



ELSEVIER

Sleep Medicine 4 (2003) 175–176

**SLEEP
MEDICINE**

www.elsevier.com/locate/sleep

Editorial

Epworth *Sleepiness* Scale?☆

A PubMed search for ‘Epworth *Sleepiness* Scale’ (ESS) at the time of this writing retrieves 226 articles. In 2002 alone, more than 60 were published, about double the number for ‘Multiple Sleep Latency Test’ (MSLT). Most of the studies appear to use the ESS as an important measure, and often the only measure of sleepiness. In the context of this explosion in use of the ESS during the past decade, the article by Miletin and Hanly in this issue of *Sleep Medicine* suggests the need for some pause and circumspection [1]. The authors review the development of the ESS and subject it to evaluation by standard criteria for scale development, validity, and reliability. They conclude that creation of the ESS did not include formal item development based on patient feedback. Content is sometimes ambiguous, and fails to include any questions unrelated to sleep propensity but nonetheless helpful in identification of sleepiness. Construct validity in clinical settings is moderate at best. Although internal consistency is good, test-retest reliability in clinical samples has not been demonstrated adequately.

Furthermore, Miletin and Hanly are in some ways more gentle than they could be. The Sleep Heart Health Study is cited as evidence of construct validity because the ESS score correlated with the apnea hypopnea index in this large sample [2]. However, the worst apneics in comparison to essentially normal subjects had an ESS score that was 9 rather than 7, only two points higher on a scale that ranges from 0 to 24. This suggests that in a sleep apnea patient with an ESS score more impressive to most clinicians – 14, for example – most of the excessive sleepiness may well be explained by factors other than the sleep apnea. Miletin and Hanly note convergence of ESS and quality of life survey results, but self-report measures often correlate with each other, especially if constructs overlap, as they may in this case. The authors note that in one study ESS scores correlated with the likelihood of falling asleep while driving. However, in this entirely retrospective study, the likelihood of falling asleep at the wheel was determined by self-report, in answer to a question that bears considerable resemblance to the eight ESS items [3]. Despite considerable statistical power in this survey of more than 4600 male

drivers, the authors themselves noted that ‘the overall effect of daytime sleepiness on accidents is less than clear cut’. The ESS showed statistical association with self-reported accidents mainly when complicated interactions with three other variables, such as driving a company car and snoring, were taken into account. Private car drivers who did not report feeling close to falling asleep at the wheel ($n = 2506$) tended to report *fewer* accidents when their ESS scores were *higher*, though the association did not reach significance.

Miletin and Hanly also raise a more basic question that is not often discussed: exactly what construct does the ESS measure? In the absence of a thorough physiological understanding of what mediates sleepiness, the best way to measure sleepiness may remain elusive for some time [4, 5]. However, many in our field have accepted sleep tendency, measured by a MSLT, as a basic objective indicator and reference measure of daytime sleepiness [6]. The ESS shares with the MSLT a focus on tendency to fall asleep. Most clinicians who use the ESS probably think of it as a subjective measure of *sleepiness*, as implied by the name of the instrument. However, the author who named the ESS has argued that it does not measure sleepiness, but rather a construct he believes to be clearly distinct, namely subjective sleep propensity [7]. Miletin and Hanly, Johns [8], and other authors [9,10] have tended to conclude that because ESS results diverge from results of the MSLT or its variants, these tests must measure different important things. Few give careful consideration to the possibility that the ESS, MSLT, or its variants may tap into undesirable confounds such as gender [11], anxiety [12], or other psychological factors [13].

Assessment of excessive daytime sleepiness is an important issue. Objective measurement is expensive, and hopefully alternatives more simple than the MSLT will one day show sufficient validity and reliability to replace it in clinical practice. The ESS filled a void when it appeared to offer an easy, brief, and inexpensive approach to assessment of chronic sleepiness. The ESS is now used by clinicians, their patients, the pharmaceutical industry, and academic centers. Unfortunately, users of the ESS may not have adequate information about its interpretation [14]. Much work remains to be done on the ESS and alternative methods for assessment of sleepiness. To be most useful, such measures must reflect well-defined constructs with good

☆ Supported by the NINDS (K02-NS02009).

* Michael S. Aldrich Sleep Disorders Laboratory, 8D8702 University Hospital, Box 0117, 1500 East Medical Center Drive, Ann Arbor, MI 48109-0117, USA. Tel.: +1-734-647-9064; fax: +1-734-647-9065.

validity and test-retest reliability. In the meantime, the ESS will continue to serve at least one important purpose: It does standardize collection of similar information between different clinicians or sites. The 226 ESS studies currently listed on PubMed are likely to have measured the same thing, even if their authors cannot be completely certain what it is that they measured.

References

- [1] Miletin MS, Hanly PJ. Measurement properties of the Epworth Sleepiness Scale. *Sleep Med* 2003;4:195–199.
- [2] Gottlieb DJ, Whitney CW, Bonekat WH, Iber C, et al. Relation of sleepiness to respiratory disturbance index: the Sleep Heart Health Study. *Am J Resp Crit Care Med* 1999;159:502–7.
- [3] Maycock G. Sleepiness and driving: the experience of U.K. car drivers. *Accid Anal Prev* 1997;29:453–62.
- [4] Chervin RD, Guilleminault C. Assessment of sleepiness in clinical practice. *Nat Med* 1995;1:1252–3.
- [5] Bliwise DL. Is the measurement of sleepiness the holy grail of sleep medicine? *Am J Resp Crit Care Med* 2001;163:1517–9.
- [6] Carskadon MA, Dement WC. The multiple sleep latency test: what does it measure? *Sleep* 1982;5:S67–S72.
- [7] Johns MW. Reply. *J Sleep Res* 2000;9:399–401.
- [8] Johns MW. Daytime sleepiness, snoring, and obstructive sleep apnea: the Epworth Sleepiness Scale. *Chest* 1993;103:30–6.
- [9] Sangal RB, Mitler MM, Sangal JM. Subjective sleepiness ratings (Epworth sleepiness scale) do not reflect the same parameter of sleepiness as objective sleepiness (maintenance of wakefulness test) in patients with narcolepsy. *Clin Neurophysiol* 1999;110:2131–5.
- [10] Sangal RB, Sangal JM, Belisle C. Subjective and objective indices of sleepiness (ESS and MWT) are not equally useful in patients with sleep apnea. *Clin Electroenceph* 1999;30:73–5.
- [11] Chervin RD, Aldrich MS. The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnea. *Neurology* 1999;52:125–31.
- [12] Naitoh P, Kelly TL. Modification of the Multiple Sleep Latency Test. In: Ogilvie RD, Harsh JR, editors. *Sleep onset: normal and abnormal processes*. New York: John Wiley & Sons, 1995. p. 327–38.
- [13] Olson LG, Cole MF, Ambrogetti A. Correlations among Epworth Sleepiness Scale scores, multiple sleep latency tests and psychological symptoms. *J Sleep Res* 1998;7:248–53.
- [14] Avidan AY, Chervin RD. ESS dot com. *Sleep Med* 2002;3:405–10.

Ronald D. Chervin*

*Sleep Disorders Center, Department of Neurology,
University of Michigan, Ann Arbor, MI, USA*
E-mail address: chervin@umich.edu (R.D. Chervin).