of sleep EEG patterns. Part I: a compilation of amended rules for their visual recognition according to Rechtschaffen and Kales," *Somnologie,* vol. 10, *(4),* pp. 159-175, 2006.
[16] S. Isanaka *et al*, "Assessing the impact of the introduction of the World Health Organization growth standards and weight-for-height z-score criterion on the response to treatment of severe acute malnutrition in children: secondary data analysis," *Pediatrics,* vol. 123, *(1),* pp. 54, Jan, 2009.
[17] J. W. Antony *et al*, "Sleep spindle refractoriness segregates periods of memory reactivation," *Current Biology,* vol. 28, *(11),* pp. 1736-1743. e4, 2018.
[18] Van Der Werf, Ysbrand D *et al*, "Reduction of nocturnal slow-wave activity affects daytime vigilance lapses and memory encoding but not reaction time or implicit learning," in *Progress in Brain Research*Anonymous Elsevier, 2011, pp. 245-255.
[19] M. Ferrara, L. De Gennaro and M. Bertini, "The effects of slow-wave sleep (SWS) deprivation and time of night on behavioral performance upon awakening," *Physiol. Behav.,* vol. 68, *(1-2),* pp. 55-61, 1999.
[20] R. Ratnavadivel *et al*, "Marked reduction in obstructive sleep apnea severity in slow wave sleep," *Journal of Clinical Sleep Medicine,* vol. 5, *(6),* pp. 519-524, 2009.