

Accuracy of the Oxford Sleep Resistance Test versus Simultaneous Electroencephalography to Detect Sleep Onset

Vincent Jobin^{1*}, Annie Mathieu¹, Pierre Rompré², Mélanie Perraton-Brillon³, Geneviève Rondeau¹ and Pierre Mayer¹

¹Department of Medicine, Centre hospitalier de l'Université de Montréal (CHUM), Canada

²Faculty of Dental Medicine, Université de Montréal, Canada

³Department of Pneumology, Hôpital Maisonneuve-Rosemont, Canada

*Corresponding author: Dr Vincent Jobin, Centre hospitalier de l'Université de Montréal (CHUM)-3840, St-Urbain Street, Montreal, QC, Canada, Tel: 514 890-8000; Fax: 5144127123; E-mail: vincent.jobin@umontreal.ca

Received date: January 27, 2017; Accepted date: February 16, 2017; Published date: February 28, 2017

Copyright: © 2017 Jobin V, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Background: The Oxford Sleep Resistance test (OSLER) is a useful tool to assess daytime vigilance. However, it has not been validated against simultaneous electroencephalography (EEG) recordings in large populations. The main objective of the study was to compare the OSLER values versus EEG-determined Sleep Onset latency (EEG-SOL).

Methods: Patients referred for assessment of daytime vigilance were recruited from a tertiary sleep clinic. Patients underwent the OSLER (4 x 40 minutes trials; if 7 consecutive stimuli are missed, the trial is terminated and sleep onset is concluded to have occurred) with simultaneous EEG recordings. Determination of EEG-SOL using American Academy of sleep Medicine (AASM) criteria to score sleep during daytime testing was compared to OSLER values.

Results: 65 OSLER were performed in 65 subjects for a total of 260 trials (65X4 trials/OSLER). In 136 out of the 260 trials (52.3%), subjects remained awake according to the OSLER, while EEG-SOL was scored in 5 of the 136 trials (3.7%). Of the 124 trials (47.7%) with sleep onset, (i.e. 7 consecutive missed stimuli) the mean sleep onset value was 14.5 ± 10.9 min and EEG-SOL was recorded before the end of the trial in 37 trials (29.8%) (Mean difference EEG-SOL vs. OSLER 4.1 ± 5.8 min).

Conclusion: Using current AASM criteria for daytime testing, EEG-determined sleep onset latency is unlikely to occur in subjects with no sleep onset in the OSLER. However, the presence of sleep onset in the OSLER cannot be used as a precise surrogate to detect EEG sleep onset.

Keywords: Sleep; Sleep apnea; Wakefulness; Continuous positive airway pressure; Electroencephalography; Polysomnography

Abbreviations: AASM: American Academy of Sleep Medicine; AHI: Apnea-Hypopnea Index; APAP: Automatically-adjusting Positive Airway Pressure; AUC: Area Under the Curve; BiPAP: Bilevel Positive Airway Pressure; BMI: Body Mass Index; CI: Confidence Interval; CPAP: Continuous Positive Airway Pressure; EEG: Electroencephalography; EEG-SOL: EEG-determined Sleep Onset Latency; EMG: Chin Electromyogram; EOG: Bilateral Electrooculogram; ESS: Epworth Sleepiness Score; LED: Light Emitting Diode; MSLT: Multiple Sleep Latency Test; MWT: Maintenance of Wakefulness Test; ODI: Oxygen Desaturation Index; OSA: Obstructive Sleep Apnea; OSLER: Oxford Sleep Resistance Test; PSG: Polysomnography; RDI: Respiratory Disturbance Index; ROC: Receiver Operating Characteristic; SOL: Sleep Onset Latency

Introduction

The current gold standard to measure the ability to stay awake is the Maintenance of Wakefulness Test (MWT) [1,2]; which requires electroencephalography (EEG) monitoring and is costly. The Oxford

Sleep Resistance test (OSLER) is a useful alternative [3]. It is a performance test where the subject is asked to respond to a Light Emitting Diode (LED), a monotonous visual stimulus occurring every three seconds. The OSLER test is not based on EEG and estimates the Sleep Onset Latency (SOL) by the time taken by the subject to fail to respond to seven consecutive visual stimuli. Compared to the MWT, this test is viewed as a simple and inexpensive; not requiring any specialized staff.

The OSLER was first used with 4 x 40 min trials and was first validated to determine sleep onset during daytime studies with non-simultaneous EEG monitoring [3]. The OSLER discriminates well untreated sleepy patients with obstructive sleep apnea (OSA) from normal controls [3]. Using different protocols (1 to 4 trials of 20 to 40 min in length), the OSLER test has subsequently proved to be a useful test, especially in OSA patients, to differentiate normality from hypersomnia as well as to demonstrate an effect of therapy [4-10]. However, little is known about its accuracy when compared to simultaneous EEG measurements of sleep onset in a large cohort of patients using current AASM 2005 definition of sleep onset in daytime studies (MWT, MSLT) versus the previous definition used when the test was first described [2,3,11]. Also, there is debate in the literature in order to determine if an OSLER test consisting of less than 4 trials can

be done with similar accuracy in order to simplify the test procedure [3,4].

We hypothesize that in a cohort of patients referred for assessment of daytime vigilance in a tertiary care university hospital: i) the OSLER test is not accurate to detect sleep onset using current AASM criteria defining sleep onset after one epoch of any stage of sleep and ii) the results of the four-trial OSLER test can be estimated by less than four trials. Therefore, the objectives of our study are to compare a SOL obtained using the OSLER *vs.* one obtained by simultaneous EEG monitoring and, to compare the mean OSLER value using 1, 2 or 3 trials with the mean value of the standard four-trial OSLER performed during the day.

Methods

Patients

Patients were recruited after an evaluation at our sleep clinic. We included all consecutive patients who were underwent a MWT after being referred for the test by a sleep specialist, including patients who were referred for a second evaluation/opinion. Exclusion criteria included poor comprehension of French or English and the presence of a physical disability that would make the OSLER test unreliable. Research Ethics Board approved the study protocol. All patients gave a written informed consent.

Experimental protocol

Some patients underwent either a full in-lab polysomnography (PSG) or a home recording, with or without treatment to control their condition (medications or positive pressure therapy). Charts were reviewed in order to obtain an Apnea-Hypopnea Index (AHI) from PSG (level 1 to 3 studies) or an Oxygen Desaturation Index (ODI level 4 study) according to the AASM classification [12]. Since some PSG studies were not done at our laboratory, scoring criteria for respiratory events could differ between studies and are here after referred to as a respiratory disturbance index.

Before the study, some patients answered a standardized questionnaire including the Epworth Sleepiness Score (ESS). They were asked to list all medication, caffeine and alcohol consumption, comorbidities, and sleep hygiene.

Measurements

OSLER trials: OSLER trials (Stowood Scientific Instruments, Oxford, UK) were set at 8:00 am, 10:00 am, 12:00 pm and 2:00 pm through 4 x 40 min duration trials. Patients were seated in a comfortable supportive armchair in a quiet sleep-recording room with low-level illumination.

Patients were instructed to touch the switch of a box, set on their thigh, with the index of their dominant hand, each time they saw a LED; which was delivered for 1 second every 3 seconds at eye's height and 2 meters from the head. Each time the patient missed a LED, a noise was emitted to the technician (seated in a different room) and a visual marker was also seen on a computer screen adjacent to EEG tracings. When 7 consecutive stimuli were missed, a different sound was heard and the trial was terminated. Each trial ended at a maximum of 40 min or when 7 consecutive omissions occurred. Patients were asked to remain in the chair and resist sleep without using extraordinary measures. They were also informed that if the LED

stopped, they should put the response box away, remain seated and try to stay awake until the technician comes in the room. Between trials, patients were asked to take a walk in order to minimize fatigue and sleepiness and maintain motivation. The same team of respiratory therapist conducted all tests.

EEG measurement: EEG was monitored simultaneously during OSLER trials using C3/A2, C4/A1, O1/A2, O2/A1 along with chin electromyogram (EMG) and bilateral electrooculogram (EOG) using a Sandman Software recording system.

Data analysis

OSLER: To determine the sleep latency for one 40 min trial, we measured the time before the occurrence of 7 consecutive diodes without response (21 seconds) [3]. This criterion was chosen because it approximates the sleep duration generally used to score one epoch of sleep with Rechtschaffen and Kales scoring rules for an overnight PSG and was the criterion used when the test was first described [3,13]. A session lasting less than 40 min is classified as presence of sleep and SOL is measured [3].

EEG-determined Sleep Onset latency (EEG-SOL): EEG-SOL was defined as single 30 seconds epochs of stage 1, 2 or 3. If no sleep occurred during the OSLER trial, no EEG-SOL was noted. To determine the number and timing of OSLER trials required to best estimate the results of four-trial OSLER; we compared mean results of all possibilities: 1, 2, 3 or 4 trials done at 8:00 am, 10:00 am, 12:00 pm and 2:00 pm comparing each trials independently and with any combination of them.

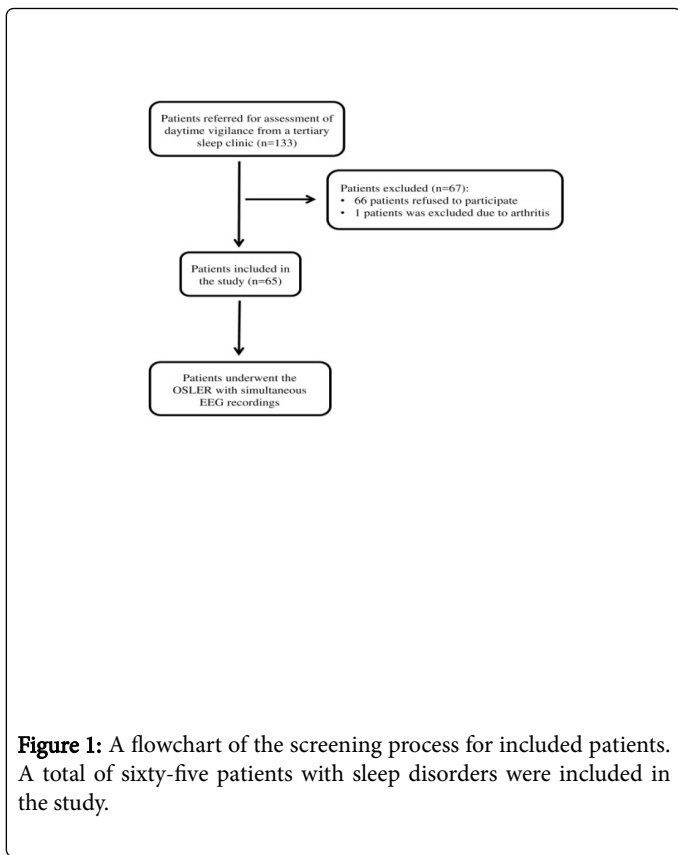
Statistical analysis

In addition to the behavioral measurement of the sleep latency obtained by the OSLER test, EEG-SOL was calculated for each patient and each trial. Values were not normally distributed and Spearman's rank correlation was used for correlation. The mean bias and variability between methods for trials in which EEG-SOL was measured were compared and displayed as a Bland-Altman plot. Lastly, to determine the number of OSLER trials needed to accurately estimate the mean value of four-trial OSLER test, we analyzed the variance between the trials by using the Wilcoxon signed-rank test and to expressed results for all possibilities (1, 2 or 3 trials); we used Bland-Altman plots and receiver operating characteristic (ROC) curves analysis. Statistical analyses were performed using the SPSS 17.0 statistical software package (SPSS Inc. Released 2007; SPSS for Windows, version 16.0, Chicago, USA). Values are expressed as means \pm SD or 95% confidence interval (CI) and medians (interquartile range). A threshold of $p < 0.05$ was used for statistical significance.

Results

Patients

Sixty-five consecutive patients (57 men and 8 women, 50.5 years \pm 10.8 years) with sleep disorders were included in the study. Clinical diagnoses of patients were: sleep disordered breathing (n=63), idiopathic hypersomnolence concomitant with sleep apnea (n=4), restless leg syndrome (n=2) and narcolepsy (n=1). For one patient, we were not able to retrieve the initial diagnosis. Five patients had more than one sleep disorder. (Figure 1).



Seven patients were taking stimulants on a regular basis and 62 patients used positive pressure therapy regularly (Continuous Positive Airway Pressure (CPAP), Automatically-adjusting Positive Airway Pressure (APAP) or Bilevel Positive Airway Pressure (BPAP)). For OSA patients, the mean Respiratory Disturbance Index (RDI) was 38.9 ± 26.9 events/hour ($n=59$). Baseline characteristics of the patients are presented in Table 1 (data were not available for all patients)

Measures of agreement between OSLER and simultaneous EEG

Analysis of OSLER: A total of 65 OSLER were done. The overall mean value for the OSLER was 27.8 ± 12 min. In 19 OSLER (29.2%), patients remained awake (four 40 min trials). In these 19 OSLER, simultaneous EEG-determined sleep was not scored in any but one test (5.2%).

Analysis of individual OSLER trials: In 136 (52.3%) of the 260 OSLER trials performed, while no sleep onset was detected on the OSLER trials, EEG-SOL was recorded using AASM 2005 criteria in 5 trials (3.7%) (Mean difference EEG-SOL vs. OSLER $23.8 \text{ min} \pm 12.3$ min). Of the 124 OSLER trials (Mean value $14.5 \text{ min} \pm 10.9$ min) where sleep onset occurred as defined by failure of the patient to respond to 7 consecutive stimuli, sleep onset by EEG monitoring was found in 37 trials (29.8%) (Mean difference EEG-SOL vs. OSLER $4.1 \text{ min} \pm 5.8$ min).

Figure 2 shows the results of OSLER trials in which EEG-determined sleep was scored. Note that sleep was scored before the OSLER trial was terminated in a significant proportion of patients.

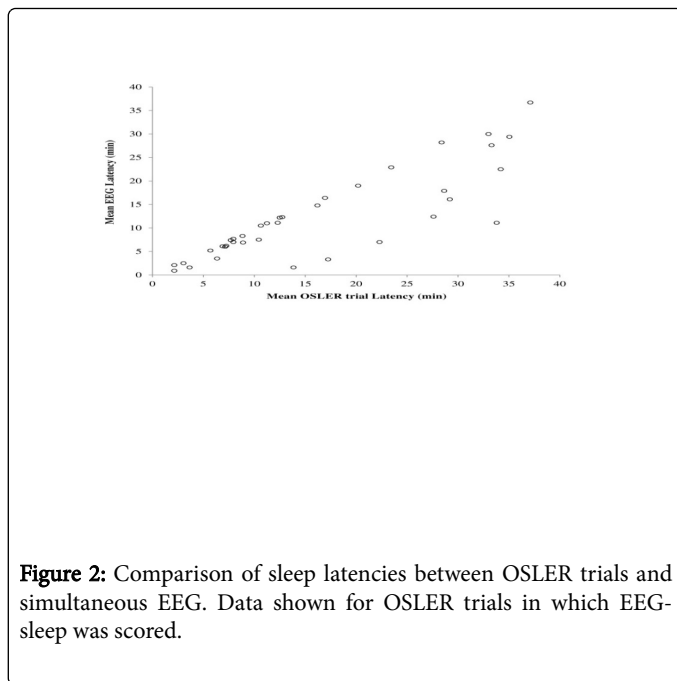


Figure 2: Comparison of sleep latencies between OSLER trials and simultaneous EEG. Data shown for OSLER trials in which EEG-sleep was scored.

In order to compare EEG-SOL with OSLER, a Bland-Altman plot was used in OSLER trials in which EEG sleep was scored (Figure 3).

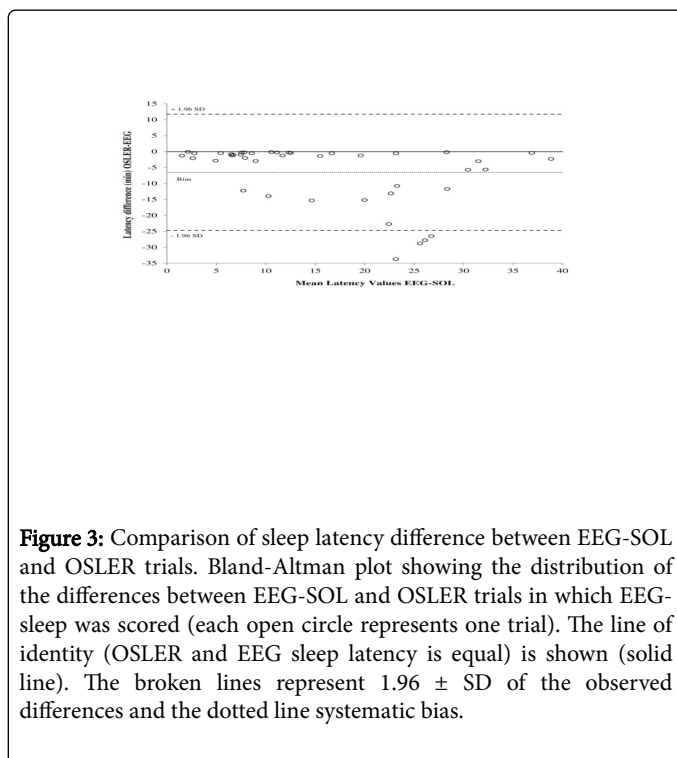


Figure 3: Comparison of sleep latency difference between EEG-SOL and OSLER trials. Bland-Altman plot showing the distribution of the differences between EEG-SOL and OSLER trials in which EEG-sleep was scored (each open circle represents one trial). The line of identity (OSLER and EEG sleep latency is equal) is shown (solid line). The broken lines represent $1.96 \pm \text{SD}$ of the observed differences and the dotted line systematic bias.

The bias towards lower values were 6.5 min for EEG-SOL [SD 9.3 95% CI (24.7, 11.7) min].

Assessment of the accurate number of OSLER trials to estimate the four-trial

OSLER test: The median OSLER value for each trial was 40.0 min with inter-quartile range 40.0 (14.8 to 40.0), 40.0 (13.9 to 40.0), 34.2 (8.6 to 40.0) and 37.1 (12.5 to 40.0) at 8:00 am, 10:00 am, 12:00 pm and 2:00 pm respectively and were not significantly different.

The accuracy of 1, 2 or 3 OSLER trials to estimate a (40 min) four-trial OSLER test was studied using ROC curves for the 14 different possibilities. The mean Area under the Curve (AUC) for 1, 2 or 3 trials were not statistically different compared to the four-trial OSLER test (data not shown). We also studied the difference in means between the use of 1, 2 or 3, vs. four-trial OSLER test by using Bland-Altman plots. We also found no statistically significant difference between deltas in means for all 14 possibilities (range 1.5 min to 1.0 min, NS). However, standard deviations for these differences were higher when only one trial was used (range 6.7 min to 10.2 min) vs. 3 trials (range 2.2 min to 3.4 min) with mid-results when 2 trials were used (range 4.8 min to 5.1 min). An example is shown in Figure 4 as a Bland-Altman plot. It can be seen that the data is less scarce around the identity line when using a combination of 3 trials (mean of the 8:00 am, 10:00 am and 12:00 pm trials - Figure 4, right hand panel) vs. a more scattered appearance around the identity line with the use of only one trial (8:00 am - Figure 4, left hand panel).

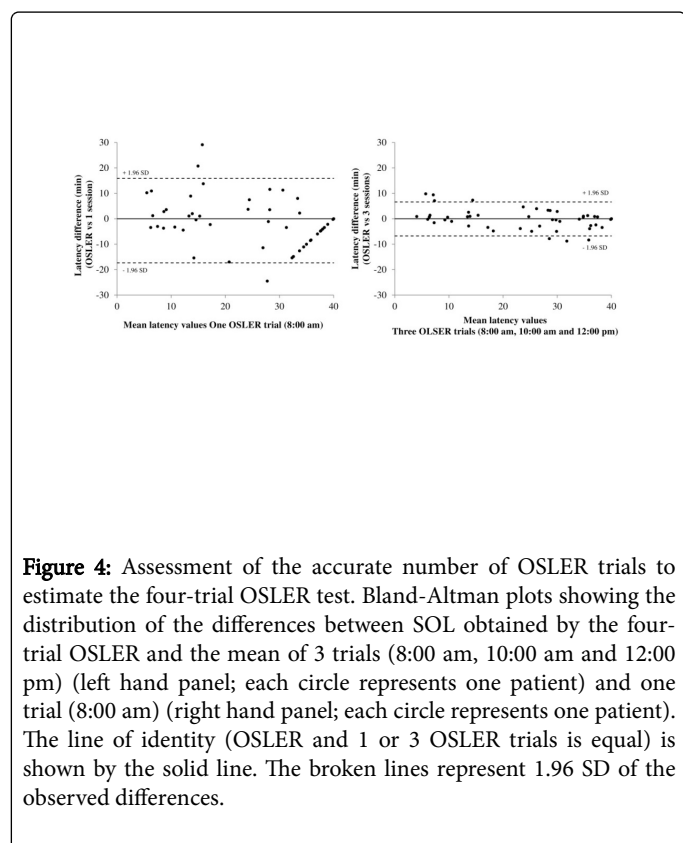


Figure 4: Assessment of the accurate number of OSLER trials to estimate the four-trial OSLER test. Bland-Altman plots showing the distribution of the differences between SOL obtained by the four-trial OSLER and the mean of 3 trials (8:00 am, 10:00 am and 12:00 pm) (left hand panel; each circle represents one patient) and one trial (8:00 am) (right hand panel; each circle represents one patient). The line of identity (OSLER and 1 or 3 OSLER trials is equal) is shown by the solid line. The broken lines represent 1.96 SD of the observed differences.

Discussion

This study assessed the accuracy of the OSLER test compared to two sets of EEG scoring using AASM criteria to score sleep in a large population mainly composed of patients with OSA on treatment referred for assessment of their daytime vigilance. Typically, patients

were men, 40 years to 60 years old, with various RDI, BMI and ESS scores (Table 1).

Data	Sixty-five patients	Minimum	Maximum
Age, years	50.5 ± 10.8	23	70
Gender Men:Women	57:08:00	-	-
RDI (events/hour) (n=59)	38.9 ± 26.9	0.3	103
Body Mass Index: BMI	33.5 ± 7.9	19	58.7
ESS (n=55)	10.3 ± 6.3	1	24

Table 1: Baseline characteristics of the study patients. Baseline characteristics such as age, gender and body mass index (BMI) were collected for all patients (n=65). However, some data were not available for all the patients (respiratory disturbance index (RDI) n=59) and Epworth sleepiness scale (ESS) n=55).

The major finding of the study is that using current AASM criteria to score sleep during daytime studies, an OSLER where no sleep onset is recorded during four 40 minutes trials is reliably associated with the absence of sleep during the procedure. However, in patients with sleep onset during OSLER trials, the occurrence of EEG-determined sleep before the end of the trial occurred in about 1 out of 3 trials, which precludes the OSLER to be used as a precise surrogate to detect EEG sleep onset. Results also pointed out that at least a three-trial OSLER is necessary to estimate with good confidence the results of the four-trial OSLER.

The OSLER, either in its original [3] (or more recently modified versions [14]) is currently being used widely in the literature to assess vigilance and attention deficits in patients with sleep disorders as well as being able to reflect the presence or absence of improvement after therapy. However, it is still not clear that it can be used to detect true physiological sleep onset due to the test design itself as well as its validation in studies with small population samples and various definitions of EEG sleep onset [15,16].

On the one hand, we observed in most trials that no EEG sleep was scored although patients failed the OSLER. Since the OSLER test relies on behavior (e.g. frequency of eye blinking, fingering and fluctuation of vigilance/sustain attention system) [17], the cooperation and the motivation of the individual is essential. Therefore, false-positive results are observed (e.g. omissions without sleep).

We observed that EEG sleep onset latency is often lower than sleep onset recorded in the OSLER. The monotonous nature of the OSLER test, i.e. the constant repetitions of the diodes one second every three seconds, permits the patients to get used to the task. This could therefore explain why periods of sleep may happen with an epoch of sleep being scored according to the rule of Rechtschaffen and Kales [13]. This could also reflect the phenomenon raised by Ogilvie stating that in some patients, there is a difference in the relationship of behavioral response and EEG criteria; certain patients can remain fairly responsive in stage 1 while others do not [18].

On the other hand, we determined the optimal number of OSLER trials needed to estimate acceptable accuracy of the four-trial OSLER test. As a group, the mean values were comparable. However, we found large standard deviations around the mean (ranging from 5 min to 10 min) that, in our opinion, precludes the utilization of the test with only one or two trials. Individually, we could observe as high as a 10 min

difference between the result of a single trial at 8:00 am and a four-trial OSLER test. It is not surprising that three trials are necessary given the expression of the circadian cycle, eye blinking and inter-individual differences in vigilance/attention capacities. We believe that these differences could be clinically significant. We agree with Priest who stated that at least a three-trial OSLER test was necessary [4]. The confirmation of this finding in this specific population is important since these patients were referred to objectively test their wake tendency.

This study presents some limitations. Our study was not designed to compare the OSLER test with MWT results, since no MWT was done and it is likely that the OSLER test influences waking tendency. One cannot infer from our findings that, in a given subject, the OSLER test would estimate the result of the MWT. The strength of our study is to better understand the nature of the OSLER test and its usefulness in a large sleep clinic. The EEG monitoring was used to better understand the capacities of the OSLER test to detect sleep onset utilizing a request to elicit a response. We did not take into account malingering that could have affected our results; the referring physician did not report this in any case. Our population was heterogeneous, as seen in a tertiary sleep clinic, and mostly composed of patients with sleep disordered breathing. The study did not have the statistical power to report on the reliability of the tests in sub-groups of patients (effect of stimulants, more than one diagnosis, hypersomnia not caused by sleep apnea). Finally, this study did not assess micro sleep (sleep intrusion of 3 seconds to 15 seconds in the EEG) which was almost always present when four or more consecutive LED were missed in the work of Priest [4]. The correlation of micro sleep and clinical endpoints needs more investigations since it could be useful for sleep specialists in is currently being investigated by other groups [9,10].

We conclude that, using AASM 2005 criteria to score sleep during daytime studies, patients with sleep disorders referred for assessment of their vigilance that successfully remain awake during the OSLER test (four 40 min trials) have a very low probability of EEG sleep onset during the procedure. However, patients who missed 7 consecutive stimuli and therefore have a sleep latency scored using the OSLER have a significant chance of not having EEG-recorded sleep. This precludes the OSLER to be used as a precise surrogate to detect EEG sleep onset latency. Given the large variability of the measurements, we suggest that at least three trials should be done to accurately estimate the value of the four-trial OSLER. More studies are needed to evaluate the OSLER as a tool to safely assess not only sleep onset but also attention and vigilance especially when individuals with sleep disorders are involved in tasks where this could be a concern.

Acknowledgements

The authors acknowledge the assistance of Anne-Marie Laurin, respiratory therapists and all the technicians of the Laboratoire du sommeil of Hôpital Hôtel-Dieu du Centre Hospitalier Universitaire de Montréal (CHUM). Thanks to Joseph Braidy, M.D., Mathieu Gaudet, M.D. and Zeina Kaddaha, M.D. for reviewing the manuscript. The sleep laboratory of the clinic is supported by Biron Soins du Sommeil (Québec, Canada) and Medigas (Québec, Canada).

References

1. Doghramji K, Mitler MM, Sangal RB, Shapiro C, Taylor S, et al. (1997) A normative study of the maintenance of wakefulness test (MWT). *Electroencephalogr Clin Neurophysiol* 103: 554-562.
2. Littner MR, Kushida C, Wise M, Davila DG, Morgenthaler T, et al. (2005) Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. *Sleep* 28: 113-121.
3. Bennett LS, Stradling JR, Davies RJ (1997) A behavioural test to assess daytime sleepiness in obstructive sleep apnoea. *J Sleep Res* 6: 142-145.
4. Priest B, Brichard C, Aubert G, Liistro G, Rodenstein DO (2001) Microsleep during a simplified maintenance of wakefulness test. A validation study of the OSLER test. *Am J Respir Crit Care Med* 163: 1619-1625.
5. Bennett LS, Barbour C, Langford B, Stradling JR, Davies RJ (1999) Health status in obstructive sleep apnea: relationship with sleep fragmentation and daytime sleepiness, and effects of continuous positive airway pressure treatment. *Am J Respir Crit Care Med* 159: 1884-1890.
6. Chouri-Pontarollo N, Borel JC, Tamisier R, Wuyam B, Levy P, et al. (2007) Impaired objective daytime vigilance in obesity-hypoventilation syndrome: impact of noninvasive ventilation. *Chest* 131: 148-155.
7. Krieger AC, Ayappa I, Norman RG, Rapoport DM, Walsleben J (2004) Comparison of the maintenance of wakefulness test (MWT) to a modified behavioral test (OSLER) in the evaluation of daytime sleepiness. *J Sleep Res* 13: 407-411.
8. West SD, Kohler M, Nicoll DJ, Stradling JR (2009) The effect of continuous positive airway pressure treatment on physical activity in patients with obstructive sleep apnoea: A randomised controlled trial. *Sleep Med* 10: 1056-1058.
9. Ferreira JJ, Neutel D, Mestre TA, Galitzky M, Thalamas C, et al. (2013) Usefulness of OSLER test in Parkinson's disease. *Sleep Med* 14: 297-298.
10. Alakuijala A, Maasilta P, Bachour A (2014) The Oxford Sleep Resistance test (OSLER) and the Multiple Unprepared Reaction Time Test (MURT) detect vigilance modifications in sleep apnea patients. *J Clin Sleep Med* 10: 1075-1082.
11. Mitler MM, Gujavarty KS, Browman CP (1982) Maintenance of wakefulness test: a polysomnographic technique for evaluation treatment efficacy in patients with excessive somnolence. *Electroencephalogr Clin Neurophysiol* 53: 658-661.
12. American Sleep Disorders Association (1994) Practice parameters for the use of portable recording in the assessment of obstructive sleep apnea. Standards of Practice Committee of the American Sleep Disorders Association. *Sleep* 17: 372-377.
13. Kales A, Rechtschaffen A (1968) A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. U. S. National Institute of Neurological Diseases and Blindness, Neurological Information Network; United States.
14. Karimi M, Hedner J, Zou D, Eskandari D, Lundquist AC, et al. (2015) Attention deficits detected in cognitive tests differentiate between sleep apnea patients with or without a motor vehicle accident. *Sleep Med* 16: 528-533.
15. Marklund M, Carlberg B, Forsgren L, Olsson T, Stenlund H, et al. (2015) Oral Appliance Therapy in Patients With Daytime Sleepiness and Snoring or Mild to Moderate Sleep Apnea: A Randomized Clinical Trial. *JAMA Intern Med* 175: 1278-1285.
16. Neutel D, Peralta R, Pires J, Bentes C, Ferreira JJ (2015) End of OSLER Test Sessions in Parkinson's Disease do not Correspond to True Sleep Onset: Results from an Exploratory Study. *Front Neurol* 6: 200.
17. Hakkanen H, Summala H, Partinen M, Tiihonen M, Silvo J (1999) Blink duration as an indicator of driver sleepiness in professional bus drivers. *Sleep* 22: 798-802.
18. Ogilvie RD, Wilkinson RT, Allison S (1989) The detection of sleep onset: behavioral, physiological, and subjective convergence. *Sleep* 12: 458-474.