Association between the Epworth Sleepiness Scale and the Multiple Sleep Latency Test in a Clinical Population

Selim R. Benbadis, MD; Edward Mascha, MS;
Michael C. Perry, R.EEG.T, RPSGT;
Barbara R. Wolgamuth, R.EEG.T;
Laurence A. Smolley, MD; and Dudley S. Dinner, MD

Background: Excessive daytime sleepiness can be evaluated with both subjective and objective tests.

Objective: To examine the association between Epworth Sleepiness Scale scores and sleep latency on the multiple sleep latency test.

Design: Case series.

Setting: Referral sleep disorders center.

Patients: 102 consecutive patients evaluated for excessive daytime sleepiness.

Measurements: Epworth Sleepiness Scale scores and mean sleep latency on the multiple sleep latency test.

Results: No significant association was seen between Epworth scores and mean sleep latency (Pearson correlation, −0.17 [95% CI, −0.35 to 0.03]; $P = 0.09$) (analysis of variance, $P = 0.13$). The mean Epworth score did not differ in three groups of patients who were defined by mean sleep latency as having normal sleep latency (>10 minutes), moderate sleep latency (5 to 10 minutes), or severe sleep latency (<5 minutes) (analysis of variance, $P = 0.13$).

Conclusions: No statistically or clinically significant association was seen between Epworth scores and mean sleep latency. The subjective Epworth Sleepiness Scale and the objective multiple sleep latency test may evaluate different, complementary aspects of sleepiness.

This paper is also available at http://www.acponline.org.

Ann Intern Med. 1999;130:289-292.

From the Cleveland Clinic Florida, Fort Lauderdale, Florida; and the Cleveland Clinic Foundation, Cleveland, Ohio. For current author addresses, see end of text.

Excessive daytime sleepiness can be evaluated with both subjective and objective methods. Of the objective methods, the multiple sleep latency test is generally considered the gold standard (1–3). This test is based on the principle that degrees of sleepiness can be measured by how quickly one falls asleep (sleep latency) when given the opportunity to do so. However, the multiple sleep latency test is cumbersome and costly, and it would be useful to know whether sleep latency can be reliably predicted by a simple subjective scale, such as the Epworth Sleepiness Scale. This test, first described in 1991 (4), consists of eight questions and yields a score of 0 to 24. It is easy to administer and is currently the most widely used subjective test for sleepiness (5–7); it was recently used in drug trials for narcolepsy (8–10).

The agreement between objective and subjective methods of sleepiness assessment has been found to be only moderate (4, 7, 9). Our purpose was to examine the association between Epworth Sleepiness Scale scores and sleep latency in a clinical population.

Methods

We analyzed 102 consecutive multiple sleep latency tests performed at our referral sleep center in patients seen in the sleep clinic. All tests included at least four naps scored according to the standard criteria for sleep staging (11). All were done according to a standard protocol (1–3), which included interpretation based on the average (mean) sleep latency. Three epochs (30-second segments) of stage 1 or one epoch of any other sleep stage was required to establish sleep onset.

Before completing the multiple sleep latency test, all patients filled out the Epworth Sleepiness Scale, which is an eight-item questionnaire that asks the patient to answer each question with a number from 0 (not at all likely to fall asleep) to 3 (very likely to fall asleep) (**Appendix Figure**). This yields a total score of 0 (minimum) to 24 (maximum). Both the multiple sleep latency test and the Epworth Sleepiness Scale were part of a standard clinical protocol.

The Pearson correlation coefficient and 95% CIs were used to assess the correlation between Epworth scores and mean sleep latency. This was calculated overall and within diagnostic groups, and the correlations among diagnostic groups were compared by using the Fisher r to z transformation method. Standard multiple sleep latency test categories (1, 2) (normal sleep latency [>10 minutes], moderate sleep latency [5 to 10 minutes], and severe sleep latency [<5 minutes]) were used and were compared with Epworth scores by using analysis of variance and the Tukey multiple comparison procedure. The ability to discriminate between pa-

Table. Association between Epworth Sleepiness Scale Scores and Mean Sleep Latency*

Epworth Scores (Quartiles)	Mean Sleep Latency		
	Normal (≥10 min)	Moderate (5–9 min)	Severe (<5 min)
	←	*n (%)*	→
<10	8 (32)	8 (32)	9 (36)
10–13	6 (26)	11 (48)	6 (26)
14–17	6 (21)	12 (41)	11 (38)
18–22	5 (20)	5 (20)	15 (60)

* Values given are numbers and percentages of patients. Pearson correlation between Epworth Sleepiness Scale scores and mean sleep latency, −0.17 (95% CI, −0.35 to 0.03).

tients with sleep latency of 10 minutes or less from those with sleep latency greater than 10 minutes was assessed by using receiver-operating characteristics analysis. This involved considering sleep latency of 10 minutes or less as the outcome and calculating the sensitivity and specificity for each observed value of the Epworth Sleepiness Scale. Nonparametric measures were used if the assumptions for the parametric methods were not met. A significance level of 0.05 was used for each hypothesis. Statistical analysis was done by using SAS software (SAS Institute, Cary, North Carolina).

Results

Eighty patients had sleep-disordered breathing (52 had sleep apnea, 13 had the upper-airway resistance syndrome, and 15 had primary snoring), 11 had periodic movements of sleep, and 1 had narcolepsy. Ten patients had less clearly defined diagnoses. Mean sleep latency (±SD) in the entire study sample was 7.5 ± 4.5 minutes.

In the total sample, no significant correlation was seen between Epworth scores and sleep latency (Pearson correlation, −0.17 [95% CI, −0.35 to 0.03]; P = 0.09). Analysis in subgroups formed according to the cause of sleepiness showed no significant correlation in patients with sleep-disordered breathing (n = 80) (correlation, −0.20 [95% CI, −0.05 to 0.40]; P = 0.08) or periodic limb movements of sleep (n = 11) (correlation, −0.22 [95% CI, −0.72 to 0.44]; $P >$ 0.2). The **Table** shows the number and percentage of patients in each sleep latency category for each quartile of the Epworth Sleepiness Scale.

No difference was seen in mean Epworth scores among the three groups defined by sleep latency with analysis of variance (overall, P = 0.13) or the Tukey multiple comparison procedure. Mean Epworth scores (±SD) were 12 ± 6 for patients with normal sleep latency, 12 ± 5 for patients with moderate sleep latency, and 14 ± 5 for patients with severe sleep latency. The **Figure** shows the distribution of the Epworth scores for the three sleep latency categories with side-by-side box plots.

A receiver-operating characteristic analysis was done to evaluate how well Epworth scores distinguished patients with sleep latency of 10 minutes or less from those with sleep latency greater than 10

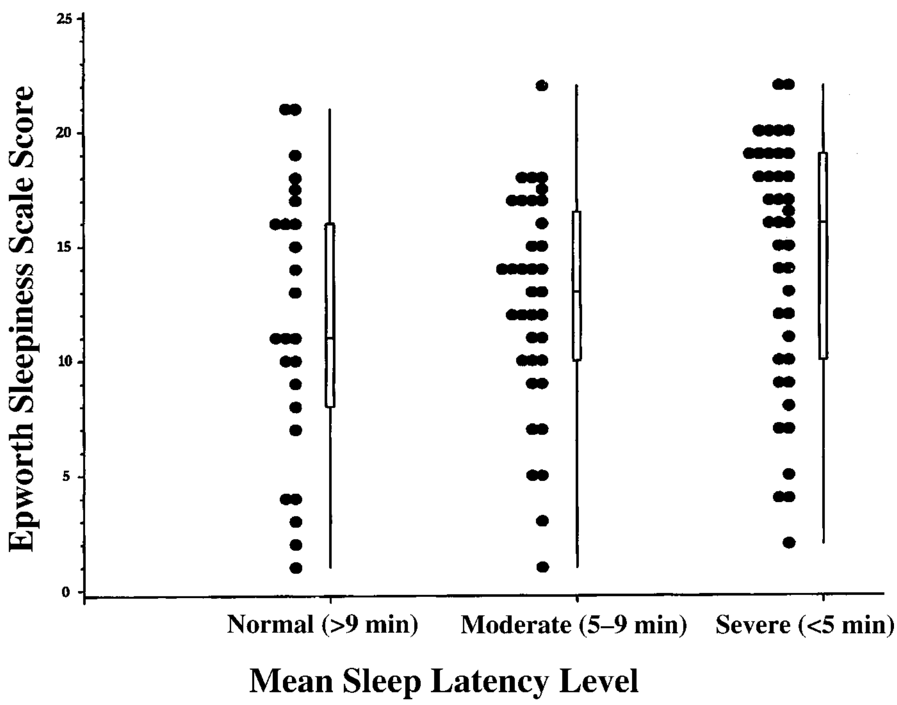

Figure. Distribution of Epworth Sleepiness Scale scores for the three sleep latency levels with side-by-side box plots. The top and bottom of each box are the 75th and 25th percentiles (quartiles) of the data, respectively. The middle line denotes the median. Box whiskers extend to data points that are within 1.5 interquartile ranges of the box.

minutes. The area under the curve was 0.61 (95% CI, 0.50 to 0.72), which is not statistically different from 0.50 (chance) ($P = 0.11$). Sensitivity and specificity were calculated for abnormal sleep latency ($\leq$10 minutes) at each observed Epworth score. No cut-point resulted in high estimates for both sensitivity and specificity. The optimal Epworth score was 11, with a sensitivity of 0.68 (95% CI, 0.56 to 0.78) and a specificity of 0.52 (95% CI, 0.31 to 0.72). The wide CIs are indicative of the moderate sample size for both sensitivity (77 abnormal sleep latencies) and specificity (25 normal sleep latencies).

Discussion

We found no statistically or clinically significant association between scores on the Epworth Sleepiness Scale (the most commonly used subjective sleepiness assessment test) and scores on the multiple sleepiness latency test (the most commonly used objective sleepiness assessment test). The results of our receiver-operating characteristics analysis were no better than random guessing; thus, the Epworth Sleepiness Scale could not reliably identify patients with abnormal and normal sleep latency. No Epworth score adequately identified patients with sleep latency of 10 minutes or less (**Table**). There was no Epworth score at which both sensitivity and specificity were high. Finally, the Epworth score could not distinguish between abnormal and normal sleep latency. As the **Figure** shows, the median Epworth score gradually increased with increases in the severity of sleep latency, but Epworth scores varied widely within each sleep latency level. Analysis by diagnostic groups also showed no association, but this analysis was limited by the small numbers of patients with conditions other than sleep-disordered breathing.

Other investigators have reported similar results. Sangal and colleagues (9) reported a significant but low correlation ($r = -0.27$; $P < 0.001$) between Epworth scores and sleep latency in 522 drug-free persons with narcolepsy. Similarly, Chervin and coworkers (7) studied a clinical population similar to ours and found a significant but moderate correlation of -0.37 ($P = 0.004$). In the original description of the Epworth Sleepiness Scale, Johns (4) reported a correlation of -0.514 ($P < 0.01$), but this was based on a smaller sample skewed toward very sleepy patients (11 had narcolepsy, 14 had idiopathic hypersomnia, and 2 had sleep apnea), with "concentrated" values for both Epworth scores and sleep latency. Correlation may not be the optimal measure of the usefulness of the Epworth scale for discrimination or prediction (12), but we and others (13, 14) also failed to show any categorical association between Epworth scores and sleep latency.

How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired?

This refers to your usual way of life in recent times. Even if you have not done some of these things recently, try to work out how they would have affected you.

Use the following scale to choose the most appropriate number for each situation:

- 0: would *never* doze
- 1: *slight* chance of dozing
- 2: *moderate* chance of dozing
- 3: *high* chance of dozing

Activity	**Chance of Dozing**
Sitting and Reading	______
Watching TV	______
Sitting inactive in a public place (meeting, theater, etc.)	______
As a passenger in a car for 1 hour without a break	______
Lying down in the afternoon when circumstances permit	______
Sitting and talking to someone	______
Sitting quietly after lunch without alcohol	______
In a car, while stopped for a few minutes in traffic	______
Total	______

Appendix Figure. The Epworth Sleepiness Scale. Each question is answered with a number from 0 (not at all likely to fall asleep) to 3 (very likely to fall asleep). This yields a total score of 0 (minimum) to 24 (maximum).

Thus, although the Epworth Sleepiness Scale may be predictive of other variables, such as general health status (6) or the respiratory disturbance index (4), our findings and those of others suggest that its association with sleep latency is tenuous.

Our study has several limitations. First, by definition, a sleep center evaluates sleepy patients; our sample was drawn from a tertiary referral center and thus may have been very sleepy. This may have resulted in a relatively homogeneous sample with high degrees of sleepiness and may have made it difficult to find a correlation. However, this sample was drawn from a true clinical population, which is where information on assessment of sleepiness may be most useful. Second, we used three epochs of stage 1 sleep to define sleep onset. Although this criterion (which defines "unequivocal sleep") is not the criterion most widely used in clinical practice, it is often used in research (1) and does not significantly affect the assessment of sleep latency (15). It is therefore unlikely that using the usual one-epoch criterion would have affected our results. Finally, with regard to the categorical associations, no universally accepted cut-off exists for abnormal sleep latency on the multiple sleep latency test, but 10 minutes is the most common (1, 2).

It is important for clinicians to know whether they should rely on objective or subjective methods to assess sleepiness. The explanation for the lack of a strong association between sleep latency and Epworth scores is uncertain. It is possible that this lack reflects a limitation of the Epworth Sleepiness Scale itself. However, many investigators have reported poor or no association between sleep latency and scores on other subjective tests, such as the Stanford Sleepiness Scale (16–19). Thus, it seems that the lack of association is not specific to the Epworth scale but reflects a general lack of association between subjective and objective measures of sleepiness. A plausible explanation is that subjective and objective tests may evaluate different aspects of sleepiness. This possibility is supported by the fact that the Epworth Sleepiness Scale may be associated with nocturnal sleep latency (on polysomnography) and the respiratory disturbance index (4) and is generally agreed upon (5, 7, 20).

In conclusion, subjective and objective methods for assessing sleepiness are probably complementary and may measure different aspects of sleepiness. Subjective scales, including the Epworth Sleepiness Scale, may be most useful as screening tools, and both types of tests should probably continue to be used in clinical practice as well as research (8–10).

Presented in part at the 11th Annual Meeting of the Associated Professional Sleep Societies, San Francisco, California, 3 June 1997.

Requests for Reprints: Selim R. Benbadis, MD, Department of Neurology and Division of Neurosurgery, University of South Florida College of Medicine, Harbourside Medical Tower, 4 Columbia Drive, Suite 730, Tampa, FL 33606; e-mail, sbenbadi@com1.med.usf.edu.

Current Author Addresses: Dr. Benbadis: Department of Neurology and Division of Neurosurgery, University of South Florida College of Medicine, Harbourside Medical Tower, 4 Columbia Drive, Suite 730, Tampa, FL 33606.
Mr. Mascha: Department of Biostatistics, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195.
Mr. Perry, Ms. Wolgamuth, and Dr. Dinner: Department of Neurology, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195.
Dr. Smolley: Department of Pulmonary Medicine, Cleveland Clinic Florida, 3000 West Cypress Creek Road, Ft. Lauderdale, FL 33309.

References

1. **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;9:519-24.
2. **Thorpy MJ.** The clinical use of the Multiple Sleep Latency Test. The Standards of Practice Committee of the American Sleep Disorders Association. Sleep. 1992;15:268-76.
3. **American Electroencephalographic Society.** Guidelines 15: guidelines for polygraphic assessment of sleep-related disorders (polysomnography). J Clin Neurophysiol. 1994;11:116-24.
4. **Johns MW.** A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. 1991;14:540-5.
5. **Johns MW.** Sleepiness in different situations measured by the Epworth Sleepiness Scale. Sleep. 1994;17:703-10.
6. **Briones B, Adams N, Strauss M, Rosenberg C, Whalen C, Carskadon M, et al.** Relationship between sleepiness and general health status. Sleep. 1996;19:583-8.
7. **Chervin RD, Aldrich MS, Pickett R, Guilleminault C.** Comparison of the results of the Epworth Sleepiness Scale and the Multiple Sleep Latency Test. J Psychosom Res. 1997;42:145-55.
8. Randomized trial of modafinil for the treatment of pathological somnolence in narcolepsy. US Modafinil in Narcolepsy Multicenter Study Group. Ann Neurol. 1998;43:88-97.
9. **Sangal RB, Mitler MM, Sangal JM.** MSLT, MWT and ESS: indices of sleepiness in 522 drug-free patients with narcolepsy [Abstract]. US Modafinil in Narcolepsy Multicenter Study Group. Sleep Res. 1997;26:492.
10. **Broughton RJ, Fleming JA, George CF, Hill JD, Kryger MH, Moldofsky H, et al.** Randomized, double-blind, placebo-controlled crossover trial of modafinil in the treatment of excessive daytime sleepiness in narcolepsy. Neurology. 1997;49:444-51.
11. **Kales A, Rechtschaffen A, eds.** A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Bethesda, MD: U.S. National Institute of Neurological Diseases and Blindness, Neurological Information Network; 1968.
12. **Greenland S, Schlesselman JJ, Criqui MH.** The fallacy of employing standardized regression coefficients and correlations as measures of effect. Am J Epidemiol. 1986;123:203-8.
13. **Smolley LA, Ivey C, Farkas M, Faucette E, Murphy S.** Epworth sleepiness scale is useful for monitoring daytime sleepiness [Abstract]. Sleep Res. 1993; 22:389.
14. **Sander AM, Mohan KK, Axelrod BN, Nahhas A, Kapen S.** The Epworth Sleepiness Scale: an unworthy adversary to clinical interview [Abstract]. Sleep Res. 1996;25:355.
15. **Benbadis SR, Perry MC, Wolgamuth BR, Mendelson WB, Dinner DS.** The multiple sleep latency test: comparison of sleep onset criteria. Sleep. 1996;8:632-6.
16. **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;7:230-8.
17. **Pressman MR, Fry JM.** Relationship of autonomic nervous system activity to daytime sleepiness and prior sleep. Sleep. 1989;12:239-45.
18. **Hoch CC, Reynolds CF 3d, Jennings JR, Monk TH, Buysse DJ, Machen MA, et al.** Daytime sleepiness and performance among healthy 80 and 20 year olds. Neurobiol Aging. 1991;13:353-6.
19. **Harnish MJ, Chard SR, Orr WC.** Relationship between measures of objective and subjective sleepiness [Abstract]. Sleep Res. 1996;25:492.
20. **Sangal RB, Sangal JM, Belisle C.** MWT and ESS measure different abilities in 41 patients with snoring and daytime sleepiness [Abstract]. Sleep Res. 1997;26:493.