REVIEW PAPER

The Clinical Use of the MSLT and MWT

Donna Arand, Ph.D[1], Michael Bonnet, Ph.D[2], Thomas Hurwitz, M.D[3], Merrill Mitler, Ph.D[4], Roger Rosa, Ph.D[5] and R. Bart Sangal, M.D[6]

[1] Kettering Medical Center, Dayton, OH, [2]Dayton Veteran's Affairs Medical Center and Wright State University, Dayton, OH, [3]Minneapolis Veteran's Affairs Medical Center and University of Minnesota, Minneapolis, MN, [4]National Institute of Neurological Disorders & Stroke Neuroscience Center, Bethesda, MD, [5]National Institute for Occupational Safety and Health, Washington, DC, [6]Sleep Disorders Institute, Troy, MI

Citation: Review by the MSLT and MWT Task Force of the Standards of Practice Committee of the American Academy of Sleep Medicine. *SLEEP* 2005;28(1):123-144.

1.0 INTRODUCTION

EXCESSIVE SLEEPINESS IS DEFINED AS SLEEPINESS OCCURRING IN A SITUATION WHEN AN INDIVIDUAL WOULD BE EXPECTED TO BE AWAKE AND ALERT. It is a chronic problem for about 5% of the general population,[1, 2] and it is the most common complaint evaluated by sleep disorder centers.[1, 3] The most common causes of sleepiness include partial sleep deprivation, fragmented sleep and medication effects. Sleepiness is also associated with sleep disorders such as sleep apnea[4] and narcolepsy[5] as well as some medical and psychiatric disorders.[6, 7] Excessive sleepiness and impaired alertness during wakefulness have been associated with increased morbidity for the individual and may threaten the safety of others as well.[8, 9]

The assessment of sleepiness is an important part of the evaluation and management of the sleep disorders patient. Historically, the Multiple Sleep Latency Test (MSLT)[10] and the Maintenance of Wakefulness Test (MWT)[11] have been the primary objective tests used for the measurement of sleepiness and alertness, respectively. Results of these tests are used to help diagnose sleep disorders, evaluate treatment and to support recommendations to the patient on the advisability of driving or performing other daily activities. The tests have also been recommended by regulatory agencies to support a decision to return a sleep disorders patient to service in a safety-sensitive position after treatment. This decision, and the measurements used to make it, may critically affect both public safety and the career of the patient. Thus, it is extremely important that evidence-based data are examined to determine the usefulness and limitations of these tests. The current review provides an examination of sleep latency data reported in the peer-reviewed literature.

Guidelines for the use of the MSLT were initially established in 1992 based on consensus.[12] Since then, the body of scientific literature using the MSLT and MWT has grown substantially. Consequently, this review was requested by the Standards of Practice Committee of the American Academy of Sleep Medicine (AASM). All authors completed AASM conflict of interest statements and were not found to have any significant conflicts with regard to the topics discussed in this review. The objectives of this paper are to review unique issues related to measuring sleepiness and to assess the database of evidence for the clinical use of the MSLT and MWT.

Disclosure Statement

Dr. Sangal has received support from Cephalon, Lilly, Takeda, Sanofi, Novartis and Aventis. Dr. Bonnet has received research support from Cephalon and Pfizer. Dr. Mitler is an owner of Pacific Sleep Medicine Services, Inc. Drs. Arand, Hurwitz and Rosa have indicated no financial conflicts of interest.

2.0 BACKGROUND

2.1 History of the Development of MSLT and MWT

Initially, definitions involving sleep onset were based on behavioral observations that often depended on features such as lack of movement, unresponsiveness, snoring, etc. After the development of the electroencephalogram (EEG), it became clear that the sleep state was associated with changes in the EEG that distinguished it from wakefulness.[13-15] A system for describing sleep architecture based on EEG changes was developed by Rechtschaffen and Kales.[16] This system made possible the quantification of sleep stages and the measurement of sleep latency, the time from lights out to the onset of EEG-measured sleep.

During the formalization of sleep disorders medicine, interest focused on the assessment of sleepiness. The Stanford Sleepiness Scale (SSS), a subjective report of sleepiness, initially provided an inexpensive and easy to administer tool.[17] However, other research at this time examined sleep and wakefulness on a 90-min day and provided the basis for the development of an objective test, the MSLT.[18-20] The 90-min day paradigm placed college students, who typically slept 7.5 - 8 hours per night, on a schedule alternating 60 min of wakefulness with 30 min of sleep. Each sleep opportunity was preceded and followed by the SSS. The correlation between the sleep latency and the SSS led to the idea that repeated measures of sleep latency throughout the day could provide an objective and reproducible measure of sleepiness.[10]

The MSLT procedure was formalized in 1977 to measure sleepiness in young normal subjects involved in sleep deprivation experiments.[21, 22] Six volunteers (age 18 - 21) underwent two nights of total sleep deprivation. They were put in bed every two hours during the wake period and told to try to fall asleep. Each test was terminated after 20 minutes if the subject did not fall asleep. If sleep occurred, the subject was awakened after one minute of stage 1 and the nap terminated to prevent accumulation of sleep during the sleep deprivation procedure. Sleep latency was measured from lights out to the first minute of stage 1. Significant correlations between the degree of deprivation and sleep latency gave face validity to using sleep latency as a biologically based measure of sleepiness.

Concern about the reliability of the SSS in patient groups quickly led to the use of the MSLT procedure in narcolepsy patients with an alteration in format that permitted 10 min of

sleep.[23, 24] The frequent occurrence of REM sleep in narcolepsy patients during the MSLT suggested a relationship between sleep onset REM periods (SOREMPs) and narcolepsy.[25, 26] The perceived usefulness of the MSLT quickly expanded to include evaluation of various sleep disorders[27, 28] effects of treatment,[4, 29] and effects of hypnotic medications.[30]

The use of the MSLT for clinical and research purposes led to the development of standardized MSLT protocols.[31] Separate clinical and research protocols were established. The research protocol was designed to limit the amount of sleep that a subject is permitted by terminating each MSLT nap after some pre-specified criterion for sleep onset has been met. The clinical version of the MSLT was designed to allow each MSLT nap to progress to a pre-specified termination time (15 min after sleep onset) in order to permit an opportunity for REM sleep to emerge. This "permit to sleep" feature is a key in the clinical use of the MSLT and figures importantly in the differential diagnosis of conditions associated with the symptom of excessive sleepiness, including narcolepsy.

2.2 What Does The MSLT Measure?

The MSLT is thought to measure physiological sleep tendency in the absence of alerting factors.[10] It is based on the assumption that physiological sleepiness decreases sleep latency. That is, the tendency to fall asleep should increase as physiological sleepiness increases.

Accurate measurement of physiological sleep tendency using the MSLT requires the absence of alerting factors that contribute to "manifest sleep tendency." Manifest sleep tendency is influenced by a host of continuously changing environmental and internal factors, such as temperature, light, noise, activity, and motivation. Seldom are manifest and physiological sleep tendency equal. However, as the alerting factors are removed, manifest sleep tendency approaches physiological sleepiness.

It has recently been suggested that sleep latency is also a measure of one's ability to transition into sleep. This characteristic has been referred to as "sleepability".[32, 33] Young adults with normal alertness and mean sleep latency (MSL) values in the 6-8 minute range even after two weeks with time in bed increased to 10 hours have previously been identified as having "high sleepability without sleepiness".[33] Several studies have documented that patients with psychophysiological insomnia, who have decreased sleep at night, actually have MSL values that are significantly longer ("low sleepability") than those of matched control subjects.[34-36] Additionally, a significant negative correlation has been reported between total sleep time at night and MSL values on the following day.[34-36] The finding that shorter prior sleep time is associated with longer MSLT latencies is the opposite of what would be expected if the MSLT were only sensitive to prior total sleep. Additionally, studies have shown that MSL is sensitive to both state and trait levels of central nervous system arousal.[37]

2.3 Maintenance of Wakefulness Test (MWT)

As experience with the MSLT grew, its use was expanded to attempt to objectively confirm reports of clinical improvement in patients after various interventions.[38] However, in severely sleepy patients, the numeric changes in the sleep latencies on the MSLT did not appear to change as much as might be suggested by the self reports. It was suspected that the MSLT did not have sensitivity at the most severe levels of sleepiness to detect clinically significant differences in this range. For extremely sleepy individuals, doubling or tripling the very short sleep latencies on the MSLT did not reflect clinical differences. This may have represented a floor effect in the MSLT, which limited its usefulness in discriminating between the sleepiest individuals or detecting changes within these individuals. The MSLT was also criticized for a ceiling effect imposed by the 20-minute nap limit that reduced the usefulness of the MSLT to discriminate between or within the more alert individuals. Additional criticism focused on the setting of the MSLT, including laboratory environment and optimal sleep conditions, as not representative of workplace conditions where sleep tendency might be very different.

In response to some of these criticisms, the MWT was developed in 1982 as an alternative to provide an assessment of ability to stay awake. The MWT procedure required the subject to sit up in a chair in a quiet and dimly lit room with instructions to stay awake. Vocalizations and movements were not allowed. Four or five trials were given beginning 1.5 to 3 hours after awakening and recurring every two hours thereafter. Each trial was terminated after 20 minutes if no sleep occurred or immediately after sleep onset. It was reasoned that instructing patients to stay awake rather than to allow themselves to fall asleep was a more accurate reflection of their ability to function and maintain alertness in common situations of inactivity. Subsequent studies demonstrated significant pre- and post- treatment differences in initial sleep latencies on the MWT in patients with excessive somnolence.[39]

2.4 What Does The MWT Measure?

The MWT measures the ability to stay awake under soporific conditions for a defined period of time.[11] It is based on the assumption that the volitional ability to stay awake is more important to know in some instances than the tendency to fall asleep. Since there is no direct biological measure of wakefulness available, this same phenomenon is calibrated indirectly by the inability or delayed tendency to fall asleep, as measured by the same EEG-derived initial sleep latency employed in the MSLT. In a large study of patients with sleep disorders, MWT sleep latencies were found to have a low but statistically significant correlation with the MSL on the MSLT.[40] The low correlation may reflect the combination of underlying sleepiness and motivational arousal present in the MWT.

2.5 What Factors Influence The MSLT and MWT?

The MSLT and MWT are affected by physiological, psychological and test protocol variables. Physiological factors that have been shown to affect MSL include: age,[41-43]circadian rhythms,[44]previous sleep quantity,[45, 46]various drugs,[30, 47-54] sleep structure and its disruption secondary to sleep and medical disorders.[28, 55-60] In general, these studies have shown that MSL decreases following sleep deprivation,[45] after hypnotic ingestion,[30, 49, 51] after barbiturate ingestion,[54] and with sleep disruption or fragmentation.[28, 55-57, 59, 60] MSL is increased following ingestion of stimulants.[47, 52, 53, 61] Age differences in MSLT means have been reported among pediatric,[41] adolescent,[62] and adult populations,[43, 56] with prepubescent children having the longest latencies

and young adults having the shortest latencies. For shift workers and late sleepers, the timing of the MSLT is an important issue since the individual's circadian phase could affect sleep latencies and the occurrence of sleep onset REM periods.[44]

Sleep tendency can be affected by psychological factors including anxiety or tension[34, 63-65] and depression,[66, 67] common co-morbidities in sleep disorder patients. Anxiety and stress can prolong sleep latency[63, 64] whereas depression may shorten the sleep latency in some patients.[66, 67] For all individuals, the anticipation of going home can lengthen sleep latency or preclude sleep on the last nap trial as demonstrated by the control data analyzed in this paper. Sleep latency measures may also be modulated by internal factors such as psychophysiological arousal.[59, 68] Several studies have shown that patients with psychophysiological insomnia have longer latencies on the MSLT.[34-36] These longer latencies have been attributed to increased central nervous system arousal.[34, 36, 55, 64]

The results of the MSLT and MWT can also be sensitive to variations in the testing protocols. In general, a basement effect in the MSLT and a ceiling effect in the MWT can limit discrimination of the sleepiest and most alert individuals, respectively. Additionally, the MSL can be increased by activity prior to testing[37, 69] caffeine use,[63] age,[43] and instructions.[29] Analysis of the normative data in this review also showed a significant effect of a 4 versus 5 nap protocol, prior nocturnal sleep quantity, and the definition of sleep latency used in calculating the MSL. In contrast, the current data analysis also shows that the MSL is not affected by administration of placebo. There is no significant difference between the clinical or research version of the MSLT on the resulting MSL.

2.6 Validity Data

The MSLT is considered the *de facto* standard for the measurement of sleepiness, since there is no other known biological measurement that directly reflects this clinical phenomenon. Consequently, evaluating the validity of the MSLT was complicated and required some unique approaches.

The MSLT is basically a formalized procedure for measuring sleep latency, so its validity is intrinsically related to the validity of the initial sleep latency as a measure of sleepiness. As stated in section 2.1 above, early studies correlated serial reports of sleepiness at individual naps with the MSL.[20] The MSL was also found to vary predictably with the amount of prior total sleep time.[45] Additionally, indirect studies with pharmacologic agents known to be soporific or stimulating produced expected changes in initial sleep latency.[70-72] Therefore, employing both convergent and divergent approaches, the MSL has face-value validity as a quantification of sleepiness, since it correlates with subjective reports, varies with prior total sleep time and reflects the sedating and alerting effects of medications.

In addition, the MSL appears to be sensitive to various forms of impaired sleep quality ranging from experimental sleep fragmentation[21, 46, 60] to clinical disorders of breathing such as apnea -in the untreated and treated states.[73-77] Thus, the MSL in the MSLT (and MWT) appears to be a marker for sleep quantity and quality and to have validity based on decades of reports.

2.7 Reliability Data

Reliability of the MSLT has been investigated for a number of variables including test-retest, number of naps, SOREMPs and scoring. The test-retest reliability of the MSLT has been examined using a four nap MSLT and reported a high test-retest reliability of 0.97 in normal healthy subjects on consistent sleep-wake schedules over a period of 4-14 months.[78] This reliability was not affected by the retest interval ($\leq$ 6 months vs. > 6 months) or the amount of sleepiness (MSL < 5 minutes vs. MSL $\geq$ 15 minutes). However, reliability was affected by the number of latency tests, decreasing to 0.85 for three tests and 0.65 for two tests. In patients with insomnia, the test-retest correlation on a five nap MSLT over a period of 3-90 weeks was 0.65.[79] One study examining test-retest reliability of SOREMPs in patients with narcolepsy found that 28 out of 30 patients had two or more SOREMPs on retest for a correlation of 0.93.[80] Scoring reliability of the clinical MSLT has also been shown to be quite high in a clinical population.[81, 82] Reported interrater reliability for MSL ranges from r = 0.85 - 0.90. Intrarater reliability for MSL was reported to be 0.87.[81] Evaluation of the presence of REM onsets during the MSLT showed an interrater agreement for more than one REM onset of 0.91 and 0.78 for intrarater agreement.[81]

2.8 Normative and Control Data Issues

It is important clinically to know the normative range of MSL values in order to distinguish those who are sleepy from those who are alert. Three studies have examined this issue in the MSLT. A study of 176 normal sleepers, including 129 young subjects (age 18-29) and 29 older subjects (age 30-80) reported a MSL of 11.1 in young subjects and 12.5 in older subjects.[43] Standard deviations (SDs) were not reported for the groups. Subsequently, this group data was reanalyzed and included in a comparative study where an overall MSL and SD of 11.5$\pm$5.1 minutes was estimated for the combined age groups.[83] The data as presented only allowed an estimate of the SD. One small database reported a median sleep latency value of 11.4 $\pm$ 4.6 minutes in 35 elderly patients[84] and another reported a MSL of 11.1 $\pm$ 4.6 minutes in 31 normal sleepers.[85] Consequently, there is no large systematically collected repository of normative MSLT data.

For the purposes of this review, an attempt was made to assemble a surrogate database of MSL values for the MSLT and MWT from the literature searches. "Control" data were identified from subjects who did not have pathology and/or who did not receive intervention. Unfortunately, in many studies prior sleep times, sleep quality, and caffeine use were not always clearly defined or used as criteria for entry into the study. Thus, these data essentially represent an actuarial database of individuals who may not have optimal sleep habits or sleep quality, but probably represent a valid contemporaneous normative group for comparison to experimental groups. Additionally, since MSLT/MWT methodologies were not always specifically described, variations outlined in 2.5 might have existed. Consequently, these subjects' data likely only approximate that of a large well-controlled systematic normative study.

2.9 Assessing Test Performance

Since the MSLT is the *de facto* standard, a problem arose in assessing its performance in the literature, especially regarding the diagnosis of narcolepsy. This involved the frequent tendency

of investigators to require a certain level of sleepiness (MSL) or REM pressure (as defined by the number of REM onsets in the MSLT) in order for individuals to gain entry in various studies. This introduced the presence of "incorporation bias" where the very parameters being tested were used as screening criteria (which can falsely elevate the sensitivity of the test). This bias precluded use of the data from subjects selected in this way. Thus, the Task Force discarded citations that included this methodology and used only those studies in which the clinical report of sleepiness or the tetrad of symptoms in narcolepsy was provided as the standard against which MSLT results were compared. Some classical measures of the performance of the MSLT in certain conditions were then possible and are reported in 6.4.

Sensitivity and specificity for the MSLT and MWT have been previously reported in patients diagnosed with narcolepsy.[83] MSLT, MWT and Epworth Sleepiness Scale (ESS) data were analyzed from a study of 530 narcolepsy patients.[86] All narcolepsy patients had to meet selection criteria that included a complaint of EDS, MSLT sleep latency criteria, and SOREMP requirements, although the latter were not strictly adhered to. The data were analyzed with that of normal subjects reported in other studies.[43, 87, 88] The MSL on MSLT had a sensitivity of 80.9% and specificity of 89.8% with a cut-off of <5 minutes. At a cutoff of < 8 minutes the sensitivity was 94.5% with a specificity of 73.3%. The MSL on the MWT had a sensitivity of 84% and a specificity of 98% using a cutoff of < 12 minutes. The ESS had a sensitivity of 93.5% and specificity of 100% with a cut-off score >10. Although the MSLT data from the narcolepsy patients clearly includes the type of incorporation bias described above, the data from the ESS, a subjective sleepiness scale, also raises concerns about incorporation bias since the narcolepsy patients were required to have a (subjective) complaint of EDS.

Sensitivity and specificity have also been reported for two or more SOREMPs in the diagnosis of narcolepsy.[89] In a 1993 study, entry criteria for narcolepsy subjects required a MSL of ≤ 8 minutes on the MSLT. It was reported that more than one SOREMP during the MSLT had a sensitivity of 78.5% and specificity of 62% if cataplexy was also required to be present for diagnosis. It was also reported that a history of cataplexy had a sensitivity of 71.5% and a specificity of 66% when more than one SOREMP was also required on the MSLT.

The approach used in the present review of the literature was directed at avoiding the type of inclusion bias described in the assessment of test performance. In the calculation of sensitivity and specificity values for MSL, all data was excluded that set MSL criteria for entry into the study. Also, sensitivity and specificity data on SOREMPs in narcolepsy and non-narcolepsy groups was excluded if SOREMP prerequisites were used for study entry.

3.0 EVIDENCE GRADING

Grading the papers that were accepted as evidence presented another problem for the task force. There were very few citations that explicitly examined the diagnostic utility of the MSLT/MWT *per se* using classical parameters such as sensitivity, specificity, and predictive values for various diagnoses. Moreover, published guidelines for assessment of diagnostic tests are still evolving and systems had to be adapted for this purpose.[90] The Task Force applied a grade based on the original purpose of the study for all topics except normative data. For the normative data, a grade of C5 was assigned to all articles since the design of the studies was unrelated to the normal control data obtained.

4.0 OTHER TESTS OF SLEEPINESS

A variety of subjective tests have also been developed to measure sleepiness. The SSS and ESS are the most common subjective tests of sleepiness. The SSS was introduced in 1972 and consists of a checklist of adjectives that are used to describe the subject's sleepiness at that moment.[17] However, its reliability in chronically sleepy patients is uncertain.[4, 38] The ESS is the most commonly used subjective test for the assessment of sleepiness. It was first described in 1991 and requires rating the tendency to fall asleep in eight common situations.[91] Subjective tests are thought to be particularly sensitive to extraneous factors such as motivation, recall bias, education level, and fatigue rather than simply the propensity to fall asleep.

The relationship between objective and subjective methods of sleepiness assessment is not clear. Correlations between scores on the SSS and the mean MSLT latency have been found to be low, whereas moderate but statistically significant correlations were reported between ESS and mean MSLT latency.[91-93] The variability and inconsistency has been attributed to measuring different aspects of sleepiness, measurements being affected by the test process and measuring sleepiness using different timeframes. It is beyond the planned scope of this paper to systematically compare objective and subjective tests for sleepiness.

5.0 METHODS

An evidence-based approach to defining and assessing the body of literature was used for this review and is outlined below.

5.1 Literature Searches

An initial Medline search for articles on the MSLT and MWT was performed on January 7, 2000. The following string of search terms was used: Multiple Sleep Latency Test or MSLT or Maintenance of Wakefulness or MWT or hypersomnia or hypersomnolence or daytime alertness, daytime sleepiness or daytime wakefulness or daytime somnolence. The search was limited to human research and only to articles published in English, French or Japanese. This generated a total of 3,162 references. There were 2,864 in English, 199 in French and 99 in Japanese. Similar searches on Psych Lit and Carl UnCover databases revealed no additional articles. An initial screening excluded the following: publications prior to 1976, articles in non-peer reviewed journals and book chapters. This resulted in 2195 potentially relevant articles.

All articles were screened based on title and English language abstract, if provided, to select studies that employed the MSLT or MWT. The inclusion criteria were: publication in peer-reviewed journal and use of the MSLT or MWT. The exclusion criteria included: abstracts, reviews, theoretical papers, editorials, case studies, drug studies and HLA studies. Studies using narcolepsy patients who were required to have SOREMPs on a previous MSLT were also excluded. Studies where the results of the MSLT were not given were also excluded. This resulted in a total of 495 potentially relevant articles. A second screening of the titles and abstracts not selected was performed by two other reviewers and an additional 150 papers were added to the potentially relevant list of articles for a total of 645 (495+ 150) papers for subsequent review.

On October 18, 2000, an additional Medline search was done to update the initial one using the same terms and conditions for the time period 1999-2000. This resulted in an additional 122 new citations. Review of these titles and abstracts, using the same inclusion and exclusion criteria as above, resulted in an additional 28 potentially relevant articles for a total of 673 (645+28). "Pearling", the process of manually scanning bibliography lists from manuscripts captured on the searches for additional relevant references not detected by Medline, netted an additional 15 citations. This resulted in a total of 688 (673+15) references.

On August 28, 2002, another Medline search was performed using the same terms and conditions for the time period 2000-2002. This resulted in 434 citations being identified. Using the same exclusion-inclusion criteria, 34 potentially relevant articles were identified for a total of (688 + 34) 722 references.

On October 9, 2003, a final Medline search was performed using the same terms and conditions for the time period 2002-3003. This resulted in 56 citations being identified. Using the same exclusion-inclusion criteria, 56 potentially relevant articles were identified for a total of (722+56) 778 references.

The titles and abstracts of the 778 papers were each screened by two reviewers to select those with adequately described data relevant to one of eight topics. Two topics were randomly assigned to each reviewer. Inclusion criteria were: study appropriate to topic, test means and SD or standard error of the means (SEM) provided. Exclusion criteria included non-standard test procedure. If there was not enough information in the title and abstract to make a determination, the article was obtained and examined in more detail to determine acceptability. The topics and numbers of references (in parentheses) identified in each group for further review included:

1) number of sleep-onset of rapid eye movement periods (SOREMPs) in narcolepsy (28)
2) sleep latency mean on MSLT in narcolepsy (40)
3) sleep latency mean on MSLT or MWT in idiopathic hypersomnia (13)
4) sleep latency mean on MSLT or MWT pre- and post-treatment of obstructive sleep apnea (OSA), periodic limb movement disorder (PLMD), insomnia and restless legs syndrome (RLS) (106)
5) sleep latency mean on MSLT or MWT in medical and neurological conditions (45)
6) sleep latency mean on MSLT or MWT on safety (33)
7) MSLT or MWT in pre- and post-drug treatment (74)

The pre- and post-drug treatment topic was added after all articles had been screened to exclude drug studies, so all 2,195 citations were again screened for this topic. Normative data was not pre-selected but obtained from other articles when found.

5.2 Final Selection and Data Extraction

All papers for each topic had two reviewers with a third reviewer to resolve differences. Only articles rejected by all reviewers on a topic were eliminated from further consideration. Full-length articles were obtained and examined for all selected references in each topic.

Data extraction sheets were developed prior to review of the articles. These included the following information: study design, number and gender of subjects, subject selection criteria, definition of sleep latency, MSLT/MWT methodology, means and SD or SEM by groups, biases and conclusions. A separate data sheet was developed for the sleep onset REM periods (SOREMP) and sleep latency data in narcolepsy that included the above information along with diagnostic criteria and number of SOREMPS.

For the first two topics, SOREMP and sleep latency in narcolepsy, the articles were randomly assigned to all reviewers to provide training in extraction of data. The extraction data were discussed among reviewers and discrepancies resolved. For the remaining topics, two reviewers, including one different from the previous reviewers for that topic, were assigned. Due to the size of the reference file only two individuals reviewed the pre- and post-drug topic. All discrepancies were resolved between reviewers or by a third reviewer and the data were entered into the evidence tables.

For the topic of narcolepsy, an attempt was made to include MSLT data only from patients whose diagnosis was clinically based rather than dependent on MSLT results. However, the requirement of a predominantly clinical diagnosis of narcolepsy was complicated. If "ASDA" or "ICSD" diagnostic criteria were cited, there was ambiguity as to whether the MSL was necessary or not for selection of cases. In these instances, papers were selected if diagnoses were apparently based upon clinical manifestations and, especially, if cataplexy was present. The 1979 ASDA nosology does not set MSL but only sleep onset REM periods (SOREMPs) as diagnostic features of the MSLT in narcolepsy.[94] The 1990 ICSD includes "mean sleep latency of less than five minutes" as one of four polysomnographic (PSG) features that would meet diagnostic criteria,[95] but the presence of cataplexy obviates the need for a PSG component.

From the original Medline literature search for the topic of MSLT, 130 articles were selected for review under the topic of narcolepsy. Articles were eliminated from consideration if sample sizes were less than five subjects, if mean and standard deviation were not reported, and if MSL data were clearly indicated as inclusion criteria for the diagnosis of narcolepsy in the study. Data from drug studies were not considered.

Forty articles were reviewed and extracted independently by two reviewers for the topic of mean sleep latency on the MSLT in narcolepsy. Inclusion criteria were: narcolepsy patients diagnosed based on predominantly clinical criteria, use of standard MSLT methodology, sleep latency data with means and SD or SEM. Exclusion criteria were: case studies, drug studies, pediatric patients, studies using the same data from previously cited publications, diagnosis or diagnostic criteria which were unclear or diagnosis that was based on sleep latency. Applying these criteria, 13 articles reported mean sleep latency data for narcolepsy in cases where diagnostic criteria were predominantly clinical and the data were entered into the evidence table.

Full-length articles from the 28 previously selected titles for SOREMPs in narcolepsy were reviewed and extracted independently by two reviewers. Inclusion criteria were all of the following: clinically diagnosed narcolepsy patients, five or more narcolepsy patients, MSL and SD or SEM, standard MSLT methodology, number of SOREMPs for narcolepsy patients. Exclusion criteria were all of the following: case studies, drug studies, studies using the same data from previously cited publications, pediatric patients, unclear diagnosis or diagnostic criteria, use of stimulant medications or if SOREMPs were used to define narcolepsy patients. Discrepancies on worksheets were resolved by a third reviewer. This process resulted in ten articles

for analysis and inclusion in the evidence tables.

From the original Medline literature search for the topic of MSLT, 13 articles were selected for review under the topic of idiopathic hypersomnolence. Articles were eliminated from consideration if sample sizes were less than four subjects, if MSL and SD data were not reported or if MSL data were clearly indicated as inclusion criteria for the diagnosis. Data from drug studies were not considered. This resulted in four articles being used for analysis and inclusion in the evidence table.

The selection of articles for MSLT or MWT in pre- and post-drug exposure resulted from a review of the initial 2,195 references using the following search terms: modafinil, Ritalin, pemoline, caffeine, flurazepam, hypnotics, triazolam amphetamine, benzodiazepines, antihistamine, and ethanol. In addition, articles were selected from a reading of all titles and abstracts if they met all of the following inclusion criteria: standard MSLT/MWT methodology, performance of pre- and post-drug MSLT/MWT, initial sleep latency data with means and SD or SEM, and clearly defined subject population. Articles were excluded if they met any of the following criteria: non-standard MSLT/MWT methodology, studies using the same data from other already cited publications, sleep subjects using other medications, case studies or less than five sleep subjects per group. This initially resulted in 45 full-length articles being obtained for review and extraction. However, due to the variety of drugs and dosages used, it was determined that only the flurazepam (four articles), caffeine (five articles) and modafinil (three articles) had an adequate number of subjects, using comparable dosages with MSLTs during the day, to be selected for data analysis. Three studies of modafinil used the MWT. Some articles included data for more than one group, so this resulted in data from 13 studies being entered into the evidence table.

Articles concerning the use of the MSLT/MWT in medical and neurological disorders were reviewed and selected from the 45 previously screened references for this topic. Selected articles were required to meet all of the following inclusion criteria: use of standard MSLT/MWT methodology, initial sleep latency data with means and SD or SEM and clearly defined patient population with a specific medical or neurological disorder. Articles were excluded if they met any of the following criteria: less than five sleep subjects per group, groups consisted of other family members or there was no screening for underlying sleep pathology. This resulted in 18 articles meeting criteria. Only six articles had control groups; the others provided only descriptive MSLT data for a single diagnostic group. None of the articles were judged useful for inclusion in the evidence tables.

Articles for data analysis concerning the use of the MSLT/MWT in OSA, insomnia, RLS and PLMD were reviewed and selected from the 106 previously screened references for this topic. Selected articles were required to meet all of the following inclusion criteria: initial sleep latency mean and SD or SEM, standard MSLT/MWT methodology. Articles were excluded if they met any of the following criteria: less than five sleep subjects per group or continuous positive airway pressure (CPAP) testing after less than one week on CPAP. It was determined that only OSA had adequate data for analysis. This resulted in 20 articles for data analysis and inclusion in the evidence tables.

Articles for data analysis concerning the indications for use of the MSLT/MWT in safety were reviewed and selected from 33 previously screened references for this topic. Selected articles were required to meet all of the following inclusion criteria: initial MSL and SD or SEM, and standard MSLT/MWT methodology. Articles were excluded if they met any of the following criteria: less than five sleep subjects per group, or graphed data was unreadable. Safety-related themes included: accidents, incidents, operational errors, performance, occupational issues, work hours and scheduling, shift work, or time zone shifts. It was determined that only two areas, driving or phase-shifted sleep (e.g., night work) had enough studies that met the inclusion criteria. This resulted in seven articles for inclusion in the evidence tables.

The selection of articles for normative data analysis was not predetermined. As articles were reviewed, normative data were taken from any article regardless of the topic if they met certain criteria. Inclusion criteria required all of the following: random sleep subject selection, a normal or control group of five or more sleepers without sleep complaints, standard MSLT protocol or standard MWT methodology, reported age, gender, initial MSL and SD or SEM on MSLT/MWT, and healthy sleep subjects. Normative-control data were excluded if they met any of the following criteria: subject selection criteria required a minimum initial MSL, sleep subjects using over-the-counter or prescription medication, sleep pathology not screened, history of psychiatric illness, irregular sleep habits, complaints about sleep, habitual napper, excessive caffeine or nicotine use not screened. This resulted in data from 33 articles for analysis including 29 articles using the MSLT and seven articles using the MWT with some studies using both tests.

As part of the normative data overview, 150 articles were selected to examine methodological variability in the MSLT/MWT. Selected articles were required to meet all of the following inclusion criteria: MSL and SD or SEM, number of subjects, ages of subjects, gender of subjects, and description of subject recruitment. Articles were excluded if they used less than five sleep subjects per group, or control subjects selection required a minimum average MSLT or MWT initial sleep latency. This resulted in 77 articles for a general survey of methodological variability but only those used in the preceding analysis of normative data were entered into the evidence table.

All extracted articles were assigned evidence levels adapted from Sackett[90] based on their original design and intent. For all normative-control data, an evidence level of C5 was routinely assigned, since the design or purpose of the study was irrelevant for this data.

6.0 RESULTS

6.1 Evidence Tables

The evidence tables are presented in the appendix Tables I-VII by topic. These can be accessed on the web at http://www.aasmnet.org.

6.2 Descriptive Summary of Each Evidence Table Topic

The evidence tables are divided by the seven topics evaluated and only include articles meeting all selection criteria. The use of the MSLT in the diagnosis of narcolepsy was reviewed for MSL and number of SOREMPs in controls and clinically diagnosed narcolepsy patients. These data are included in Evidence Tables I and II. MSL data using the MSLT were examined in patients diagnosed clinically with idiopathic hypersomnia and are presented in

Evidence Table III. Data on the MSL from MSLT or MWTs performed before and after treatment of OSA are presented in the OSA pre- and post-treatment Evidence Table IV. The medical-neurological topic focused on MSLT and MWT findings in patient groups diagnosed with any specific medical or neurological disorder. The data were from very diverse groups and were not entered into an evidence table. The safety topic focused on the MSL results from shifted circadian periods as well as results from studies of simulated or actual automobile driving. Studies with data on shifted circadian periods are presented in Evidence Table V. The pre- and post-drug topic examined the MSL in the MSLT and MWT before and after the administration of hypnotic (flurazepam) or stimulating medication (caffeine or modafinil). Data from the three types of drugs are grouped by drug class in Evidence Table VI. The normative data topic examined the MSL and SD in control data for the MSLT and MWT. It also examined the effect of age and methodological differences on MSL including various sleep latency definitions, different protocols (clinical and research), nap termination criteria, number of nap trials, and the effect of prior total sleep time. This data are presented in Evidence Table VII. All MSL data are presented in minutes with mean $\pm$ SD unless specified otherwise.

6.2.1 MSL in narcolepsy

Thirteen papers using the MSLT were used as evidence for MSL in narcolepsy. These articles were judged to be reasonably free of the inclusion bias of a MSL criterion for diagnosis. Four of these articles included MSL data for a comparison group of normal control subjects.[23, 68, 96, 97] These papers included consideration of a total of 39 narcolepsy cases with weighted MSL of 3.0 ± 3.1. By comparison, weighted MSL for a total of 40 control subjects was 10.5 ± 4.6. This difference is significant ($p<.001$).

Nine articles using the MSLT were judged to be reasonably free of the inclusion bias of a MSL criterion but do not include any normal subject control data.[24, 28, 98-104] These papers include 255 patients with narcolepsy. The MSL was 3.1 ± 2.9.

From these data it appears that most patients with clinical narcolepsy have clear objective hypersomnolence with a MSL on the MSLT of less than the traditional limit of five minutes. However, using a five minute cut point, about 16% of narcolepsy patients would have a MSL above this cutoff whereas about 16% of normal control subjects would fall below this cutoff.

The MWT was used in one large study of 530 narcolepsy patients.[86] This study used MSL criteria from the MSLT for patient inclusion, but MWT data were not part of the patient inclusion criteria. The MWT used a 20 minute protocol and a sleep latency definition of three epochs of stage 1 or one epoch of any other stage of sleep. Results of the 20-minute MWT showed a MSL of 6.0 with a SD of 4.8 for patients with narcolepsy. Normal control data for a 20 minute MWT using the same definition of sleep latency were found in a paper by Doghramji.[88] That paper reported a MSL of 18.7 $\pm$ 2.6.

6.2.2 SOREMPs in the diagnosis of narcolepsy

There were ten studies that met criteria for data on SOREMPs for use of the MSLT in clinically diagnosed narcolepsy patients. All selected articles were judged to be reasonably free of SOREMP requirements for diagnosis or classification purposes. Diagnostic criteria were not always clear, but the presence of cataplexy was considered sufficient for a clinical diagnosis of narcolepsy. The data show that SOREMPs are very common in narcolepsy patients[28, 68, 89, 100, 102-106] with the majority of narcolepsy patients having more than two SOREMPs.[24, 28, 103-106] The presence of SOREMPs was not linked to the presence of cataplexy.[89, 105] SOREMPs did not occur exclusively in narcolepsy patients but were found frequently in patients with OSA.[102, 106] The number of SOREMPs was found to increase with decreasing MSL on the MSLT.[89, 106] Analysis of the articles with appropriate data showed a sensitivity of 0.78[24, 28, 68, 89, 102, 103, 105, 106] and a specificity of 0.93[68, 102, 105, 106] for two or more SOREMPs for the diagnosis of narcolepsy. Sensitivity and specificity numbers were also recalculated without data from the study by Aldrich[105] due to the fact that these measures do not respond well when large numbers of comparison subjects meet criteria. The high false positives in that study may be due to the fact that children were included (age range 6 - 79) and a disproportionately large comparison group of 1,251 patients had sleep-related breathing disorders with 7% demonstrating two or more SOREMPs. Without the Aldrich paper the sensitivity was 0.79 and the specificity was 0.98 for two or more SOREMPs for the diagnosis of narcolepsy.

One study that did not make inclusion criteria looked at SOREMPs on the MWT in a small group of narcolepsy patients and normal controls.[11] The MWT protocol allowed ten minutes of sleep if subjects (Ss) fell asleep while sitting up. Interestingly, normal controls did not have any SOREMPs in the five trials whereas narcolepsy patients had an average of 3.2$\pm$2 SOREMPs. It should be noted, however, that the initial diagnosis of narcolepsy in these patients included MSLT criteria.

6.2.3 MSL in idiopathic hypersomnia

Four articles reported mean sleep latency data for idiopathic hypersomnolence (IH) in cases where diagnostic criteria were not entirely dependent upon MSL.[28, 107-109] One paper reviewed a large clinical database of hypersomnolence cases in which clinical diagnosis was emphasized and report that a diagnosis of IH was assigned to only 1% of the cases resulting in a ratio of IH to narcolepsy of 16:100.[107] Additionally, one study distinguished two types of IH.[108] Polysymptomatic IH was defined as a long nocturnal sleep requirement with difficulty awakening while monosymptomatic IH involved excessive daytime sleepiness alone. The MSL for each group was 10.4 ±5.2 and 7.8 ± 3.9, respectively. Overall, these papers include a total of 92 cases with a weighted MSL of 6.2 ± 3.0. This value appears to be intermediate between those for narcolepsy patients and those for normal control subjects.

6.2.4 MSL in pre- and post- treatment of OSA

MSLT in OSA patients and normal controls

Two papers met inclusion criteria, but they seemed to include the same OSA patients.[75, 110] It is unclear if the normal controls in one are a subset of the normal controls in the other. The normal control MSL was 12.8 ± 4.1, and the OSA MSL was 7.2 $\pm$ 6.0, which was between 1 and 1.5 SD less than the normal control mean.

MWT (40 minute protocol) in OSA patients and normal controls

One paper met inclusion criteria.[76] The MSL in normal controls

was 34.1 ± 5.5. In OSA patients, the MSL was 23.2 ± 10.2, almost 2 SD from the normal control MSL values.

MWT (20 minute protocol) in OSA patients and normal controls

One paper met inclusion criteria.[111] Normal control MSL on the MWT was 18.8 ± 3.3. In OSA patients, the MSL was 11.0 ± 5.6, which is more than 2 SD from the normal control MSL.

Effect of CPAP treatment on MSLT values in OSA

Ten papers met inclusion criteria.[73-75, 112-118] Studies testing the effect of less than a week of CPAP were excluded. Nine of the ten papers showed statistically significant increases in MSL with CPAP. A weighted MSL before treatment was 6.8 ± 4.2, which is within 1 SD of normal control mean. Weighted MSL after treatment was 11.6 ± 5.3. Both pre- and post-treatment mean were within 1 SD of the normal control mean.

Effect of CPAP treatment compared to placebo on MSLT values in OSA

Four papers met inclusion criteria.[119-122] Only two of the four papers showed a longer MSL in OSA patients treated with CPAP than those treated with placebo. Weighted MSL on CPAP was 9.8 ± 4.5 compared to 8.3 ± 4.7 on placebo.

Effect of CPAP treatment on MWT (40 minute protocol) values in OSA

Three papers met inclusion criteria.[76, 77, 123] All showed statistically significant improvement in MSL on the 40 min MWT in OSA. Weighted MSL before treatment was 18.8 ± 9.9 compared to 26.3 ± 10 after treatment. Pre-treatment MSL was more than 1 SD less than the normal control mean. Post-treatment MSL was within 1 SD of the normal control mean.

Effect of CPAP treatment on MWT (20 minute protocol) values in OSA

No papers were found addressing this issue.

Effect of CPAP treatment compared to placebo treatment on MWT (40 minute protocol) in OSA

One paper met inclusion criteria and did not find a significant difference in MSL between the CPAP group (16.2 ± 19.6) and a placebo group (14.4 ± 8.5).[124]

6.2.5 Medical and neurological disorders

There were 18 studies that met initial criteria for use of the MSLT in patients with medical or neurological disorders. This included two studies on children[125, 126] and one that used both the MSLT and MWT.[127] The studies included 13 different disorders and a number of slightly different methodologies. Due to the diversity of disorders, none of the studies were entered into an evidence table.

Eleven studies had no comparison groups but reported MSLT data for the patient group studied. This descriptive MSLT data (mean and SD) included patients with Prader-Willi[128] (6.9 ± 5.03), muscular dystrophy[129] (7.3 ± 2.2), and post-traumatic narcolepsy/closed head injury[130] (3.34 ± 1.87), rheumatoid arthritis[131] (16.1 ± 2.6), chronic obstructive pulmonary disease[132] (11.0 ± 3.7) and congestive heart failure - Cheyne Stokes breathing pattern[133] (11.3 ± 4.8 and 4 ± 1.1). A study of 30 epilepsy patients showed an overall MSLT MSLlatency of 8.4 ± 8, and when the epilepsy patients were divided into clinically sleepy and non-sleepy groups there was a significant difference between groups (7.6 ± 7.1 and 10.3 ± 9.5 respectively).[134] A study on Parkinson's patients found that those with REM onsets were significantly sleepier than those without REM onsets (4.6 ± 0.9 and 7.4 ± 0.7 respectively).[135] Patients with chronic renal failure showed no difference in MSL before and after hemodialysis (6.5 ± 4.1 and 6.3 ± 5.7 respectively).[136] Another study of hemodialysis patients showed a MSL of 10.2 + 4.2 on non-dialysis days.[137] A study of children with beta-thalassemia and congenital deserythropoietic anemia reported MSL of 7.8 + 3.5 and 10.7 + 7.5 respectively.[126]

Six studies using the MSLT included control groups. One study reported that muscular dystrophy (MD) patients were significantly sleepier than controls (13.5 ± 3.57 and 18.3 ± 1.82, respectively).[138] A study on patients with retinosa pigmentosa found that they were significantly sleepier than controls (10 ± 5 and 17 ± 3 respectively).[127] Another study found no difference between multiple sclerosis patients and controls (16.2 and 15.86, respectively).[139] A study on symptomatic and asymptomatic Ss exposed to the rhinovirus showed no differences between groups (7.84 ± 3.6 and 8.75 ± 2.8).[140] A study on post-traumatic stress in Viet Nam veterans reported a mean of 13.5 ± 4.1 in post-traumatic stress patients compared to 12.7 ± 4.4 in controls, but no statistics were done.[141] One study on children with attention deficit hyperactivity disorder showed that they were sleepier than controls (16.7 ± 5.4 and 18.9 ± 3).[125] No statistics were done on overall means, however patients were significantly sleepier on each of the first 3 naps than controls.

One study was found that used the MWT in a patient group. Patients with retinitis pigmentosa were significantly less alert than controls (21 ± 9 and 29 ± 2).[127]

Two drug studies are also noteworthy. Patients with epilepsy who were taking a sedating medication (phenobarbital) were shown to be significantly sleepier on the MSLT than controls or patients taking a non-sedating medication (carbamazapine) (MSL of 9, 12.9 and 12.5 min respectively).[54] In contrast, a study of Parkinson patients treated with modafinil showed no changes in MWT values after treatment.[142]

There were few studies that met criteria to evaluate the usefulness of the MSLT in patients with medical or neurological disorders typically associated with excessive sleepiness. Most studies were descriptive and lacked control groups. These studies used a widely accepted latency of greater than ten minutes as a control comparison to draw conclusions about sleepiness in the patient group. However, such a comparison may not be appropriate since the study on MD patients reported a MSL of 13.5 but indicated that they were significantly sleepier than the control group in the study.[138] In the two studies that used control groups for comparison to patients with disorders associated with sleepiness, only one found the expected result. The study of MD patients found that they were sleepier than controls, whereas a study on multiple sclerosis patients found no difference between patients and controls.[139] The strongest support for the usefulness of the MSLT in patients with neurological disorders comes from a drug study on

epilepsy patients demonstrating a significant decrease in MSLT latencies in patients given sedating anti-epileptic medications compared to controls or patients given non-sedating anti-epileptic medication.[54] However, these groups also differed by type of epilepsy (generalized versus partial). Overall, there is very little data concerning the use of the MSLT in patients with medical or neurological illnesses.

6.2.6 Assessment of safety

Two major safety themes emerged among the final selection of 16 articles. One set of studies examined sleep latencies in response to sleep that was shifted away from the normal circadian phase position. Another set of studies examined sleep latencies in association with simulated or actual driving. One miscellaneous study used a modified MSLT procedure to examine daytime sleep latencies and perception of being asleep.[143]

Phase-Shifted Sleep: Sleep Medications and Simulated Night Shift

The studies of phase-shifted sleep measured sleep latency associated with changes in sleep quantity or quality, circadian arousal, or treatments with sleep medications or bright light. The implications of these studies for safety are based on the assumption that increased sleepiness from phase shifting may increase the probability of operational errors or performance decrements. Among the studies of phase-shifted sleep in the initial sample, nine articles measured sleep latencies in subjects remaining awake overnight in the laboratory during a simulation of night shift work. Of those articles, six met the criteria for final inclusion. Another four studies measured sleep latencies in subjects who traveled to a different time zone and slept at a time that was different from their home sleep circadian phase. None of those articles met the criteria for final inclusion.

In all selected studies of phase-shifted sleep, a "research" MSLT protocol was used, requiring subjects to be awakened at the first signs of EEG sleep. Among the studies of overnight wakefulness in the laboratory, only one directly compared multiple nighttime sleep latencies to daytime latencies in the same subjects.[144] The protocol for that study included administration of placebo, triazolam (0.25 and 0.50 mg), or flurazepam (15 or 30 mg) prior to daytime sleep before the overnight vigil. Results indicated that overnight latencies were lower than daytime latencies in all placebo and drug conditions and lowest following doses of flurazepam, a long-acting benzodiazepine. Average differences between daytime and nighttime latencies ranged from 2.3 to 3.8 minutes across placebo conditions and 0.5 to 6.6 minutes across medication conditions. Average differences between daytime and nighttime latencies ranged from 0.5 to 2.8 minutes across all conditions. Latencies across the night shift in all conditions ranged from approximately 15 - 18 minutes at the beginning of shifts to two to three minutes around 6:00 to 8:00 AM (values estimated from the figures).

Across all studies of simulated night shift the mean overnight MSLT scores during no-drug or placebo conditions ranged from 7.2 to 10.3 minutes with standard deviations between 3.2 and 5.9 minutes.[144-150] Following medication treatment for daytime sleep, overnight MSLT scores ranged from 5.8 to 12.1 minutes with standard deviations of 1.9 to 4.8 minutes. Drug minus placebo differences ranged from -3.8 to 2.0 minutes across studies (i.e., nighttime MSL scores both increased and decreased relative to placebo nights). Total daytime sleep time prior to the overnight vigil ranged from 190 to 214 minutes in placebo conditions and 300 to 320 minutes in drug conditions.

Four overnight studies also employed the MWT. Latencies on those tests during no-drug or placebo conditions ranged from 12.2 to 18.3 minutes with standard deviations ranging from 2.8 to 7.4 minutes. Following medication treatment in three studies, latencies on the wakefulness tests ranged from 13.9 to 19.2 minutes with standard deviations ranging from 1.3 to 5.1 minutes. Drug minus placebo differences ranged from 0.9 to 1.7 minutes across studies (i.e. slight increases in MWT MSL scores in the drug conditions).

Sleep latencies across the night following placebo/no-drug or benzodiazepine (temazepam or triazolam) administration during the day are shown in Figure 1. Values in the figure are the weighted average of the reported means of each study at each time point.[145-148, 150] In placebo/no-drug conditions, MSLT scores ranged from 13.3 minutes at the beginning of the shift to 4.6 minutes in the early morning hours. During drug conditions, MSL

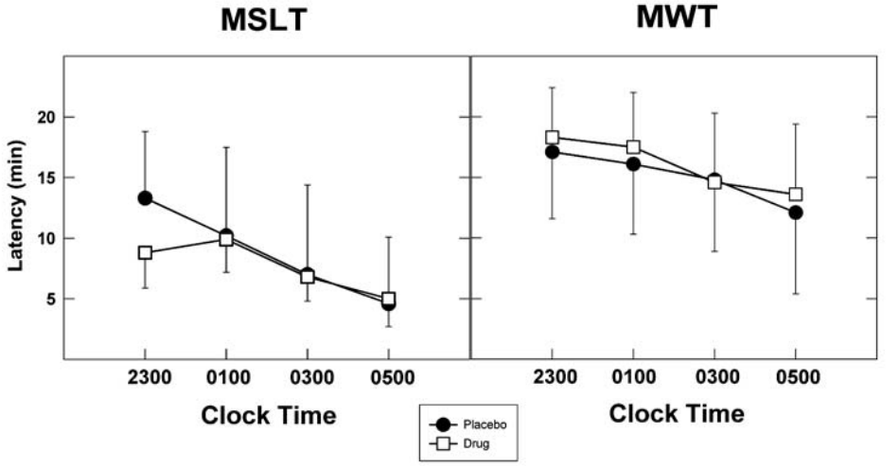

Figure 1—Sleep latencies during simulated night shifts following daytime sleep with administration of temazepam, triazolam, placebo, or no drug. Values at each time point are the weighted average of the reported means from several published studies. Error bars are standard deviations.

scores ranged from 8.8 to 9.9 minutes early in the shift to 5.0 minutes in the early morning hours. MWT scores overnight in placebo/no-drug conditions ranged from 17.1 minutes at the beginning of the shift to 12.1 minutes in the early morning hours. MWT scores overnight in drug conditions ranged from 18.3 minutes at the beginning of the shift to 13.6 minutes in the early morning hours.

Phase-Shifted Sleep: Bright Light and Simulated Night Shift

Bright light exposure during the night shift increased average nighttime MSLT scores compared to dim light.[149] From early to late in the shift, latencies decreased from approximately 12 to 3 minutes under dim light and from 16 to 7 minutes under bright light (values estimated from the figures).

Phase-Advanced Sleep in Sleepy or Alert Subjects

One study used the MSLT to divide subjects into alert (latencies of 12 or more minutes) or moderately sleepy (latencies of ten or less minutes) and then tested the degree of sleep disturbance induced by phase-advanced sleep (earlier bedtimes) in each group.[151] Sleep was disturbed in both groups, but less disturbed in the sleepy group.

Driving

Ten studies of actual or simulated driving were identified and seven met the inclusion criteria. Long distance driving (one study), alcohol (one study), sleep disorders (two studies), or CPAP treatment (three studies) were examined in these studies. Four studies of sleep disorder patients employed a "clinical" MSLT procedure and a fifth study[152] did not specify a procedure. The remaining studies of driving used a "research" MSLT procedure.

Motor vehicle accident histories and MSL on the MSLTs were examined in one study entered into the evidence table that reported no latency differences among groups having zero, one, or more than one accident in a five-year period.[153] Mean latencies in those groups ranged from 9.3 to 9.7 minutes. Similar results were reported in a study that was not included in the evidence table because standard deviations were not reported.[154] In that study, MSLT latency scores did not differ among sleep disorders patients who had motor vehicle crashes and those who had not. Group mean sleep latencies among apnea, narcolepsy, and other sleepy EDS patients with a crash history ranged from 4.8 to 9.4 minutes, whereas group means among patients without a crash history ranged from 3.3 to 8.2 minutes.

One study evaluated MSLT and simulated driving performance (tracking and divided attention) among narcolepsy and sleep apnea patients tested in the laboratory.[110] Compared to a matched control group, simulated driving was significantly worse in both patient groups. Apnea group latencies (mean = 7.2 minutes) were significantly shorter than the control group (mean = 13.2 minutes) and the narcolepsy group latencies were shorter than both groups (mean = 4.9 minutes). Driving performance (tracking error) was negatively correlated with MSL in both patient groups ($r = -.42$ for apnea; $r = -.32$ for narcolepsy).

CPAP treatment for sleep apnea was associated with increased MSL and a reduction in the rate of self-reported motor vehicle crashes[73] or improved performance on a simulated driving task.[75] In the latter study, increases in MSL following treatment were significantly correlated with improvements in tracking performance on the driving task ($r = .65$). CPAP treatment also was associated with increased MWT latency scores and improved actual driving performance among professional bus drivers.[76]

One study measured daytime sleep latencies and performance in a driving simulator following four or eight hours of sleep and ingestion of either ethanol or placebo.[155] Significantly longer latencies were observed after eight hours of sleep in the placebo group compared to all other conditions (10.7 ± 3.4). Latencies following four hours of sleep and either ethanol or placebo, or eight hours of sleep with ethanol did not differ (MSL = 4.7 vs. 6.3 minutes, SDs = 3.1 to 4.7). Simulated driving performance was worst in the 4-hour ethanol group but no correlations were reported between driving performance and sleep latencies.

Sleepiness among long distance drivers was measured at a highway rest stop in one study by using two latency tests taken in rapid succession (10-minute interval).[156] Long distance drivers had greater sleep debts and shorter latencies than a control group who had not driven and had normal amounts of sleep (Test 1 = 11.5 vs. 15.5 minutes; Test 2 = 14.0 vs. 18 minutes, respectively). Total sleep times and latencies were positively correlated in the driving group.

6.2.7 Pre- and Post-Drug Studies - Validity

Studies using the MSLT after administration of medications were examined to assess the face validity of the MSLT. Specifically, one would expect the MSL to become longer after administration of stimulants and to become shorter after administration of sedatives. A broad range of studies in which the MSLT was used after stimulant or sedative administration were reviewed. Multiple studies meeting the general criteria of this review were found for the MSLT after caffeine, modafinil, and flurazepam administration. Multiple studies meeting the criteria of the review were found for the MWT after modafinil administration.

Three studies used the MSLT after modafinil and placebo administration.[71, 157, 158] A total of 131 Ss received placebo and 127 Ss received modafinil in the these experiments. The weighted mean and standard deviations for the placebo and modafinil groups across the three studies was respectively 5.0 ± 3.7 and 7.4 ± 4.2 minutes ($t = 4.869$, $p < .001$). Five studies were identified where caffeine and placebo were administered prior to the MSLT.[63, 72, 159-161] In these studies, which involved 91 Ss, the weighted means for the MSL after placebo and caffeine administration were 9.6 ± 4.1 and 13.8 ± 3.8 minutes ($t = 6.676$, $p < .001$). Finally, four studies were identified where flurazepam or placebo was administered the night before the MSLT was given.[49, 70, 72, 162] In these studies, 55 Ss received placebo and 56 Ss received flurazepam. The respective MSL for placebo and flurazepam were 11.7 ± 4.8 and 8.9 ± 4.7 minutes ($t = 3.104$, $p < .01$). These data suggest that the MSL is characteristically increased after the administration of stimulant medication and reduced after the administration of sedating medication, providing support for the validity of the MSLT.

Three studies employed the MWT after modafinil and placebo

administration.[71, 158, 163] A total of 194 Ss received placebo and 190 Ss received modafinil in these experiments. The MSL for the placebo and modafinil groups across the three studies were 8.2 ± 5.4 and 10.4 ± 6.3 minutes respectively ($t = 3.672$, $p < .001$). Analysis of the two studies using narcolepsy patients[158, 163] showed similar results following administration of placebo (164 Ss) and modafinil 200mg (160 Ss) with MSL 6.62 ± 5.66 and 9.13 ± 7.0 respectively ($t=2.38$, $p<.01$). These data suggest that the MSLT is characteristically increased after the administration of stimulant medication and therefore provide support for the validity of the MWT.

6.2.8 Normative data

The issue of normative MSL data for the MSLT and MWT is complicated because many factors affect the results and must be taken into consideration. Factors that have been reported to affect the MSL include age and prior total sleep time (TST), as well as many variations in the methodologies such as number of naps, SL definitions and termination criteria whose effects are not clear. These minor variations in protocol are common but difficult to track. A review of sleep latency definitions and termination criteria in 77 articles that used the MSLT protocol showed that only 35% reported using the recommended sleep onset definition[31] (i.e. first epoch of Stage 1) whereas 22% specified other SL definitions. Many articles (43%) did not specify any SL definition or include a methodology reference. The findings were identical for nap termination criteria. Moreover, the stated methodology was not always strictly followed in some studies. Given the methodological variability underlying the published literature, data from normal control subjects (Ss) was analyzed to evaluate the effect of many specific factors on MSL. This provided the basis for combining data when no significant differences were found and provided the support for calculating normative values for relevant factors for each test.

The MSLT database reviews were examined for all papers that reported MSLT means and standard deviation from at least four naps placed in the morning and afternoon for normal or control Ss. Inclusion was based upon review of the data table of reviewer notes for abstracts plus an additional three studies identified by the reviewers. Weighted means are reported. Analyses containing SD are based upon pooled (and weighted) variance. The data are based on 27 papers, some with multiple groups. Studies with normal Ss given placebo were included and the data combined with normal controls not given placebo because no significant difference in MSL was found in narcolepsy patients if placebo was given or not given.[164] However, no active drug data are included.

One study reported a significant effect for age.[43] Since it was unknown if results differed as a function of clinical vs. research version on the MSLT, data were first tabulated by age and test version (Table 1). The differences between clinical and research versions did not differ significantly at any age range except for 30's (clinical = 10.4, research = 12.5; $t_{190}= 2.60$, $p = .01$). These findings contradict the hypothesis that accumulated sleep in clinical MSLTs would result in increased MSL. Therefore, Table 2 presents combined MSLT values by age. The data are also plotted in Figure 2. As expected from the previous literature, significant age-related differences were found with t-tests and are summarized in Table 2. Eighty-year-old Ss had longer MSL than all other age groups. Fifty-year old Ss had longer MSL than all groups under age 40. This analysis predicts that MSL values will increase by approximately 0.6 minutes per decade.

Mean values from all studies doing five naps were compared with the mean values from all studies doing four naps (studies with age 50 or greater not included). There were ten studies (with 12 total groups) with 372 Ss reporting four naps[34, 43, 78, 165-171] and 13 studies (15 total groups) with 284 Ss reporting five naps[22, 23, 68, 96, 97, 110, 172-178] The overall mean for four nap MSLT studies was 10.4 ± 4.3. The overall means for five nap MSLT studies was 11.6 ± 5.2. This difference was statistically significant ($t_{654} = 2.889$ $p < .01$). Using two SD from the mean, 95% of the normal values for the four nap MSLT would fall between 1.8-19 minutes while the values for the five nap MSLT would fall between 1.2-20 minutes.

The difference between scoring sleep onset to the first non-wake epoch of sleep versus scoring to 'consolidated' sleep (defined here as the lesser of the first of three epochs of stage 1 or the first epoch of non-stage 1 sleep) was also calculated. This analysis was done in one data set that had previously been scored with both criteria.[166] Based upon the definitions, it is not possible for the latency to consolidated sleep to be less than the latency to the first non-wake epoch. Therefore, it is not surprising that a significant difference was found. A repeated measures analysis of variance (ANOVA) with effects for time of nap (4) and scoring method (2) was performed. The F-value for scoring method was significant ($F_{1, 48} = 23.8$, $p < .00001$). The respective means were 7.4 and 8.1 minutes. Although statistically different, this small difference may not be clinically significant.

The effect of prior sleep time on MSLT results was also examined. Ten of the 27 studies (one with two groups) with prior sleep time greater than 435 minutes[22, 23, 34, 63, 166, 171, 176-179] were compared to 7 studies (one with three groups) with prior sleep time less than 425 minutes.[43, 133, 167, 169, 172, 173, 180] The mean total sleep time respectively was 451 and 397 minutes. However, it should be noted that four of the studies with short prior sleep time were performed with Ss in their 50's or 60's. The overall mean for the MSLT after long prior sleep was 9.8 ± 4.8. The overall means for the MSLT after short sleep was 11.2 ± 4.1. This difference was statistically significant ($t_{397} = 3.04$. $p < .001$). The difference was also significant when the four studies from older individuals were eliminated from the data set. These data indicate that MSL is longer when prior TST is shorter. The implication is that participants in studies reporting shorter TST probably did not suffer from partial sleep deprivation as a methodological flaw but were more likely just unable to sleep longer.

Based upon the significant difference found between the four nap and five nap protocol above, the age analysis was repeated and split for the four or five nap protocol. Due to concerns about specific end effects of the last nap, studies that had more than five naps were not included in the five nap group (even if individual nap data were available). Studies were also eliminated if there was a very broad age range of participants (without means). In the remaining groups, statistical comparisons could be done in three age blocks (20's, 30's, and 40's). A significant difference in the predicted direction was found for 20 year-old participants ($t_{270} = 1.984$, $p < .05$) and 30 year-old participants ($t_{165} = 3.059$, $p < .01$). The individual data are presented in Table 3 and the means are plotted in Figure 3. (with standard error of the mean, the lines

are the linear regression). The overall mean for the 4-nap studies reported in Table 3 was 10.3 ± 4.2. The mean for the 5-nap studies was 11.5 ± 5.3 ($t_{612} = 3.04$, $p < .01$).

Three articles using the MWT of 40 and 20 min where found which contained normative data. Two of these articles used MWT trial durations of 40 minutes[76, 88] and one used MWT trial durations of 20 minutes.[111] One of the 40 minute MWT studies also reported 20 minute MWT data derived from the 40 minute test and that is not included here but is reviewed below. The 40 minute MWT studies included a total of 74 normal subjects using four trials that were terminated one minute after sleep onset. The average MSL for both studies was 35.2 ± 7.8 when sleep latency was defined as three epochs of stage 1 or one epoch of any other stage. Using four naps and the same definition of sleep latency, the 20 minute MWT included 15 normal subjects (mean age 45.2) and reported a MSL of 18.8 + 3.3.

One of the 40 minute MWT studies had a large data sample of 64 Ss patients and reported MSL for 20 minute and 40 minute MWT using two definitions of sleep latency.[88] The 20 minute MWT data was derived from the 40 minute test. The MSL was calculated for sleep onset defined in two ways. One definition referred to the first epoch of any stage of sleep or the first consecutive ten seconds of stage 1. The second definition referred to sustained sleep defined as three epochs of stage 1 or the first epoch of any stage of sleep (SUMWT). The MSL to the first epoch or ten seconds of sleep on the 40 minute MWT and 20 minute MWT were 30.4 ± 11.2 and 18.1 ± 3.6 respectively. The MSL to sustained sleep were 35.2 ± 7.9 and 18.7 ± 2.6 for the 40 and 20 SUMWT, respectively.

The availability of the raw data from this data set allowed additional analysis to be performed to examine normative values for both the 40 minute (Table 4) and the 20 minute (Table 5) MWT scores.[181] Data were analyzed on 64 subjects who were considered as not having sleep problems. MSL was calculated across four naps using the latency to sustained sleep. Because these subjects did not have sleep problems, many of them scored the respective highest values of 40 and 20 for the average scores. On the 40 minute MWT 38 of 64 subjects (59%) did not fall asleep on any trial using the sustained sleep latency definition. Using the first epoch or ten seconds of stage 1 definition 27 of 64 subjects (42%) did not fall asleep on any trial. This ceiling effect results in the data not being normally distributed. Therefore Kaplan-Meier survival analysis was used to analyze the data. One patient (subject 39) had a very low score (asleep earlier) than the other

Table 1—Comparison of MSL for clinical and research MSLT protocols by age group.

	CLINICAL PROTOCOLS			RESEARCH PROTOCOLS		
DECADE	**Age**	**#SS**	**Mean SD**	**Age**	**#SS**	**Mean SD**
20'S	24	29	12.7+71[75]	21	12	10.6+6.1[34]
				29	14	11.1+71[79]
				23	49	7.4+31[71]
				25	12	8.0+3.11[80]
				24	16	11+31[69]
				26.5	89	10.9+51[77]
				23	18	9.8+4.61[73]
				25	18	14.5+7.81[70]
				29	16	10.3+3.31[76]
				28	11	8.4+6.81[78]
SUMMARY		**29**	**12.7+7**		**255**	**10.1+4.9**
30's	33.4	139	10+3.71[65]	39	10	9.5+5.3[63]
	34	17	13.4+41[74]	36	10	15.5+5.41[72]
				35	11	13.8+1.4[57]
				30	5	9.9+5.61[76]
SUMMARY*		**156**	**10.4+3.7**		**36**	**12.5+4.5**
40's						
	43	10	10.3+6.4[96]	44	11	12.3+1.2[57]
	45	11	10.4+6.3[68]	44	9	12.1+1.1[23]
	46	21	13.2+2.41[10]			
	49	10	10+6.6[97]			
SUMMARY		**52**	**11.4+5.1**		**20**	**12.2+1.2**
50's				54	11	12.1+1.1[57]
60's				65	9	12.4+11[33]
				66	13	11.9+1.6[57]
				68	32	10.5+6.51[67]
SUMMARY					**54**	**11.2+5.2**
80's				83	22	15.2+61[75]

* Means 30 year-old groups differ at the .01 level

Table 2—Summary and statistical comparison of MSLT MSL by age group

Age	Mean	SD	# SS	Note
10	10.0	4.5	25	
20	10.4	5.4	284	Differ from 50, 80
30	10.8	3.9	192	Differ from 50, 80
40	11.7	4.4	72	Differ from 80
50	12.1	1.1	11	Differ from 80
60	11.2	5.2	54	Differ from 80
70			0	
80	15.2	6	22	Differ from all

patients. Analysis was performed both with and without this patient. Potential effects due to three covariates were also considered. These were gender, age group and body mass index. For statistical purposes, age was divided into four 10-year age groups and BMI was divided into three categories of <24.5, 24.5-<29.5, and 29.5+ as factors in the survival analysis. The only statistically significant result was due to age group, but the only age group that was different was the 30-39 groups, who had shorter MSL than the older subjects. One other study[179] in a group of 14 subjects with a mean age of 29 $\pm$ 12.2 reported a sleep latency similar to that reported in the 30-39 year old group from Doghramji et al.[88] These results imply that the reported significant age effect for the MWT is probably not a chance finding. Further, these age results are similar to those reported for the MSLT (see Figure 4). Regression across age shows an increase in sleep latency of approximately 2.5 minutes per decade for the MWT. The absence of gender and body mass index affects on MSL is probably due to that fact that these 64 subjects do not have sleep problems, so the "usual" effects due to gender and BMI have already been removed by excluding patients with sleep problems.

Table 4 reports the basic descriptive information on the 40 minute MWT mean scores, overall and by the three covariate groups; Table 5 reports on the 20 minute MWT mean scores. The data for the tables were generated by the survival analysis. Figuress 5 and 6 provide survival plots using the sustained definition of sleep latency for the 20 and 40 minute MWT, respectively. Figure 7 provides a survival plot for the 40 minute MWT defining sleep latency as either ten seconds of stage 1 or the first epoch of any stage and Table 6 provides the percentiles for the MSL in Figure 7. So using Figure 6, for example, a MSL of 25 minutes on the 40 minute MWT shows 1 minus survival of approximately 15%, with +2 SEM of approximately 25%. This indicates that 15% of normal patients had a MSL of 25 or lower. Using the 95% confidence interval (+2 SEM), the actual percentage scoring 25 or lower may go as high as 25%. If a sleep patient has a MSL of 10 minutes, the percentage of normal patients with this score or lower is approximately 2-6%. One would assume that a patient with such a low MSL is not likely "normal".

7.0 DISCUSSION

Most studies that used the MSLT and MWT were not designed to address the issues in this review. Consequently, much of the data was excluded because it did not meet the criteria established to address the selected topics. In many cases, details were vague or omitted from articles. This places additional limitations on the use and interpretation of results reported. However, the data in this review represent the current evidence available, and the findings are generally consistent with a variety of studies.

This review revealed numerous variations in test methodology being used, along with many studies that omitted any description of the test methodology. Common omissions concerning methodology .included: amount of prior sleep time, definition of sleep latency, number of naps, age range, and use of research or clinical protocols. Even when a common reference was used to describe the methodology for the performance of the MSLT or MWT, it was sometimes evident from other information in the paper that the cited methodology was not always followed. Analysis of control data was done to determine the impact of these factors on MSL and it was found that prior total sleep time, definition of sleep latency, number of naps, and age resulted in clinically significant differences in MSL on the MSLT. There was not a significant difference on MSL between the clinical and research MSLT protocols. Given the variations in methodology and lack of detailed information in studies, it is unlikely that the summarized studies for each topic in this review had identical test methodologies.

Another methodological issue involved patient selection crite-

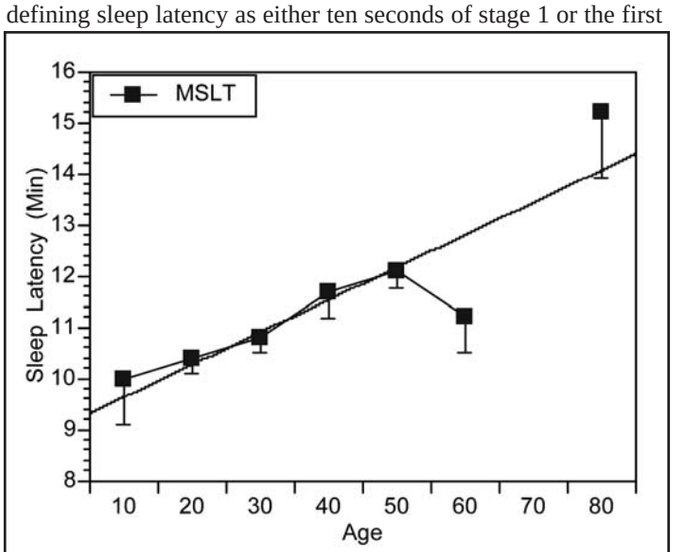

Figure 2—MSL with SEM on the MSLT by age. Linear regression line is plotted.

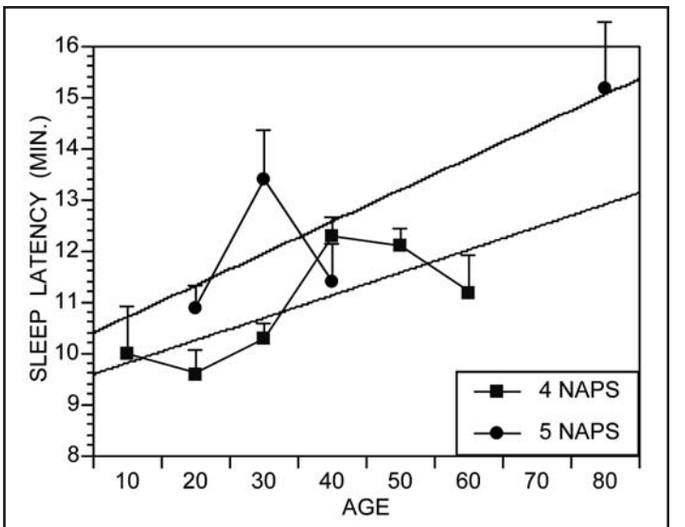

Figure 3—MSL with SEM on the MSLT in four and five nap protocols. Linear regression lines are plotted.

ria. The use of a questionnaire or clinical interview rather than polysomnography to screen for sleep disorders sometimes resulted in subjects being dropped from the study when sleep disorders were discovered or the finding was simply noted in the results and the data included. Drug usage in subject and patient populations was often only screened for stimulant or hypnotic medication through questionnaires or interviews. Urine screens were not typically performed and other classes of medication were often permitted in patient populations. The possible presence and effect of these extraneous variables on the MSL data cannot be determined although these presumably random variables would be expected to have minimum impact on the overall data.

Although there are some limitations with the data, the validity of the MSLT as a measure of sleepiness has strong support from drug studies. It was reasoned that SL should become longer following administration of stimulant medication and shorter following administration of sedative medication. Studies using the MSLT after caffeine, modafinil, and flurazepam administration showed that MSL is characteristically increased after the administration of stimulant medication and reduced after the administration of sedating medication. Studies using the MWT after administration of modafinil showed an increase in MSL. These data suggest that both the MSLT and MWT respond as expected following administration of medication and therefore provide support for the validity of the MSL in the MWT and MSLT. However, the results provide only indirect support for the MSLT since the test is the *de facto* standard and there is no other objective measure to use for comparison.

Determining normative MSL values is complicated. There are different values for different sample characteristics. However, the MSL seems to be relatively stable regardless of most minor variations. For the MSLT, the MSL is around ten minutes with a two SD range of about 2 - 19 minutes. This appears to be a normal distribution of scores. However, the SDs are relatively large for the MSLT, so the MSL does not discriminate normal and sleepy populations very well due to extensive overlap of MSL values. Moreover, the use of a SD cutoff may not always be appropriate for the MSLT and MWT since there may not be a normal distribution of some scores. For example, the MSLT is more susceptible to a floor effect in very sleepy individuals such as narcolepsy patients, whereas the MWT has a ceiling effect for the most alert individuals.

As can be seen from Table 3, MSL values varied from 7.4 to 15.2 minutes in the normal populations sampled with SDs ranging from 1.1 to 7 minutes. Sleep latencies were significantly different depending on a) the criterion used for sleep onset, although the actual difference between the first epoch of any sleep stage and 3 epochs was small; b) the nap protocol, with the 5-nap protocol having latencies about 1-.2 minutes longer (see Figure 3); and c) age (see Figure 2). These variables interact so that reported differences in latencies vary from 9.6 $\pm$ 4.8 minutes in 20 year olds in a 4-nap protocol to 15.2 $\pm$ 6 minutes in 80 year olds with a 5-nap protocol. Interestingly, the age effects found on the MSLT were replicated on the MWT. This emphasizes the impor-

Table 3—Comparison of MSLT MSL by age group in 4 and 5 nap protocols

	Four Nap Protocol			Five Nap Protocol		
DECADE	**Age**	**#SS**	**Mean SD Reference**	**Age**	**#SS**	**Mean SD Reference**
10's	15	25	10+4.61[66]			
20's	29	14	11.1+71[79]	23	18	9.8+4.61[73]
	23	49	7.4+31[71]	24	29	12.7+71[75]
	25	12	8.0+3.11[80]	26.5	89	10.9+51[77]
	24	16	11+31[69]	29	16	10.3+3.31[76]
	25	18	14.5+7.81[70]	28	11	8.4+6.81[78]
SUMMARY*		**109**	**9.6+4.8**		**163**	**10.9+5.4**
30's	33.4	139	10+3.71[65]	34	17	13.4+41[74]
	35	11	13.8+1.4[57]			
SUMMARY*		**150**	**10.3+3.6**		**17**	**13.4+4**
40's	44	11	12.3+1.2[57]	43	10	10.3+6.4[96]
				45	11	10.4+6.3[68]
				46	21	13.2+2.41[10]
				49	10	10+6.6[97]
SUMMARY		**11**	**12.3+1.2**		**52**	**11.4+5.2**
50's	54	11	12.1+1.1[57]			
60's	65	9	12.4+1.91[33]			
	66	13	11.9+1.6[57]			
	68	32	10.5+6.51[67]			
SUMMARY		**54**	**11.2+5.2**			
80's				83	22	15.2+61[75]

(*p<.05 indicating a significant difference between 4 and 5 nap means)

Table 4—Statistics for the Average of Four 40-Minute MWT

Group	n	Mean (SEM)v	75%tile (SEM)	# scoring 40+
Male	27	36.19 (1.12)	32.75 (1.78)	16
Female	37	34.54 (1.48)	32.75 (6.21)	22
Female, excl	36	35.31 (1.31)	32.75 (2.92)	22
Age 30-39	18	30.86 (2.04)	22.58 (2.53)	5
Age 40-49	20	36.52 (1.88)	NA	15
Age 40-49, excl	19	38.07 (1.18)	NA	15
Age 50-59	15	36.73 (1.59)	34.75 (5.35)	10
Age 60-69	11	38.03 (1.06)	NA	8
BMI <24.5	32	35.09 (1.38)	32.50 (1.76)	18
BMI 24.5-<29.5	26	34.88 (1.69)	31.63 (5.84)	17
BMI 24.5-<29.5, excl	25	36.00 (1.34)	32.75 (4.68)	17
BMI 29.5+	6	37.54 (1.17)	34.75 (2.31)	3
Total 64	64	35.24 (0.98)	32.75 (1.37)	38
Total 63	63	35.69 (0.89)	32.75 (1.46)	38

Excl: statistics excluding the potential outlier
NA: 75th percentile or its standard error cannot be estimated

Statistical Tests:

N=64			
	Gender:	log rank=0.01, df=1,	p=0.9230
	Age group:	log rank=13.62, df=3,	p=0.0035
	BMI group:	log rank=0.33, df=2,	p=0.8461
N=63			
	Gender:	log rank=0.01, df=1,	p=0.9425
	Age group:	log rank=15.59, df=3,	p=0.0014
	BMI group:	log rank=0.69, df=2,	p=0.7074

Table 5—Statistics for the Average of Four 20-Minute MWT

Group	n	Mean (SEM)	75%tile (SEM)	# scoring 20+
Male	27	18.92 (0.35)	17.75 (1.35)	18
Female	37	18.52 (0.50)	NA	26
Female, excl	36	18.84 (0.40)	NA	26
Age 30-39	18	17.52 (0.68)	14.33 (1.66)	8
Age 40-49	20	18.88 (0.70)	NA	16
Age 40-49, excl	19	19.50 (0.36)	NA	16
Age 50-59	15	19.18 (0.44)	NA	11
Age 60-69	11	19.59 (0.30)	NA	9
BMI <24.5	32	18.64 (0.43)	17.50 (2.07)	22
BMI 24.5-<29.5	26	18.54 (0.59)	18.88 (1.98)	18
BMI 24.5-<29.5, excl	25	19.00 (0.39)	NA	18
BMI 29.5+	6	19.58 (0.34)	19.75 (2.31)	4
Total 64	64	18.69 (0.33)	18.25 (1.03)	44
Total 63	63	18.87 (0.27)	18.88 (1.11)	44

Excl: statistics excluding the potential outlier
NA: 75th percentile or its standard error cannot be estimated

Statistical Tests:

N=64			
	Gender:	log rank=0.04, df=1,	p=0.8375
	Age group:	log rank=8.57, df=3,	p=0.0355
	BMI group:	log rank=0.01, df=2,	p=0.9946
N=63			
	Gender:	log rank=0.17, df=1,	p=0.6843
	Age group:	log rank=10.23, df=3,	p=0.0167
	BMI group:	log rank=0.11, df=2,	p=0.9462

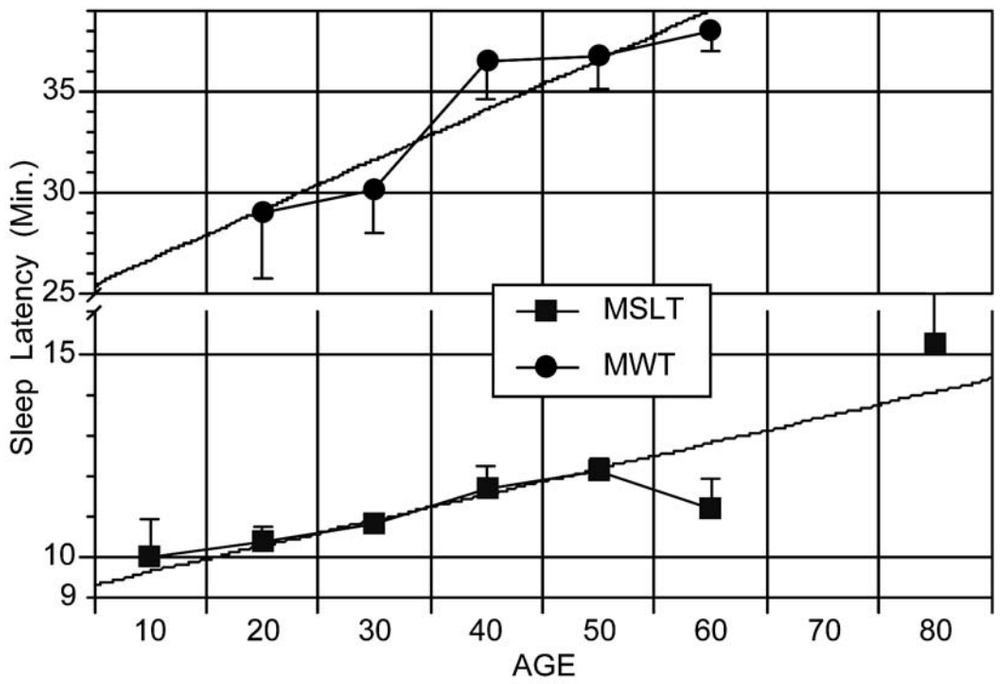

Figure 4—MSL with SEM on the MSLT and the 40 min MWT by age. Linear regression lines are plotted.

tance of age specific normative values. However, the use of a SD to determine normative ranges may still be problematic since a large percentage of normal Ss score 20 minutes on the 20-minute MWT protocol. The 40-minute MWT may be more normally distributed because fewer individuals received the 40-minute score. However, additional work needs to examine distribution issues in these tests (see Tables 4 & 5).

In narcolepsy patients, studies reporting SOREMP data showed a high sensitivity and very high specificity for two or more SOREMPs for the diagnosis of narcolepsy. The specificity of SOREMPs is also increased if other sleep disorders (especially sleep apnea) are excluded. The presence of two or more SOREMPs occurs in the majority of narcolepsy patients on a single MSLT with repeat testing identifying others.[28] Additionally, the MSL in narcolepsy patients is around 2 - 3 minutes compared to normal controls at 10.4 minutes. Therefore, MSLT latency data are strongly discriminatory in the diagnosis of narcolepsy but must be viewed in context with the sleep history and other medical information.

The use of the MSLT in idiopathic hypersomnia shows that the MSL is between that of normal controls and narcolepsy patients. However, the SD is again large and does not discriminate between normal controls and sleepy individuals.

In studies of OSA patients, the MSL from the MWT and MSLT overlapped extensively with those of normal Ss. The MSL from the 20 minute MWT provided the best separation between groups. This was followed closely by the 40 minute MWT. The MSL from the MSLT demonstrated significant improvement with CPAP treatment, but still fell within one SD of the pre-treatment mean and showed no statistical difference between CPAP- and placebo-treated OSA patients. Similar results were found with the 40 minute MWT. Unfortunately, there were no studies using the 20 minute MWT following CPAP treatment. Consequently, it seems that the MSL change between pre- and post-treatment for an individual is probably meaningful, although comparison of these data with the normative values is not helpful.

There are little data on the use of the MSLT or MWT in patients with medical or neurological conditions associated with sleepiness. This data may be difficult to obtain because of special concerns with these patient populations. It maybe difficult to obtain a homogeneous patient group, especially if the diagnosis can only be made on clinical grounds or the disease is progressive. In diseases that involve impairment or abnormality in brain function, there is concern about the patient's ability to understand or follow test instructions, as well as the ability to reliably score sleep onset when the EEG is abnormal.

Studies of safety indicate that sleep latency is sensitive to circadian variation when subjects remain awake in the laboratory under simulated night shift conditions. Both MSLT and MWT latencies decreased across the night to low points in the early morning hours. The decreases were greater for MSL on the MSLT. Thus, subjects remained awake longer when requested to remain awake, but still fell asleep under conditions of high circadian sleepiness. This indicates that the MSLT and MWT respond appropriately when sleep wake cycles are reversed and that circadian variation shortens the MSL.

Only a few studies examined driving performance and MSL measures. In general, sleep latencies were not correlated with roadway crash history for most patient groups. However, they were positively correlated with TST before long-distance driving in subjects without sleep disorders. In laboratory tests, patients with narcolepsy or sleep apnea had decreased scores on the

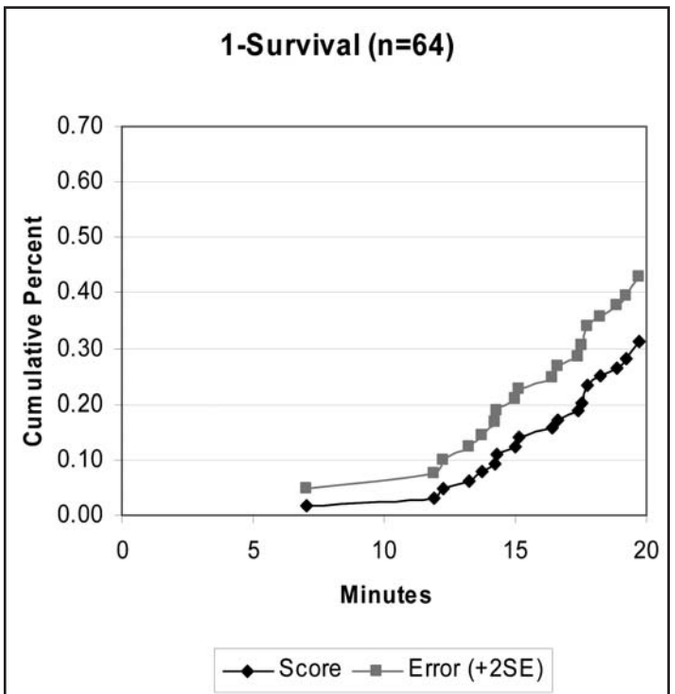

Figure 5—Survival plot for 20 minute MWT. The graph shows the result from one survival analysis. The X-axis is the average score of four naps. The Y-axis is the cumulative proportion. The lower line shows 1 minus the estimated survival curve for each score. The upper line marks the estimated error bound of two SEM.

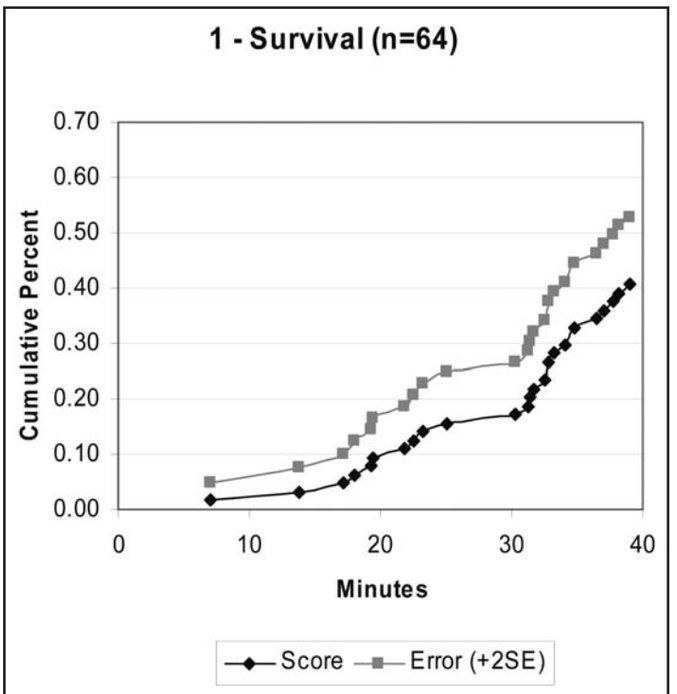

Figure 6—Survival plot for the 40 minute MWT defining sleep latency as three epochs of stage 1 or the first epoch of any stage. The graph shows the result from one survival analysis. The X-axis is the average score of four naps. The Y-axis is the cumulative proportion. The lower line shows 1 minus the estimated survival curve for each score. The upper line marks the estimated error bound of two SEM.

MSLT and poorer simulated driving performance. Apnea patients treated with CPAP had improved scores on both the MSLT and MWT, performed better at simulated driving, and also reported better driving performance on the road. Both sleep deprivation and alcohol use decreased sleep latencies and resulted in poorer driving performance. Although it appears that sleepiness and alcohol have a negative effect on driving performance, it is not clear how MSL relates to driving performance.

The data on the use of the MSLT or MWT to evaluate safety is very limited. It is not clear how these results translate to risk when performing overnight work or other real-world activities. Although some of the studies employed experienced shift workers, none of the studies linked sleep latency measurements to actual errors or incidents in the field. Thus, based on current studies, the predictive validity of MSL for safety is not well-established. Consequently, determination of an individual's excessive sleepiness and fitness for duty in safety-sensitive positions should not rely solely on the MSLT or MWT.

In general, this review indicates that the MSLT and MWT have different strengths and weaknesses. The MSLT appears valuable for the diagnosis of narcolepsy based on SOREMPs. Additionally, the MSL on the MSLT seems to reflect an individual's response to treatment or ingestion of medication. This provides face validity for the MSLT as a measure of sleepiness. In comparison, the MWT values also reflect an individual's response to treatment and medication. The instruction to stay awake in the MWT results in longer MSL values compared to the MSLT so consequently the MWT may have more face validity as a measure to evaluate an individual's ability to function during wakefulness. Both the MSLT and MWT have similar weaknesses. In both tests the MSL values overlap extensively in normal and patient populations limiting the ability to distinguish between the groups. In addition, the MSLT appears limited by a floor effect, while the MWT is limited by a ceiling effect. However, the ceiling effect is slightly diminished in the 40 minute MWT when latencies are scored to the first epoch of stage 1 since MSL decreases from the extreme upper values. Finally, there is probably non-linearity in both tests, meaning that a latency change from 2 to 4 minutes may not be equivalent to a change from 12 to 14 minutes.

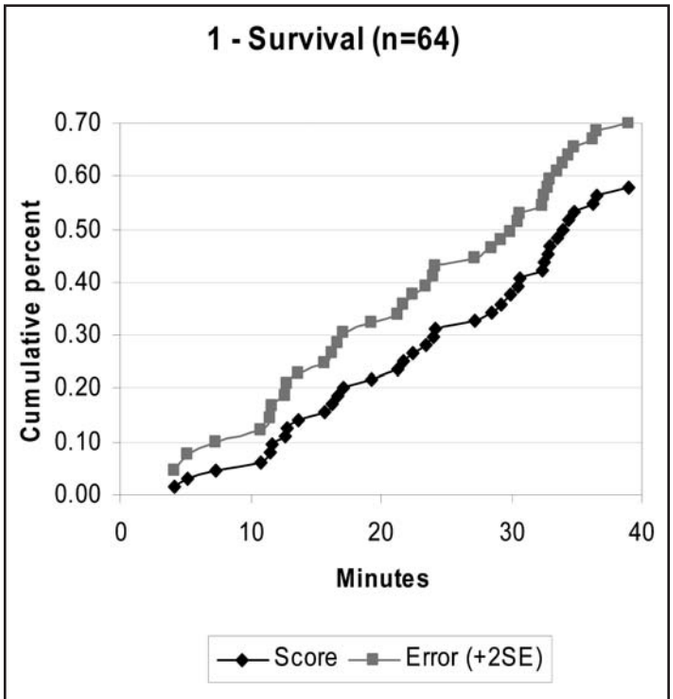

Figure 7—Survival plot for the 40 minute MWT defining sleep latency as the first ten seconds of stage 1 or the first epoch of any stage. The graph shows the result from one survival analysis. The X-axis is the average score of four naps. The Y-axis is the cumulative proportion. The lower line shows 1 minus the estimated survival curve for each score. The upper line marks the estimated error bound of two SEM.

SUMMARY AND CONCLUSION

The studies examined in this review indicate that the MSL is sensitive to conditions expected to increase sleepiness. MSL are generally lower following sleep loss, following use of sedating medications, during wakefulness in the late night or early morning hours, and among patients with sleep disorders associated with excessive sleepiness such as narcolepsy or obstructive sleep apnea. However, the wide range in MSL makes it difficult to establish a specific threshold value for excessive sleepiness or to discriminate patients with sleep disorders from non-patients. Some of this variation may be attributable to methodological differences and some may be attributable to individual differences in sleep tendency (e.g., related to age).

The studies analyzed in this review indicate that the MSL on both the MSLT and MWT does not discriminate well between patients with sleep disorders and normal populations. This is due to large SD as well as floor or ceiling effects in the tests. However, the MSL shows appropriate change from initial testing to subsequent testing following treatment or manipulations intended to alter sleepiness or alertness. Additionally the presence of two or more SOREMPs on the MSLT is a common finding in narcolepsy patients. However, SOREMPs are not exclusive to narcolepsy patients but are frequent in untreated sleep apnea

Table 6—Percentiles for mean sleep latencies on the 40-min MWT using a sleep latency definition of 10 sec of stage 1or the first epoch of any stage.

Percentile	MSL	Percentile	MSL
pct0-1.5	4.2	pct30-31	24.1
pct2-3	5.3	pct31.5-32.5	27.1
pct3.5-4.5	7.3	pct33-34	28.5
pct5-6	10.8	pct34.5-35.5	29.3
pct6.5-7.5	11.5	pct36-37	29.9
pct8-9	11.6	pct37.5	30.2
pct9.5-10.5	12.6	pct38-39	30.5
pct11-12	12.8	pct39.5-40.5	30.6
pct12.5	13.2	pct41-42	32.4
pct13-14	13.6	pct42.5-43.5	32.5
pct14.5-15.5	15.8	pct44-45	32.8
pct16-17	16.3	pct45.5-46.5	32.9
pct17.5-18.5	16.8	pct47-48	33.5
pct19-20	17.1	pct48.5-49.5	34.0
pct20.5-21.5	19.3	pct50	34.2
pct22-23	21.3	pct50.5-51.5	34.4
pct23.5-24.5	21.8	pct52-53	34.8
pct25	22.1	pct53.5-54.5	36.3
pct25.5-26.5	22.4	pct55-56	36.5
pct27-28	23.5	pct56.5-57.5	39.0
pct28.5-29.5	24.0	pct58-100	40.0

patients. This underscores the necessity of ruling out or treating other sleep disorders before interpreting SOREMPs for diagnostic purposes. Finally, the MSL is sensitive to circadian changes but a relationship between MSL and evaluation of safety in real life operations has not been established.

Analysis of MSLT normative data, showed that there are many methodological factors that affect MSL. The four factors that significantly affected the MSL–were: number of naps, age, sleep latency definition and prior TST. Type of protocol (research or clinical) did not result in significant differences in MSL. The significant effect for number of naps demonstrated longer latencies with a five nap protocol due to a more prominent "last nap effect." Age effects included longer latencies for 50 and 80 year old age groups compared to all younger age groups. (A significant age effect was also found for the MWT). The significant difference due to sleep latency definition showed that latency to the first epoch resulted in a significantly shorter MSL than the latency to sustained sleep; however the 0.7 minute difference is probably not clinically significant. Finally, prior TST was inversely related to MSL showing that shorter TST (mean of 6.6 hours) resulted in longer MSL than longer TST (mean of 7.5 hours). This suggests that shorter normal sleepers are not sleep deprived. However, when normal sleepers are sleep deprived, the MSL is decreased. Given all the factors that affect MSL it is not possible to determine one number to represent a normal control mean and SD for the MSLT. At a minimum, MSL comparisons should be made to data from a similar age group and for the number of nap opportunities allowed.

In general, a diagnosis of excessive sleepiness should be made with extreme care and with as much clinical information as possible. The diagnosis can affect an individual's job or ability to drive and, in some cases, public safety. Based on current evidence, the MSL should not be the sole criterion for determining sleepiness or for certifying a diagnosis or response to treatment. Interpretation of test results should be made within the context of the individual patient history and as part of other medical information and testing.

Future research would benefit from standardized protocols for the performance of the MSLT and MWT sufficiently detailed to limit variability and provide control of extraneous factors. Subsequent research should be directed at addressing the specific operating characteristics of the MSLT or MWT across different patient groups and ages as well as evaluating other objective and subjective tests for excessive sleepiness. Evaluation of all tests on varying sleep wake cycles should also be undertaken. Since the MWT may be particularly useful in situations where the ability to remain awake is being assessed, studies examining the relationship between MWT data, safety, and occupational risk in "real life" circumstances are needed.

REFERENCES

1. Bixler EO, Kales A, Soldatos CR, Kales JD, Healey S. Prevalence of sleep disorders in the Los Angeles metropolitan area. Am J Psychiatry. 1979;136(10):1257-62.
2. Lavie P. Sleep habits and sleep disturbances in industrial workers in Israel: main findings and some characteristics of workers complaining of excessive daytime sleepiness. Sleep. 1981;4(2):147-58.
3. Coleman RM, Roffwarg HP, Kennedy SJ, et al. Sleep-wake disorders based on a polysomnographic diagnosis. A national cooperative study. JAMA. 1982;247(7):997-1003.
4. Dement WC, Carskadon MA, Richardson GS. Excessive daytime sleepiness in the sleep apnea syndrome. In: Guilleminault C, Dement WC, eds. Sleep Apnea Syndromes. New York: Alan R. Liss; 1978:23-46.
5. Dement WC. Daytime sleepiness and sleep "attacks." In: Guilleminault C, Dement WC, Passouant P, eds. Narcolepsy. Orofino, ID: Spectrum; 1976:17-42.
6. Guilleminault C, Dement WC. 235 cases of excessive daytime sleepiness. Diagnosis and tentative classification. J Neurol Sci. 1977;31(1):13-27.
7. Guilleminault C, Billiard M, Montplaisir J, Dement WC. Altered states of consciousness in disorders of daytime sleepiness. J Neurol Sci. 1975;26(3):377-93.
8. Mitler MM, Carskadon MA, Czeisler CA, Dement WC, Dinges DF, Graeber RC. Catastrophes, sleep, and public policy: consensus report. Sleep. 1988;11(1):100-9.
9. National Commission on Sleep Disorders Research. (Department of Health and Human Services). Wake up America: national sleep alert. 1992.
10. Carskadon MA, Dement WC. The multiple sleep latency test: what does it measure? Sleep. 1982;5 Suppl 2:S67-72.
11. Mitler MM, Gujavarty KS, Browman CP. Maintenance of wakefulness test: a polysomnographic technique for evaluation treatment efficacy in patients with excessive somnolence. Electroencephalogr Clin Neurophysiol. 1982;53(6):658-61.
12. Thorpy MJ. The clinical use of the Multiple Sleep Latency Test. The Standards of Practice Committee of the American Sleep Disorders Association. Sleep. 1992;15(3):268-76.
13. Davis H, Davis PA, Loomis AL, Harvey EN, Hobart G. Changes in human brain potentials during the onset of sleep. Science. 1937;86:448-450.
14. Blake H, Gerard RW. Brain potentials during sleep. American Journal of Physiology. 1937;119:692-703.
15. Loomis AL, Harvey EN, Hobart GA. Cerebral states during sleep, as studied by human brain potentials. Journal of Experimental Psychology. 1937;21:127-144.
16. Rechtschaffen A, Kales A, (Eds). A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Los Angeles: BIS/BRI, UCLA; 1968.
17. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: a new approach. Psychophysiology. 1973;10(4):431-6.
18. Dement WC, Kelley J, Laughlin E, et al. Life on the basic rest-activity cycle (BRAC): sleep studies of a ninety minute day. Psychophysiology. 1972;9:132.
19. Carskadon MA, Dement WC. Sleep studies on a 90-minute day. Electroencephalogr Clin Neurophysiol. 1975;39(2):145-55.
20. Carskadon MA, Dement WC. Sleepiness and sleep state on a 90-min schedule. Psychophysiology. 1977;14(2):127-33.
21. Carskadon MA, Dement WC. Sleep tendency: an objective measure of sleep loss. Sleep Research. 1977;6:200.
22. Carskadon MA, Dement WC. Effects of total sleep loss on sleep tendency. Percept Mot Skills. 1979;48(2):495-506.
23. Richardson GS, Carskadon MA, Flagg W, Van den Hoed J, Dement WC, Mitler MM. Excessive daytime sleepiness in man: multiple sleep latency measurement in narcoleptic and control subjects. Electroencephalogr Clin Neurophysiol. 1978;45(5):621-7.
24. Mitler MM, Van den Hoed J, Carskadon MA, et al. REM sleep episodes during the Multiple Sleep Latency Test in narcoleptic patients. Electroencephalogr Clin Neurophysiol. 1979;46(4):479-81.
25. Rechtschaffen A, Wolpert EA, Dement WC, Mitchell SA, Fisher C. Nocturnal Sleep of Narcoleptics. Electroencephalogr Clin Neurophysiol. 1963;15:599-609.
26. Takahashi Y, Jimbo M. Polygraphic study of narcoleptic syndrome, with special reference to hypnagogic hallucination and cataplexy. Folia Psychiatr. Neurol. Jap. Suppl.7. 1964:343-7.

27. Hartse KM, Zorick F, Sicklesteel J, Piccione P, Roth T. Nap recordings in the diagnosis of daytime somnolence. Sleep Research. 1979;8:190.
28. van den Hoed J, Kraemer H, Guilleminault C, et al. Disorders of excessive daytime somnolence: polygraphic and clinical data for 100 patients. Sleep. 1981;4(1):23-37.
29. Hartse KM, Roth T, Zorick FJ, Zammit G. The effect of instruction upon sleep latency during multiple daytime naps of normal subjects. Sleep Research. 1980;9:123.
30. Dement W, Seidel W, Carskadon MA. Daytime alertness, insomnia and benzodiazepines. Sleep. 1982;5(Suppl.1):S28-45.
31. Carskadon MA, Dement WC, Mitler MM, Roth T, Westbrook PR, Keenan S. Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. Sleep. 1986;9(4):519-24.
32. Lavie P, Scherson A. Ultrashort sleep-walking schedule. I. Evidence of ultradian rhythmicity in "sleepability'. Electroencephalogr Clin Neurophysiol. 1981;52(2):163-74.
33. Harrison Y, Horne JA. "High sleepability without sleepiness". The ability to fall asleep rapidly without other signs of sleepiness. Neurophysiol Clin. 1996;26(1):15-20.
34. Bonnet MH, Arand DL. 24-Hour metabolic rate in insomniacs and matched normal sleepers. Sleep. 1995;18(7):581-8.
35. Haynes SN, Fitzgerald SG, Shute GE, Hall M. The utility and validity of daytime naps in the assessment of sleep-onset insomnia. J Behav Med. 1985;8(3):237-47.
36. Stepanski E, Zorick F, Roehrs T, Young D, Roth T. Daytime alertness in patients with chronic insomnia compared with asymptomatic control subjects. Sleep. 1988;11(1):54-60.
37. Bonnet MH, Arand DL. Activity, arousal, and the MSLT in patients with insomnia. Sleep. 2000;23(2):205-12.
38. Roth T, Hartse KM, Zorick F, Conway W. Multiple naps and the evaluation of daytime sleepiness in patients with upper airway sleep apnea. Sleep. 1980;3(3-4):425-39.
39. Mitler MM, Gujavarty KS, Sampson MG, Browman CP. Multiple daytime nap approaches to evaluating the sleepy patient. Sleep. 1982;5 Suppl 2:S119-27.
40. Sangal RB, Thomas L, Mitler MM. Maintenance of wakefulness test and multiple sleep latency test. Measurement of different abilities in patients with sleep disorders [see comments]. Chest. 1992;101(4):898-902.
41. Palm L, Persson E, Elmqvist D, Blennow G. Sleep and wakefulness in normal preadolescent children. Sleep. 1989;12(4):299-308.
42. Valencia-Flores M, Campos RM, Mendez J, et al. Multiple sleep latency test (MSLT) and sleep apnea in aged women. Sleep. 1993;16(2):114-7.
43. Levine B, Roehrs T, Zorick F, Roth T. Daytime sleepiness in young adults. Sleep. 1988;11(1):39-46.
44. Richardson GS, Carskadon MA, Orav EJ, Dement WC. Circadian variation of sleep tendency in elderly and young adult subjects. Sleep. 1982;5 Suppl 2:S82-94.
45. Rosenthal L, Roehrs TA, Rosen A, Roth T. Level of sleepiness and total sleep time following various time in bed conditions. Sleep. 1993;16(3):226-32.
46. Carskadon MA, Dement WC. Nocturnal determinants of daytime sleepiness. Sleep. 1982;5 Suppl 2:S73-81.
47. Roehrs T, Papineau K, Rosenthal L, Roth T. Sleepiness and the reinforcing and subjective effects of methylphenidate. Exp Clin Psychopharmacol. 1999;7(2):145-50.
48. Roth T, Roehrs T, Koshorek G, Sicklesteel J, Zorick F. Central effects of antihistamine. Sleep Research. 1986;15:43.
49. Carskadon MA, Seidel WF, Greenblatt DJ, Dement WC. Daytime carryover of triazolam and flurazepam in elderly insomniacs. Sleep. 1982;5(4):361-71.
50. Roehrs T, Lumley M, Asker D, Zorick F, Roth T. Ethanol and caffeine effects on daytime sleepiness. Sleep Research. 1986;15:41.
51. Roehrs T, Kribbs N, Zorick F, Roth T. Hypnotic residual effects of benzodiazepines with repeated administration. Sleep. 1986;9(2):309-16.
52. Bishop C, Roehrs T, Rosenthal L, Roth T. Alerting effects of methylphenidate under basal and sleep-deprived conditions. Exp Clin Psychopharmacol. 1997;5(4):344-52.
53. Mitler MM, Hajdukovic R. Relative efficacy of drugs for the treatment of sleepiness in narcolepsy. Sleep. 1991;14(3):218-20.
54. Manni R, Ratti MT, Perucca E, Galimberti CA, Tartara A. A multiparametric investigation of daytime sleepiness and psychomotor functions in epileptic patients treated with phenobarbital and sodium valproate: a comparative controlled study. Electroencephalogr Clin Neurophysiol. 1993;86(5):322-8.
55. Stepanski E, Lamphere J, Badia P, Zorick F, Roth T. Sleep fragmentation and daytime sleepiness. Sleep. 1984;7(1):18-26.
56. Carskadon MA, Brown ED, Dement WC. Sleep fragmentation in the elderly: relationship to daytime sleep tendency. Neurobiol Aging. 1982;3(4):321-7.
57. Levine B, Roehrs T, Stepanski E, Zorick F, Roth T. Fragmenting sleep diminishes its recuperative value. Sleep. 1987;10(6):590-9.
58. Philip P, Stoohs R, Guilleminault C. Sleep fragmentation in normals: a model for sleepiness associated with upper airway resistance syndrome. Sleep. 1994;17(3):242-7.
59. Pressman MR, Fry JM. Relationship of autonomic nervous system activity to daytime sleepiness and prior sleep. Sleep. 1989;12(3):239-45.
60. Roehrs T, Merlotti L, Petrucelli N, Stepanski E, Roth T. Experimental sleep fragmentation. Sleep. 1994;17(5):438-43.
61. Mitler MM. Evaluation of treatment with stimulants in narcolepsy. Sleep. 1994;17(8 Suppl):S103-6.
62. Carskadon MA, Harvey K, Duke P, Anders TF, Litt IF, Dement WC. Pubertal changes in daytime sleepiness. Sleep. 1980;2(4):453-60.
63. Bonnet MH, Arand DL. Caffeine use as a model of acute and chronic insomnia. Sleep. 1992;15(6):526-36.
64. Bonnet MH, Arand DL. The consequences of a week of insomnia. II: Patients with insomnia. Sleep. 1998;21(4):359-68.
65. Bonnet MH, Arand DL. The use of lorazepam TID for chronic insomnia. Int Clin Psychopharmacol. 1999;14(2):81-9.
66. Reynolds CF, 3rd. Sleep in affective disorders. In: Kryger MH, Roth T, Dement W, eds. Principles and practice of sleep medicine. 1st ed. Philadelphia: W.B. Saunders; 1989:413-5.
67. Benca RM. Mood disorders. In: Kryger MH, Roth T, Dement W, eds. Principles and practice of sleep medicine. 2nd ed. Philadelphia: W.B. Saunders; 1994:899-913.
68. Broughton R, Aguirre M, Dunham W. A comparison of multiple and single sleep latency and cerebral evoked potential (P300) measures in the assessment of excessive daytime sleepiness in narcolepsy-cataplexy. Sleep. 1988;11(6):537-45.
69. Bonnet MH, Arand DL. Sleepiness as measured by modified multiple sleep latency testing varies as a function of preceding activity. Sleep. 1998;21(5):477-83.
70. Bliwise D, Seidel W, Karacan I, et al. Daytime sleepiness as a criterion in hypnotic medication trials: comparison of triazolam and flurazepam. Sleep. 1983;6(2):156-63.
71. Kingshott RN, Vennelle M, Coleman EL, Engleman HM, Mackay TW, Douglas NJ. Randomized, double-blind, placebo-controlled crossover trial of modafinil in the treatment of residual excessive daytime sleepiness in the sleep apnea/hypopnea syndrome. Am J Respir Crit Care Med. 2001;163(4):918-23.
72. Johnson LC, Spinweber CL, Gomez SA. Benzodiazepines and caffeine: effect on daytime sleepiness, performance, and mood. Psychopharmacology (Berl). 1990;101(2):160-7.
73. Cassel W, Ploch T, Becker C, Dugnus D, Peter JH, von Wichert P. Risk of traffic accidents in patients with sleep-disordered breathing: reduction with nasal CPAP. Eur Respir J. 1996;9(12):2606-11.
74. Fietze I, Quispe-Bravo S, Schiller W, et al. Respiratory arousals in mild obstructive sleep apnea syndrome. Sleep. 1999;22(5):583-9.

75. George CF, Boudreau AC, Smiley A. Effects of nasal CPAP on simulated driving performance in patients with obstructive sleep apnoea. Thorax. 1997;52(7):648-53.
76. Hakkanen H, Summala H, Partinen M, Tiihonen M, Silvo J. Blink duration as an indicator of driver sleepiness in professional bus drivers. Sleep. 1999;22(6):798-802.
77. Fitzpatrick MF, Alloway CE, Wakeford TM, MacLean AW, Munt PW, Day AG. Can patients with obstructive sleep apnea titrate their own continuous positive airway pressure? Am J Respir Crit Care Med. 2003;167(5):716-22.
78. Zwyghuizen-Doorenbos A, Roehrs T, Schaefer M, Roth T. Test-retest reliability of the MSLT. Sleep. 1988;11(6):562-5.
79. Seidel W, Dement W. The Multiple Sleep Latency Test: test-retest reliability. Sleep Research. 1981;10:105.
80. Folkerts M, Rosenthal L, Roehrs T, et al. The reliability of the diagnostic features in patients with narcolepsy. Biol Psychiatry. 1996;40(3):208-14.
81. Drake CL, Rice MF, Roehrs TA, Rosenthal L, Guido P, Roth T. Scoring reliability of the multiple sleep latency test in a clinical population. Sleep. 2000;23(7):911-3.
82. Benbadis SR, Qu Y, Perry MC, Dinner DS, Warnes H. Interrater reliability of the multiple sleep latency test. Electroencephalogr Clin Neurophysiol. 1995;95(4):302-4.
83. Johns MW. Sensitivity and specificity of the multiple sleep latency test (MSLT), the maintenance of wakefulness test and the Epworth sleepiness scale: failure of the MSLT as a gold standard. J Sleep Res. 2000;9(1):5-11.
84. Bliwise DL, Carskadon MA, Seidel WF, Nekich JC, Dement WC. MSLT-defined sleepiness and neuropsychological test performance do not correlate in the elderly. Neurobiol Aging. 1991;12(5):463-8.
85. Steinberg R, Schonberg C, Weess HG, Schneider C, Pritzel M. The validity of the multiple sleep latency test. Journal of Sleep Research. 1996;5, Supplement 1:220.
86. Randomized trial of modafinil for the treatment of pathological somnolence in narcolepsy. US Modafinil in Narcolepsy Multicenter Study Group. Ann Neurol. 1998;43(1):88-97.
87. Johns M, Hocking B. Daytime sleepiness and sleep habits of Australian workers. Sleep. 1997;20(10):844-9.
88. Doghramji K, Mitler MM, Sangal RB, et al. A normative study of the maintenance of wakefulness test (MWT). Electroencephalogr Clin Neurophysiol. 1997;103(5):554-62.
89. Moscovitch A, Partinen M, Guilleminault C. The positive diagnosis of narcolepsy and narcolepsy's borderland. Neurology. 1993;43(1):55-60.
90. Sackett DL. Rules of evidence and clinical recommendations for the management of patients. Can J Cardiol. 1993;9(6):487-9.
91. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. 1991;14(6):540-5.
92. Cook Y, Schmitt F, Berry D, Gilmore R, Phillips B, Lamb D. The effects of nocturnal sleep, sleep disordered breathing and periodic movements of sleep on the objective and subjective assessment of daytime somnolence in healthy aged adults. Sleep Research. 1988;17:95.
93. Vignatelli L, Plazzi G, Barbato A, et al. Italian version of the Epworth sleepiness scale: external validity. Neurol Sci. 2003;23(6):295-300.
94. Sleep Disorders Classification Committee of the Association of Sleep Disorders Centers. Diagnostic classification of sleep and arousal disorders. Sleep. 1979;2(1):1-137.
95. American Academy of Sleep Medicine. The International Classification of Sleep Disorders, Revised Rochester, MN: American Academy of Sleep Medicine; 1997.
96. Alloway CE, Ogilvie RD, Shapiro CM. The alpha attenuation test: assessing excessive daytime sleepiness in narcolepsy-cataplexy. Sleep. 1997;20(4):258-66.
97. Newman J, Broughton R. Pupillometric assessment of excessive daytime sleepiness in narcolepsy-cataplexy. Sleep. 1991;14(2):121-9.
98. Genton P, Benlakhel K, Disdier P, et al. [Diagnosis of narcolepsy-cataplexy: importance of continuous recording in ambulatory EEG. Report of 20 cases]. Neurophysiol Clin. 1995;25(4):187-95.
99. Guilleminault C, Pelayo R. Narcolepsy in prepubertal children. Ann Neurol. 1998;43(1):135-42.
100. Nykamp K, Rosenthal L, Helmus T, et al. Repeated nocturnal sleep latencies in narcoleptic, sleepy and alert subjects. Clin Neurophysiol. 1999;110(9):1531-4.
101. Rye DB, Dihenia B, Weissman JD, Epstein CM, Bliwise DL. Presentation of narcolepsy after 40. Neurology. 1998;50(2):459-65.
102. Walsh JK, Smitson SA, Kramer M. Sleep-onset REM sleep: comparison of narcoleptic and obstructive sleep apnea patients. Clin Electroencephalogr. 1982;13(1):57-60.
103. Zachariev Z, Djurkova A. Clinico-polysomnographic diagnostics of narcolepsy-cataplexy. Folia Med (Plovdiv). 1999;41(2):5-12.
104. Zorick F, Roehrs T, Koshorek G, et al. Patterns of sleepiness in various disorders of excessive daytime somnolence. Sleep. 1982;5 Suppl 2:S165-74.
105. Aldrich MS, Chervin RD, Malow BA. Value of the multiple sleep latency test (MSLT) for the diagnosis of narcolepsy. Sleep. 1997;20(8):620-9.
106. Amira SA, Johnson TS, Logowitz NB. Diagnosis of narcolepsy using the multiple sleep latency test: analysis of current laboratory criteria. Sleep. 1985;8(4):325-31.
107. Bassetti C, Aldrich MS. Idiopathic hypersomnia. A series of 42 patients. Brain. 1997;120 (Pt 8):1423-35.
108. Billiard M, Merle C, Carlander B, Ondze B, Alvarez D, Besset A. Idiopathic hypersomnia. Psychiatry Clin Neurosci. 1998;52(2):125-9.
109. Dolenc L, Besset A, Billiard M. Hypersomnia in association with dysthymia in comparison with idiopathic hypersomnia and normal controls. Pflugers Arch. 1996;431(6 Suppl 2):R303-4.
110. George CF, Boudreau AC, Smiley A. Comparison of simulated driving performance in narcolepsy and sleep apnea patients. Sleep. 1996;19(9):711-7.
111. Browman CP, Mitler MM. Hypersomnia and the perception of sleep-wake states: some preliminary findings. Percept Mot Skills. 1988;66(2):463-70.
112. Guilleminault C, Stoohs R, Duncan S. Snoring (I). Daytime sleepiness in regular heavy snorers. Chest. 1991;99(1):40-8.
113. Guilleminault C, Stoohs R, Clerk A, Simmons J, Labanowski M. From obstructive sleep apnea syndrome to upper airway resistance syndrome: consistency of daytime sleepiness. Sleep. 1992;15(6 Suppl):S13-6.
114. Kribbs NB, Pack AI, Kline LR, et al. Effects of one night without nasal CPAP treatment on sleep and sleepiness in patients with obstructive sleep apnea. Am Rev Respir Dis. 1993;147(5):1162-8.
115. Sangal RB, Sangal JM. Abnormal visual P300 latency in obstructive sleep apnea does not change acutely upon treatment with CPAP. Sleep. 1997;20(9):702-4.
116. Sforza E, Lugaresi E. Daytime sleepiness and nasal continuous positive airway pressure therapy in obstructive sleep apnea syndrome patients: effects of chronic treatment and 1-night therapy withdrawal. Sleep. 1995;18(3):195-201.
117. Morisson F, Decary A, Petit D, Lavigne G, Malo J, Montplaisir J. Daytime sleepiness and EEG spectral analysis in apneic patients before and after treatment with continuous positive airway pressure. Chest. 2001;119(1):45-52.
118. Zorick FJ, Roehrs T, Conway W, Potts G, Roth T. Response to CPAP and UPPP in apnea. Henry Ford Hosp Med J. 1990;38(4):223-6.
119. Engleman HM, Martin SE, Deary IJ, Douglas NJ. Effect of CPAP therapy on daytime function in patients with mild sleep apnoea/hypopnoea syndrome. Thorax. 1997;52(2):114-9.
120. Engleman HM, Martin SE, Deary IJ, Douglas NJ. Effect of contin-

uous positive airway pressure treatment on daytime function in sleep apnoea/hypopnoea syndrome. Lancet. 1994;343(8897):572-5.
121. Engleman HM, Martin SE, Kingshott RN, Mackay TW, Deary IJ, Douglas NJ. Randomised placebo controlled trial of daytime function after continuous positive airway pressure (CPAP) therapy for the sleep apnoea/hypopnoea syndrome. Thorax. 1998;53(5):341-5.
122. Barbe F, Mayoralas LR, Duran J, et al. Treatment with continuous positive airway pressure is not effective in patients with sleep apnea but no daytime sleepiness. a randomized, controlled trial. Ann Intern Med. 2001;134(11):1015-23.
123. Meurice JC, Marc I, Series F. Efficacy of auto-CPAP in the treatment of obstructive sleep apnea/hypopnea syndrome. Am J Respir Crit Care Med. 1996;153(2):794-8.
124. Engleman HM, Kingshott RN, Wraith PK, Mackay TW, Deary IJ, Douglas NJ. Randomized placebo-controlled crossover trial of continuous positive airway pressure for mild sleep Apnea/Hypopnea syndrome. Am J Respir Crit Care Med. 1999;159(2):461-7.
125. Lecendreux M, Konofal E, Bouvard M, Falissard B, Mouren-Simeoni MC. Sleep and alertness in children with ADHD. J Child Psychol Psychiatry. 2000;41(6):803-12.
126. Tarasiuk A, Abdul-Hai A, Moser A, et al. Sleep disruption and objective sleepiness in children with beta-thalassemia and congenital dyserythropoietic anemia. Arch Pediatr Adolesc Med. 2003;157(5):463-8.
127. Ionescu D, Driver HS, Heon E, Flanagan J, Shapiro CM. Sleep and daytime sleepiness in retinitis pigmentosa patients. J Sleep Res. 2001;10(4):329-35.
128. Hertz G, Cataletto M, Feinsilver SH, Angulo M. Sleep and breathing patterns in patients with Prader Willi syndrome (PWS): effects of age and gender. Sleep. 1993;16(4):366-71.
129. van Hilten JJ, Kerkhof GA, van Dijk JG, Dunnewold R, Wintzen AR. Disruption of sleep-wake rhythmicity and daytime sleepiness in myotonic dystrophy. J Neurol Sci. 1993;114(1):68-75.
130. Lankford DA, Wellman JJ, O'Hara C. Posttraumatic narcolepsy in mild to moderate closed head injury. Sleep. 1994;17(8 Suppl):S25-8.
131. Hirsch M, Carlander B, Verge M, et al. Objective and subjective sleep disturbances in patients with rheumatoid arthritis. A reappraisal. Arthritis Rheum. 1994;37(1):41-9.
132. Orr WC, Shamma-Othman Z, Levin D, Othman J, Rundell OH. Persistent hypoxemia and excessive daytime sleepiness in chronic obstructive pulmonary disease (COPD). Chest. 1990;97(3):583-5.
133. Hanly P, Zuberi-Khokhar N. Daytime sleepiness in patients with congestive heart failure and Cheyne-Stokes respiration. Chest. 1995;107(4):952-8.
134. Drake ME, Jr., Weate SJ, Newell SA, Padamadan H, Pakalnis A. Multiple sleep latency tests in epilepsy. Clin Electroencephalogr. 1994;25(2):59-62.
135. Arnulf I, Konofal E, Merino-Andreu M, et al. Parkinson's disease and sleepiness: an integral part of PD. Neurology. 2002;58(7):1019-24.
136. Hanly PJ, Gabor JY, Chan C, Pierratos A. Daytime sleepiness in patients with CRF: impact of nocturnal hemodialysis. Am J Kidney Dis. 2003;41(2):403-10.
137. Parker KP, Bliwise DL, Bailey JL, Rye DB. Daytime sleepiness in stable hemodialysis patients. Am J Kidney Dis. 2003;41(2):394-402.
138. Giubilei F, Antonini G, Bastianello S, et al. Excessive daytime sleepiness in myotonic dystrophy. J Neurol Sci. 1999;164(1):60-3.
139. Taphoorn MJ, van Someren E, Snoek FJ, et al. Fatigue, sleep disturbances and circadian rhythm in multiple sclerosis. J Neurol. 1993;240(7):446-8.
140. Drake CL, Roehrs TA, Royer H, Koshorek G, Turner RB, Roth T. Effects of an experimentally induced rhinovirus cold on sleep, performance, and daytime alertness. Physiol Behav. 2000;71(1-2):75-81.
141. Hurwitz TD, Mahowald MW, Kuskowski M, Engdahl BE. Polysomnographic sleep is not clinically impaired in Vietnam combat veterans with chronic posttraumatic stress disorder. Biol Psychiatry. 1998;44(10):1066-73.
142. Hogl B, Saletu M, Brandauer E, et al. Modafinil for the treatment of daytime sleepiness in Parkinson's disease: a double-blind, randomized, crossover, placebo-controlled polygraphic trial. Sleep. 2002;25(8):905-9.
143. Rosenthal L, Nykamp K, Day R, et al. The detection of brief daytime sleep episodes. Sleep. 1999;22(2):211-4.
144. Seidel WF, Cohen SA, Bliwise NG, Roth T, Dement WC. Dose-related effects of triazolam and flurazepam on a circadian rhythm insomnia. Clin Pharmacol Ther. 1986;40(3):314-20.
145. Casagrande M, Ferrara M, Curcio G, Porcu S. Assessing nighttime vigilance through a three-letter cancellation task (3-LCT): effects of daytime sleep with temazepam or placebo. Physiol Behav. 1999;68(1-2):251-6.
146. Porcu S, Bellatreccia A, Ferrara M, Casagrande M. Acutely shifting the sleep-wake cycle: nighttime sleepiness after diurnal administration of temazepam or placebo. Aviat Space Environ Med. 1997;68(8):688-94.
147. Porcu S, Bellatreccia A, Ferrara M, Casagrande M. Performance, ability to stay awake, and tendency to fall asleep during the night after a diurnal sleep with temazepam or placebo. Sleep. 1997;20(7):535-41.
148. Porcu S, Ferrara M, Urbani L, Bellatreccia A, Casagrande M. Smooth pursuit and saccadic eye movements as possible indicators of nighttime sleepiness. Physiol Behav. 1998;65(3):437-43.
149. Thessing VC, Anch AM, Muehlbach MJ, Schweitzer PK, Walsh JK. Two- and 4-hour bright-light exposures differentially effect sleepiness and performance the subsequent night. Sleep. 1994;17(2):140-5.
150. Walsh JK, Schweitzer PK, Anch AM, Muehlbach MJ, Jenkins NA, Dickins QS. Sleepiness/alertness on a simulated night shift following sleep at home with triazolam. Sleep. 1991;14(2):140-6.
151. Roehrs T, Salin-Pascual R, Merlotti L, Rosenthal L, Roth T. Phase advance in moderately sleepy and alert normals. Sleep. 1996;19(5):417-22.
152. Beersma DG, Dijk DJ, Blok CG, Everhardus I. REM sleep deprivation during 5 hours leads to an immediate REM sleep rebound and to suppression of non-REM sleep intensity. Electroencephalogr Clin Neurophysiol. 1990;76(2):114-22.
153. Young T, Blustein J, Finn L, Palta M. Sleep-disordered breathing and motor vehicle accidents in a population-based sample of employed adults. Sleep. 1997;20(8):608-13.
154. Aldrich MS. Automobile accidents in patients with sleep disorders. Sleep. 1989;12(6):487-94.
155. Roehrs T, Beare D, Zorick F, Roth T. Sleepiness and ethanol effects on simulated driving. Alcohol Clin Exp Res. 1994;18(1):154-8.
156. Philip P, Ghorayeb I, Leger D, et al. Objective measurement of sleepiness in summer vacation long-distance drivers. Electroencephalogr Clin Neurophysiol. 1997;102(5):383-9.
157. Damian MS, Gerlach A, Schmidt F, Lehmann E, Reichmann H. Modafinil for excessive daytime sleepiness in myotonic dystrophy. Neurology. 2001;56(6):794-6.
158. Randomized trial of modafinil as a treatment for the excessive daytime somnolence of narcolepsy: US Modafinil in Narcolepsy Multicenter Study Group. Neurology. 2000;54(5):1166-75.
159. Rosenthal L, Roehrs T, Zwyghuizen-Doorenbos A, Plath D, Roth T. Alerting effects of caffeine after normal and restricted sleep. Neuropsychopharmacology. 1991;4(2):103-8.
160. Muehlbach MJ, Walsh JK. The effects of caffeine on simulated night-shift work and subsequent daytime sleep. Sleep. 1995;18(1):22-9.
161. Lumley M, Roehrs T, Asker D, Zorick F, Roth T. Ethanol and caffeine effects on daytime sleepiness/alertness. Sleep. 1987;10(4):306-12.
162. Dement WC. Objective measurements of daytime sleepiness and performance comparing quazepam with flurazepam in two adult

populations using the Multiple Sleep Latency Test. J Clin Psychiatry. 1991;52 Suppl:31-7.
163. Broughton RJ, Fleming JA, George CF, et al. Randomized, double-blind, placebo-controlled crossover trial of modafinil in the treatment of excessive daytime sleepiness in narcolepsy. Neurology. 1997;49(2):444-51.
164. Lammers GJ, Arends J, Declerck AC, Kamphuisen HA, Schouwink G, Troost J. Ritanserin, a 5-HT2 receptor blocker, as add-on treatment in narcolepsy. Sleep. 1991;14(2):130-2.
165. Bishop C, Rosenthal L, Helmus T, Roehrs T, Roth T. The frequency of multiple sleep onset REM periods among subjects with no excessive daytime sleepiness. Sleep. 1996;19(9):727-30.
166. Carskadon MA, Wolfson AR, Acebo C, Tzischinsky O, Seifer R. Adolescent sleep patterns, circadian timing, and sleepiness at a transition to early school days. Sleep. 1998;21(8):871-81.
167. Edinger JD, Fins AI, Sullivan RJ, Jr., et al. Do our methods lead to insomniacs' madness?: Daytime testing after laboratory and home-based polysomnographic studies. Sleep. 1997;20(12):1127-34.
168. Hartse KM, Roth T, Zorick FJ. Daytime sleepiness and daytime wakefulness: the effect of instruction. Sleep. 1982;5 Suppl 2:S107-18.
169. Martin SE, Engleman HM, Deary IJ, Douglas NJ. The effect of sleep fragmentation on daytime function. Am J Respir Crit Care Med. 1996;153(4 Pt 1):1328-32.
170. Mattmann P, Loepfe M, Scheitlin T, et al. Day-time residual effects and motor activity after three benzodiazepine hypnotics. Arzneimittelforschung. 1982;32(4):461-5.
171. Bonnet MH, Arand DL. Situational insomnia: consistency, predictors, and outcomes. Sleep. 2003;26(8):1029-36.
172. Bonnet MH, Arand DL. The consequences of a week of insomnia. Sleep. 1996;19(6):453-61.
173. Manni R, Ratti MT, Barzaghi N, et al. Daytime sleepiness in healthy university students: a multiparametric study. Ital J Neurol Sci. 1991;12(3):303-9.
174. Mitler MM, Nelson S, Hajdukovic R. Narcolepsy. Diagnosis, treatment, and management. Psychiatr Clin North Am. 1987;10(4):593-606.
175. Reynolds CF, 3rd, Jennings JR, Hoch CC, et al. Daytime sleepiness in the healthy "old old": a comparison with young adults. J Am Geriatr Soc. 1991;39(10):957-62.
176. Roth T, Roehrs T, Koshorek G, Sicklesteel J, Zorick F. Sedative effects of antihistamines. J Allergy Clin Immunol. 1987;80(1):94-8.
177. Seidel WF, Ball S, Cohen S, Patterson N, Yost D, Dement WC. Daytime alertness in relation to mood, performance, and nocturnal sleep in chronic insomniacs and noncomplaining sleepers. Sleep. 1984;7(3):230-8.
178. Seidel WF, Cohen S, Bliwise NG, Dement WC. Cetirizine effects on objective measures of daytime sleepiness and performance. Ann Allergy. 1987;59(6 Pt 2):58-62.
179. Bonnet MH, Arand DL. Arousal components which differentiate the MWT from the MSLT. Sleep. 2001;24(4):441-7.
180. Martin SE, Wraith PK, Deary IJ, Douglas NJ. The effect of nonvisible sleep fragmentation on daytime function. Am J Respir Crit Care Med. 1997;155(5):1596-601.
181. Mitler MM, Doghramji K, Shapiro C. The maintenance of wakefulness test: normative data by age. J Psychosom Res. 2000;49(5):363-5.