Forty versus twenty minutes trials maintenance of wakefulness test regimen for driving licensing.

Limor Arzi MD, Roni Shreter MD, Baruch El-Ad MD, Ron Peled MD, Giora Pillar MD, PhD

The Sleep Laboratory, Rambam Medical Center and Faculty of Medicine, Technion, Haifa 31096, ISRAEL.

Correspondence:
Giora Pillar, MD, PhD
Sleep Laboratory
Rambam Medical Center,
Haifa 31096
Tel: 972-4-8538585
Fax: 972-4-8542441
E-mail gpillar@tx.technion.ac.il

Running head: MWT40 vs MWT20 for driving licensing
Abstract:

Study Objectives:
Objective assessment of the ability to maintain wakefulness, although very important, is still equivocal. A recent study from our lab has shown that the Maintenance of Wakefulness Test, when performed with the 20 minute protocol (MWT20), is unreliable in assessing patients who are highly motivated to maintain wakefulness. In this study we sought to examine whether the 40 minute protocol (MWT40) is a better tool in assessing such individuals.

Methods:
One hundred and sixty four consecutive subjects referred to our sleep lab by the Medical Institute for Driving Safety were studied. All subjects underwent a full night polysomnography (PSG) followed by an MWT, 4 trials of 40 minutes each. All subjects knew that if they failed the wakefulness test their driving license would be revoked.

Results:
Forty one subjects out of 164 (25%) fell asleep at least once. Of 39 subjects with severe OSA (RDI>40/h), 19 fell asleep (48.7%). Of 13 subjects with minimal oxygen saturation level below 65%, 7 fell asleep (53%). In the MWT20, only 7% of patients with severe OSA fell asleep at least once.

Conclusions:
We conclude that the MWT40 is superior to the MWT20 in detecting difficulties maintaining wakefulness in a highly motivated population. However, our results yield significantly lower detection of difficulties maintaining wakefulness than those reported in healthy subjects, suggesting that the MWT40 is also highly affected by motivation. We believe that, for a highly motivated population (such as for driving license validation), different average sleep latency threshold should be used than in general population.

Keywords: Excessive Daytime Sleepiness (EDS), Maintenance of Wakefulness Test (MWT), Motivation, Obstructive Sleep Apnea (OSA)
Abbreviations:

BMI: Body Mass Index
CPAP: Continuous positive airway pressure
ECG: electrocardiography
EEG: electroencephalography
EMG: electromyography
EOG: electrooculography
ESS: Epworth Sleepiness Scale
MSLT: Multiple Sleep Latency Test
MWT: Maintenance of Wakefulness Test
OSA: Obstructive Sleep Apnea
PSG: Polysomnography
PVT: Psychomotor Vigilance Test
RDI: Respiratory Disturbance Index
Introduction
Prevention of traffic accidents is a challenge for authorities worldwide. Many traffic accidents, in particular fatal ones, are caused by drivers who suffer from excessive daytime sleepiness and fall asleep while driving. Excessive Daytime Sleepiness (EDS) may result from various sleep disorders such as: Obstructive Sleep Apnea (OSA), insomnia, jet lag, sleep deprivation, narcolepsy and others. Excessive daytime sleepiness impairs the individual’s daytime performance and may lead to traffic and work-related accidents. A case-control study by Teran-Santos et al., on groups of drivers who were involved in traffic accidents, determined that drivers who suffer from sleep apnea (over 10 apneas/hypopneas per hour) have a higher likelihood to be involved in traffic accidents. Treatment of OSA with Continuous Positive Airway Pressure (CPAP) can reduce the likelihood for accidents. Therefore, diagnosing and treating these patients is an important public health issue.

It is difficult to assess daytime sleepiness and maintenance of wakefulness objectively and accurately. While the quality of sleep and the severity of sleep disorders can be objectively quantified, there is no clear correlation between any of these variables and the degree of daytime sleepiness, assessed by the currently available tests. These tests consist of subjective questionnaires such as the Epworth Sleepiness Scale (ESS) and objective methods such as Multiple Sleep Latency Test (MSLT), the Maintenance of Wakefulness Test (MWT), and the more recently introduced Psychomotor Vigilance Test (PVT).

While in the MSLT the patient is instructed to try falling asleep during five 20-minute daytime trials, in the MWT the subject is instructed to stay awake in soporific circumstances during four 20 or 40-minute trials. Thus, the MSLT test quantifies the drive to sleep while the MWT test evaluates the ability to maintain wakefulness. It is quite common that people who fall asleep in the MSLT manage to maintain wakefulness in the MWT. Hence, logically, the MWT is a preferable test for examining the ability of drivers to remain alert, in cases of suspected excessive daytime sleepiness.
The diagnostic criteria for the MWT were determined in 1997 by Doghramji et al. In a research conducted on 64 healthy subjects, a score of less than two standard deviations from the average was considered diagnostic of a severe problem in the ability to maintain wakefulness. The threshold value for the MWT20 protocol was lower than 10.9 minutes and for the MWT40 lower than 19.4 minutes. In a similar trial by Banks et al., the threshold value for the MWT40 was 26.1 minutes. However, these tests were performed in experimental settings with no potential real-life factors which may affect the results. The clinical use of the MSLT and MWT have been recently thoroughly reviewed by Arand et al. Among other conclusions, they suggested that the MWT is affected by many factors besides daytime sleepiness and that it doesn't discriminate well between normal population and patients with sleep disorders.

In a recent study from our lab, we reported that under conditions of high motivation, when failing the MWT test may lead to losing one's driving license, only 5 out of 54 potentially sleepy drivers fell asleep in any of five trials of the MWT20. The calculated average MWT scores were higher than in any of the previously published papers. Some of the patients in that study had severe untreated OSA and others had been involved in a previous sleep-related accident. From the subgroup of twenty one subjects with severe untreated OSA only one fell asleep. We, therefore, concluded that motivation had a significant effect on the MWT20, and that the test is not reliable in the evaluation of sleepiness in highly motivated subjects. Thus, in the current study, we sought to assess whether the MWT40 regimen is a more reliable tool to quantify difficulties maintaining wakefulness in a motivated population, although, a study by Bonnet et al. reported that financial reward could not motivate 12 healthy volunteers to increase their sleep latency in MWT40 protocol, suggesting it is less likely affected by motivation and a better measure of alertness than the MSLT. In the current study we utilized a different motivator (the threat of losing a driving license) and sought to compare the MWT40 (tested on a new dataset of patients) to the MWT20 regimens (previously reported patients), in a population of potentially sleepy patients.
Methods

Subjects:
The study was consisted of newly tested one hundred and sixty four consecutive subjects who were referred to the Technion sleep lab from June 2006 until January 2007 by the Medical Institute for Driving Safety of the Ministry of Transportation. The subjects were referred in order to evaluate their ability to maintain wakefulness for the purpose of renewal of their driving license. Many of the subjects were drivers who were previously diagnosed with a sleep disorder. Among them thirty nine subjects had OSA, ten of them were already being treated with nightly CPAP. Other subjects were suspected of having excessive daytime sleepiness due to variety of reasons, including: snoring, obesity, or previous involvement in a sleep related car accident. Seventeen of the subjects were professional drivers, one was an aviation pilot. All subjects were referred to our center for the purpose of undergoing an MWT, and all were informed that if they fail the test (i.e. fall asleep) they may lose their driving license. These patients were compared to patients from a previously reported paper with similar recruitment method and similar population, tested by the MWT20 protocol33.

Study Protocol:
All participants filled out the Technion sleep questionnaire35,36 and the Epworth Sleepiness Scale. On the night previous to the MWT the subjects underwent a medical interview and then a complete full night polysomnography (PSG). On the following day the MWT40 was performed.

Test Procedures:
Overnight polysomnography:
Each patient was studied in a private room. The recording included two electroencephalography (EEG) channels (C3-A2, O2-A1), two electrooculography (EOG) channels, an electromyography (EMG) channel (submental), leg movements (tibialis anterior EMG), an electrocardiography (ECG) channel, airflow (both thermistor and nasal pressure), respiratory effort (thoracic belt), oxygen saturation (pulse oximeter), and quantitative snoring
intensity (by a dB-meter placed one meter above the bed). A technician viewed the patients during the night by a closed circuit monitor. Bedtime was between 22:00-23:00 and the patients were woken up at 6:30-7:00, to start the MWT at 8:00.

The results of the recording were scored by an experienced technician based on the Chicago criteria with the following definitions: apnea was defined as a cessation of airflow for 10 seconds or more. Hypopnea was defined as any notable decline in airflow that was accompanied by an arousal (defined as an appearance of 3-second alpha wave or frequency change in the EEG or an increase in submental EMG signal) or a decline in oxygen saturation level (at least 3%). The Respiratory Disturbance Index (RDI) was calculated by dividing the total amount of respiratory events by the total sleep time. The minimal oxygen saturation was the lowest saturation value due to a respiratory event (excluding artifacts) and the maximal snoring intensity was the maximal recorded volume in decibels that was due to snoring (excluding speech, cough or other noises).

Maintenance of wakefulness test (MWT):
Each of the subjects was placed in a private, dimly lit and quiet room and was told to sit down in bed (in a reclining position, with the back at about 45 degrees to the bed surface) and maintain wakefulness for 40 minutes. The subject was not allowed to perform any unusual action in order to maintain wakefulness (such as read, talk or make repetitive movements). The recording included two EEG channels, two EOG channels, an EMG channel placed on the submental muscle, an ECG channel and quantitative snoring intensity. Trial was stopped if the subject fell asleep or if the subject maintained wakefulness for 40 minutes. For the purpose of terminating a study, sleep was defined as three consecutive 30-second epochs of stage 1 sleep, or any epoch of any other sleep stage. For the MWT score, sleep latency was calculated as the latency to the first epoch of any sleep stage. During the test four such trials were performed at 8:00, 10:00, 12:00 and 14:00. Subjects were given breakfast and lunch between the trials. Caffeine-
containing drinks were not allowed, as well as any drugs other than those medically prescribed.

For each trial the sleep latency was recorded (the time was 40 minutes if the patient maintained wakefulness). The calculated score of the test was the average score of all four trials.

Results

Table 1 summarizes the demographic and nocturnal PSG variables of the 164 subjects. The average age was 50±14 years, the average Body Mass Index (BMI) was 35±6.5 kg/m² and the average RDI was 25±23. About a quarter of the participants had substantially disturbed nocturnal sleep, such as total sleep time of less than 5 hours with no REM detected, or sleep latency of greater than 4 hours. Over 20% of the participants had severe OSA with RDI of up to 85/h.

Forty one out of the 164 subjects (25%) fell asleep in one or more of the MWT40 trials, with the following characteristics: 15 (9.1%) subjects fell asleep in only one trial, 11 (6.7%) in two trials, 5 (3%) in three trials and 10 (6.1%) fell asleep in all four trials, the results are displayed in figure 1. In the MWT20 study that was conducted on a similar population only 5 out of 54 subjects fell asleep in any of the trials (9.2%, p<0.05), the distribution of the subjects is displayed in figure 2.

In tables 2, 4 and 5 we sorted and grouped our subjects according to different medical and nocturnal parameters. Since data were missing on a few of our subjects in table 2 four patients are excluded, in table 4 five patients are excluded and in table 5 three patients are excluded.

By grouping the subjects according to their RDI (table 2), it can be seen that in the severe OSA group (RDI>40), 19 out of 39 subjects (48.7%) fell asleep at least in one of the MWT40 trials, and 11 (28%) fell asleep more than once. When we look in the three lower severity categories, the percentage of subjects that fell asleep is significantly smaller (p<0.05 between RDI>40 and other groups). In the MWT20 study there was no clear association between RDI score and the number of subjects that fell asleep (table 3).

Table 4 shows the data sorted according to the minimal oxygen saturation level recorded during the PSG test. The results show that when the minimal
saturation was 95%-100%, only three out of 20 subjects (15%) fell asleep in any of the MWT trials, while in the severe OSA group (minimal oxygen saturation below 65%), 7 out of 13 subjects fell asleep (53%, p<0.05). Dividing the subjects into subgroups according to their BMI (table 5), demonstrates that there was an increased tendency to fall asleep in the MWT40 with increasing BMI, although the differences were not statistically significant.

**Discussion**

Our study shows that the MWT40 protocol is superior to the MWT20 protocol in detecting difficulties to maintain wakefulness in a highly motivated albeit sleepy population. While with the MWT20 less than 10% of potentially sleepy subjects fell asleep in any of the trials, with the MWT40 this proportion more than doubled, reaching 25%.

The purpose of this study was to evaluate the ability of the MWT40 to detect failure to maintain wakefulness in people with suspected excessive daytime sleepiness and high motivation to show normal daytime alertness. Evaluating daytime sleepiness has always been a challenge. In our study, the challenge was even greater because our subjects were highly motivated to maintain wakefulness. While the MWT20 protocol indicated low resolution ability of the test (only 10% of potentially sleepy subjects fell asleep in any trial, with no association between falling asleep during the daytime test and any nocturnal sleep parameters), with the MWT40 regimen this figure increased. Although it is very challenging to quantify the ability to maintain wakefulness, and although we do not have a gold standard to compare our results to, our study showed that with the MWT40 more subjects fell asleep, with a significantly higher proportion of patients with severe OSA falling asleep. These findings suggest that the MWT40 protocol is superior to the MWT20 protocol in detecting subjects unable to stay awake. Furthermore, in our current study we found a significant negative correlation between RDI/nocturnal sleep disruption and the ability to maintain wakefulness, which was not found with the MWT20 protocol. It is a common-sense assumption that, with longer trials,
it becomes harder for sleepy patients to maintain wakefulness, even when highly motivated.

Obviously the MWT40 is not perfect. In a study of healthy subjects reported by Doghramji et al.29, the average mean sleep latency was 35.2±7.9 minutes and the lower normal limit was below 19.4 minutes. In contrast, the mean sleep latency of our subjects who were suspected of suffering from EDS was 37±6.6 minutes. Only eight of our subjects had sleep latency below the reported low-normal cut-off value of 19.4 minutes, four of whom had severe OSA (all were untreated), one suffered from restless legs syndrome, and three had no previous diagnosis. The fact that our subjects were sitting in a reclining position (rather than lying down) further emphasizes these results. Thus, we believe that high motivation of the subjects may still contribute significantly to their ability to maintain wakefulness. Our subjects' results indicate that they are even more alert than healthy subjects without sleep disorders but also without high motivation to stay awake. Although Bonnet et al. 34 found that financial reward could not motivate 12 healthy volunteers to increase their sleep latency in the MWT40 protocol, we believe that losing a driver's license is a stronger motivator than a monetary one, and healthy volunteers may be differently affected by motivation than patients. Although the mean MWT sleep latency in our highly motivated subjects is not shorter than the mean sleep latency reported in normal subjects, as determined by Doghramji et al. 27, we still believe that the average sleep latency is a good parameter for identifying people with difficulty maintaining wakefulness. We think that greater importance should be attribute to the number of trials subjects fall asleep in. Comparing number of trials containing sleep between the general population, people with sleep disorders and highly motivated populations and correlating the MWT results with outcomes (i.e. driving accidents) needs to be supported by future research.

Our study had several limitations. One of them is the lack of control group of healthy subjects and another is the lack of outcome data. We have scored hypopneas based on the Chicago criteria which probably have resulted in higher respiratory indices than if scored according to the new AASM
guidelines. The main general limitation is the lack of a "gold standard" test, an objective way to determine the "real" tendency of these subjects to fall asleep. Comparing the results of the MWT40 to the MWT20 in different cohorts is also not ideal. However, although different subjects were studied, their general characteristics were very similar and did not differ between the cohorts (i.e. they were both referred by the Medical Institute for Driving Safety of the Ministry of Transportation in order to evaluate their ability to maintain wakefulness for the purpose of renewal of their driving license, and they were of similar age: 50.7 vs 49.2 years, NS; similar BMI: 35.2 vs 34Kg/m², NS, and similar RDI: 25.5 vs 26.3/h, NS). Only the time of study really differed between the cohorts without any meaningful change in referral characteristics. It therefore seems that there is a substantial effect of the test protocol itself. Thus we believe that our results preclude using the MWT20 protocol for the purpose of driving license validation in the future. Whether PVT or any other objective measures of the ability to maintain wakefulness is better than the MWT40 remains to be determined. In addition, it is difficult to quantify motivation. Although we believe that losing one’s driving license highly motivates a subject to stay awake, individual differences are always an issue.

In conclusion, despite these limitations, we believe our study shows that, for the evaluation of drivers with suspected EDS, the MWT40 is a more appropriate test than the MWT20. The test, however, is still affected by motivation and as such is not completely reliable. The evaluation of EDS is still challenging. Setting a threshold for study failure or developing other methods should be further explored.
Figure 1: Number of falling asleep cases from 164 subject who underwent MWT40, 123 (75%) never fell asleep, 15 (9.1%) fell asleep at one trial, 11 (6.7%) at two trials, 5 (3%) at three trials and 10 (6.1%) at all four trials.

Figure 2: Data retrieved from Shreter et al.33: Number of falling asleep cases from 54 subjects who underwent MWT20, 59 (90.7%) never fell asleep, 2 (3.7%) fell asleep once, and three (5.6%) fell asleep twice or more (up to five trials).
## Tables

### TABLE 1: Demographics and Nocturnal Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>50.66</td>
<td>14.10</td>
<td>20-78</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>107.60</td>
<td>22.11</td>
<td>50-165</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>35.17</td>
<td>6.49</td>
<td>17.76-52.08</td>
</tr>
<tr>
<td>Sleep latency (min)</td>
<td>28.52</td>
<td>34.51</td>
<td>0-285</td>
</tr>
<tr>
<td>Sleep period time (min)</td>
<td>443.96</td>
<td>33.10</td>
<td>338-528</td>
</tr>
<tr>
<td>Total sleep time (min)</td>
<td>343.98</td>
<td>69.59</td>
<td>117-480</td>
</tr>
<tr>
<td>RDI (kg/m²)</td>
<td>25.46</td>
<td>23.05</td>
<td>0-85</td>
</tr>
<tr>
<td>Lowest SaO₂ (%)</td>
<td>84.04</td>
<td>11.6</td>
<td>51-100</td>
</tr>
<tr>
<td>Stage 2 sleep (%)</td>
<td>61.66</td>
<td>14.1</td>
<td>27-95</td>
</tr>
<tr>
<td>Stage 3-4 sleep (%)</td>
<td>17.09</td>
<td>12.5</td>
<td>0-56</td>
</tr>
<tr>
<td>REM time (%)</td>
<td>20.98</td>
<td>7.8</td>
<td>0-41</td>
</tr>
</tbody>
</table>

### Table 2: Results sorted by RDI in MWT40

<table>
<thead>
<tr>
<th>RDI</th>
<th>&lt;5</th>
<th>5 - 19</th>
<th>20-39</th>
<th>≥40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects (n)</td>
<td>33</td>
<td>50</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>Fell asleep (n)</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>once</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>twice</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>≥3 times</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Stayed awake (n)</td>
<td>28</td>
<td>40</td>
<td>33</td>
<td>20</td>
</tr>
<tr>
<td>% falling asleep</td>
<td>15.15%</td>
<td>20.00%</td>
<td>13.16%</td>
<td>48.72%</td>
</tr>
<tr>
<td>Average MWT40 (min)</td>
<td>38.18</td>
<td>37.69</td>
<td>38.09</td>
<td>34.93</td>
</tr>
</tbody>
</table>

### Table 3: Results sorted by RDI in MWT20 (Data retrieved from Shreter et al.33):

<table>
<thead>
<tr>
<th>RDI</th>
<th>&lt;5</th>
<th>5-19</th>
<th>20-39</th>
<th>≥40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects (n)</td>
<td>14</td>
<td>14</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Fell asleep (n)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Stayed awake (n)</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>% falling asleep</td>
<td>7.14%</td>
<td>14.29%</td>
<td>8.33%</td>
<td>7.14%</td>
</tr>
<tr>
<td>Average MWT20 (min)</td>
<td>19.92</td>
<td>19.90</td>
<td>19.64</td>
<td>19.67</td>
</tr>
</tbody>
</table>
### Table 4: Results sorted by minimal oxygen saturation level in MWT40:

<table>
<thead>
<tr>
<th>Saturation (%)</th>
<th>95-100</th>
<th>85-94</th>
<th>65-84</th>
<th>&lt;65</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects (n)</td>
<td>20</td>
<td>83</td>
<td>43</td>
<td>13</td>
</tr>
<tr>
<td>Fell asleep (n)</td>
<td>3</td>
<td>16</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>once</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>twice</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>≥3 times</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Stayed awake (n)</td>
<td>17</td>
<td>67</td>
<td>31</td>
<td>6</td>
</tr>
<tr>
<td>% falling asleep</td>
<td>15.00%</td>
<td>19.28%</td>
<td>27.91%</td>
<td>53.85%</td>
</tr>
<tr>
<td>Average MWT40 (min)</td>
<td>38.94</td>
<td>37.48</td>
<td>36.1</td>
<td>36.18</td>
</tr>
</tbody>
</table>

### Table 5: Results sorted by BMI in MWT40:

<table>
<thead>
<tr>
<th>BMI (kg/m^2)</th>
<th>&lt;25</th>
<th>25-30</th>
<th>&gt;30-40</th>
<th>&gt;40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects (n)</td>
<td>10</td>
<td>18</td>
<td>96</td>
<td>37</td>
</tr>
<tr>
<td>Fell asleep (n)</td>
<td>2</td>
<td>3</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td>Stayed awake (n)</td>
<td>8</td>
<td>15</td>
<td>72</td>
<td>26</td>
</tr>
<tr>
<td>% falling asleep</td>
<td>20.00%</td>
<td>16.67%</td>
<td>25.00%</td>
<td>29.73%</td>
</tr>
<tr>
<td>Average MWT40 (min)</td>
<td>38.18</td>
<td>38.01</td>
<td>36.53</td>
<td>37.64</td>
</tr>
</tbody>
</table>
References


