FAILURE OF THE MSLT AS A GOLD STANDARD TEST OF DAYTIME SLEEPINESS

Murray Johns
Epworth Sleep Centre, Epworth Hospital, Melbourne, Vic, 3121.

Excessive daytime sleepiness (EDS) is an important symptom that needs to be quantified, but there is uncertainty about how best to do that. Three of the most commonly used tests, the multiple sleep latency test (MSLT), the maintenance of wakefulness test (MWT) and the Epworth sleepiness scale (ESS) give results that are significantly correlated in a statistical sense, but are not closely related. The MSLT is often assumed to be the gold standard test and a "rule of thumb" has been adopted for its interpretation. However, this has not been validated. The purpose of this investigation was to help clarify this problem.

Methods: Previously published data from several investigations were used to calculate the reference range of normal values for each of these three tests, defined by the mean ± 2SD or by the 2.5 and 97.5 percentiles. Other previously published results from each test were also available for narcoleptic patients who were drug-free at the time and who by definition had EDS (1). This enabled the sensitivity and specificity of the 3 tests and their receiver operator characteristic curves to be compared for the first time in their ability to distinguish the EDS of narcolepsy from the daytime sleepiness of normal subjects.

Results: The reference range of mean sleep latencies (SL) in the MSLT for normal subjects is 11.5 ± 5.1 (SD) mins when calculated from the mean ± 2 SD i.e. 1.3 – 20 mins. However, such mean SL's are usually skewed in their distribution, so a more accurate definition of the reference range would be the 97.5 and 2.5 percentiles i.e. 3.2 – 20 mins. This is very different from the "rule of thumb" which many people assume gives a reference range of 10-20 mins.

The mean MSLT-SL had a sensitivity of only 80.9% and a specificity of 89.8% in distinguishing narcoleptics from normals when a cut-off SL < 5 mins was used, as suggested by the "rule of thumb". By contrast, the MWT-SL had a sensitivity of 84.3% and a specificity of 98.4% at a cut-off at < 12 mins. The ESS had a sensitivity of 93.5% and specificity of 100% with a cut-off score > 10. The receiver operator characteristic curves clearly show that, no matter what cut-off's are used, the ESS is more accurate than the MWT, and the MSLT is the least accurate of the 3 tests of daytime sleepiness.

Discussion: The "rule of thumb" which is so commonly used for the interpretation of the MSLT is grossly misleading and should be abandoned. The MSLT is the least accurate of the 3 tests of daytime sleepiness and the ESS the most accurate. The MSLT can no longer be considered a gold standard. More research is needed to develop an objective test that is as accurate as the ESS, but without its reliance on subjective reports, which is the main disadvantage of the ESS.

1. Sangal RB, Mitler MM, Sangal JAM. US modafinil in narcolepsy multicenter study group – MSLT, MWT and ESS: indices of sleepiness in 522 drug-free patients with narcolepsy. Sleep Res 1997; 26:492.

ASA 99.