Sleep and Breathing (2020) 24:369–377
https://doi.org/10.1007/s11325-019-01866-8

METHODS • REVIEW

Inside the clinical evaluation of sleepiness: subjective and objective tools

Simone Baiardi[1,2] · Susanna Mondini[3]

Received: 3 February 2019 / Revised: 29 April 2019 / Accepted: 13 May 2019 / Published online: 29 May 2019

Abstract

Purpose To critically review the available tools for evaluating excessive daytime sleepiness (EDS) in clinical practice.

Methods Objective tests and subjective scales were divided into three groups in accordance with the different dimensions of sleepiness they measure, namely physiological, manifest, and introspective. Strengths, weaknesses, and limitations of each test have been analysed and discussed along with the available recommendations for their use in clinical practice.

Results The majority of the tests developed for sleepiness evaluation do not have practical usefulness outside the research setting. The suboptimal correlation between different tests mainly depends on the different dimensions of sleepiness they analyse. Most importantly in-laboratory tests poorly correlate with sleepiness in real-life situations and, to date, none is able to predict the risk of injuries related to EDS, especially on an individual level.

Conclusions There exists not the one best test to assess EDS, however, clinicians can choose a more specific test to address a specific diagnostic challenge on the individual level. The development of novel performance tests with low cost and easy to administer is advisable for both screening purposes and fitness for duty evaluations in populations at high risk of EDS-related injuries, for example professional drivers.

Keywords Excessive daytime sleepiness · Epworth sleepiness scale · Multiple sleep latency test · Maintenance of wakefulness test · Psychomotor vigilance test · Driving simulation

Introduction

Sleepiness is a subjective feeling that expresses the individual need of sleep and may be thought of as a physiological state, like hunger, that is favoured by sleep deprivation and, conversely, is at least partially reduced by the attainment of an adequate amount of sleep. Sleepiness temporally characterizes the transition between full alertness and definite sleep states and is accompanied by multiple physiological changes, from cognitive and behavioural to biological modifications, which concur to create a complex and multifaceted condition. Although the process of falling asleep is mostly regulated by the integration of multiple internal stimuli including the homeostatic and circadian drives [1], the interaction with the environment also plays a significant role, given that several external cues and stimuli may influence alertness levels (e.g. sleepiness is usually promoted by monotonous situations, whereas it can be at least partially inhibited by noisy conditions). Given these introductory remarks, it is not surprising that, as for other physiological states, the precise limits distinguishing between normal (i.e. physiological) and pathological sleepiness are not well defined and often complicated to assess. In clinical practice, we refer to excessive daytime sleepiness (EDS), or alternatively to hypersomnolence, to indicate pathological sleepiness. EDS is defined as the inability to maintain wakefulness and alertness during the major waking episodes of the day with sleep occurring unintentionally or at inappropriate times [2]. This definition affords the chance to introduce further relevant aspects: first, to be considered pathological, sleepiness should interfere with individual

This article is part of the Topical Collection on *Excessive Daytime Sleepiness*

✉ Simone Baiardi
simone.baiardi6@unibo.it

[1] Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy

[2] IRCCS Istituto delle Scienze Neurologiche di Bologna, Ospedale Bellaria, Via Altura 1/8, 40139 Bologna, Italy

[3] Neurology Unit, Sant'Orsola-Malpighi University Hospital, Bologna, Italy

functioning during daytime; second, wakefulness (i.e. arousal) is not merely a synonym of alertness. Indeed, the latter is defined as the ability to sustain wakefulness and depends on the individual adaptation to actively respond to specific situations or tasks. According to the updated International Classification of Sleep Disorders (ICSD-3) [2], abnormal sleepiness should persist almost daily for at least 3 months to match the definition of EDS. It is relevant to highlight that this definition has been developed to identify sleep disorders associated with hypersomnolence (e.g. hypersomnias), whereas in clinical practice, physicians may be aware that short-lasting, abnormal sleepiness can also result from acute sleep deprivation, substance use, or other associated medical conditions.

Although epidemiological surveys are almost limited to subjective assessment, the prevalence of EDS is estimated to range from 10 to 25% in the general population [3], representing one of the most frequent complaints reported to clinicians and a significant individual and social issue. Indeed, EDS has a negative impact on several individual psychocognitive functions, including mood, concentration, memory, and attention, with consequent worsening of performance in a broad range of activities, which raises concerns on public safety while working and driving [4]. Given that the underlying causes of EDS are in most cases treatable (e.g. sleep disorders) or reversible (e.g. insufficient sleep time), the awareness and early recognition of EDS may improve the quality of life of affected subjects and at the same time reduce driving accidents and work-related injuries.

After a comprehensive review of scientific literature searching the PubMed database for English-language articles using the terms (alone and in combination) "excessive daytime sleepiness", "hypersomnia", "alertness", "multiple sleep latency test", "maintenance of wakefulness test", "Epworth sleepiness scale", "performance", "driving simulation" (including all commonly used abbreviations of these terms), here, we critically discuss the strengths and weaknesses of the most diffuse tools used in clinical practice for the evaluation of patients with suspected pathological sleepiness. Therefore, we have arbitrarily chosen to omit from this review a large part of research works given that they do not yet have clinical application.

Sleepiness assessment

Although sleepiness is largely a subjective phenomenon, it may be evaluated by both subjective and objective tools. Given its complex and multifaceted characteristics, the tests developed to measure EDS often allow us to analyse only a particular appearance of the phenomenon. Accordingly, to provide a more comprehensive interpretation of individual sleepiness often multiple tests should be associated, although they sometimes produce contrasting results [5].

In this regard, the collection of an accurate anamnesis as well as the physical examination is fundamental to address the diagnostic work-up correctly and to choose the most appropriate test to assess sleepiness [6]. Other objectives of clinical evaluation should include the distinction between physiological and pathological sleepiness, the definition of EDS features, the differential diagnosis with clinical mimics such as fatigue, the characterization of the individual sleep habits, and the identification of clues suggesting the underlying aetiology (Table 1). Furthermore, clinical history should consider the negative consequences of EDS and whether they constitute a risk for patient safety (e.g. professional drivers).

Tools for the evaluation of sleepiness may be classified into subjective and objective tests, but from a clinical point of view, it is more appropriate to distinguish them according to the dimensions of sleepiness they measure. A broadly accepted model proposed by Carskadon and Dement recognizes three main dimensions of sleepiness: introspective, physiological, and manifest [7] (Fig.1).

Introspective sleepiness

Introspective sleepiness refers to the individual self-assessment of the phenomenon and generally constitutes the first step of clinical evaluation of patients with suspected EDS. Given the subjective nature of sleepiness, the individual self-assessment is potentially the most accurate, but, at the same time, it is highly dependent on the ability to be aware, to recognize, and to report this feeling. In other words, when asked to quantify sleepiness, a subject might over- or under-report it on the basis of differences in internal perception and in individual estimation of "normal" sleepiness or of intentional misrepresentation. Several subjective scales have been developed for the self-quantification of sleepiness and, according to the temporal interval they evaluate, they are categorized into two main groups, which analyse sleepiness in a defined moment of the day ("state" scales) or over a defined period of time ("trait" scales).

The scales belonging to the first group answer to the question "how does the patient feel now?" providing a momentary assessment of sleepiness and are therefore sensitive to changes in the intensity of the symptom. Given these proprieties, they are useful in research applications to evaluate, through repeated administrations, the circadian oscillation of sleepiness or the variation of its levels that are induced by drugs or sleep deprivation. Among the state scales, those that need to be mentioned are the Karolinska sleepiness scale, which is a 10-point self-administering scale ranging from 1 (patient extremely alert) to 10 (extremely sleepy) [8], and the Stanford sleepiness scale, which has a similar structure but ranges from

Table 1 Overview of most frequent conditions associated to excessive daytime sleepiness

Behavioural
Acute sleep deprivation
Insufficient sleep syndrome
Medications
Antihistamines, barbiturates, opioids, hypnotics, anxiolytics, antidepressants, neuroleptics, antiepileptics, beta-blockers, dopamine agonists
Central disorders of hypersomnolence
Narcolepsy types 1 and 2
Idiopathic hypersomnia
Kleine-Levine syndrome
Sleep-related breathing disorders
Obstructive sleep apnea disorders
Central sleep apnea syndromes
Sleep-related hypoventilation and/or hypoxemia disorders
Sleep-related movement disorders
Periodic limb movement disorder
Other sleep disorders[#]
Insomnia
Restless legs syndrome/Willis-Ekbom disease
Delayed sleep-wake phase disorder
Advanced sleep-wake phase disorder
Jet lag
Other circadian rhythm sleep-wake disorders
Psychiatric
Substance abuse
Depression
Other neurological disorders
Multiple sclerosis
Neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease)
Stroke
Traumatic brain injury
Myotonic dystrophy
Neoplasms
Toxic or metabolic encephalopathies (e.g. hepatic and renal failure)
Other medical conditions
Hypothyroidism
Fibromyalgia

[#] when leading to sleep loss

1 to 7 [9]. These scales provide a discrete rating and for each point a brief description of intensity of the symptom; conversely, the visual analogue scales allow a continuous quantification of sleepiness and the intermediate degrees of sleepiness between the maximum and minimum extremities are not specified in detail, not forcing the subject into discrete definitions.

Sleepiness trait scales are more useful to evaluate patients with suspected EDS since they ask the subjects to average their levels of sleepiness in different situations and various moments of the day providing information about the symptoms over a prolonged period of time. The Epworth sleepiness scale (ESS), the most popular scale of this group, was originally developed by Murray W. Johns in 1991 to evaluate daytime sleepiness in healthy subjects and in patients with a variety of sleep disorders [10]. This questionnaire asks the patient to rate on a 4-point scale, ranging from zero to three, the probability to doze off in eight situations of daily life, that are known to be soporific, based on their usual habits. The total score is the sum of the eight items' scores and ranges between 0 and 24; the higher the final score is, the more pronounced the sleepiness feeling. A cut-off of ≥ 10 points is accepted to indicate EDS. Several factors contributed to the large diffusion of ESS for sleepiness evaluation: the questionnaire is self-administered, easy to complete and score, costless and therefore suitable for a quick evaluation of EDS in the outpatient setting. Currently, ESS is considered a useful tool worldwide for sleepiness evaluation/screening and has been translated and validated in several languages and countries. Recently, a modified version of the ESS adapted for children and adolescents has been developed and validated [11, 12]. Concerning psychometric proprieties of the ESS, it has a good internal consistency in clinical samples, which makes it a useful tool for the comparison of groups of patients [13]. Conversely, the utility of the ESS as a screening tool in non-clinical samples (e.g. students, community samples) is debated, given the lower internal consistency in these populations [14, 15]. Moreover, physicians should be aware that some patients, such as professional drivers, could intentionally underscore symptoms due to a perceived threat to their driving license, making the ESS unreliable [16]. A further limit which may impact on the clinical use of ESS concerns the limited evidence of test-retest reliability [13, 17], in particular when the questionnaire administrations are performed in a short time interval, whereas the poor correlation (i.e. from moderate to weak) with other objective tests should not surprise since they evaluate sleepiness from different standpoints [13]. Finally, some authors challenged the conventional self-administration of the ESS, reporting an increased accuracy when the scale is physician administered [18].

Despite the lesser diffusion, the Sleep-Wake Activity Inventory (SWAI) [19] and the Toronto Hospital Alertness Test (THAT) [20] are validated and useful tools for the evaluation of sleepiness and alertness, respectively, in the clinical setting.

Physiological sleepiness

Physiological sleepiness, also called "sleep pressure", refers to the biological drive that regulates the process of falling asleep. From this standpoint, the rapidity with which a subject

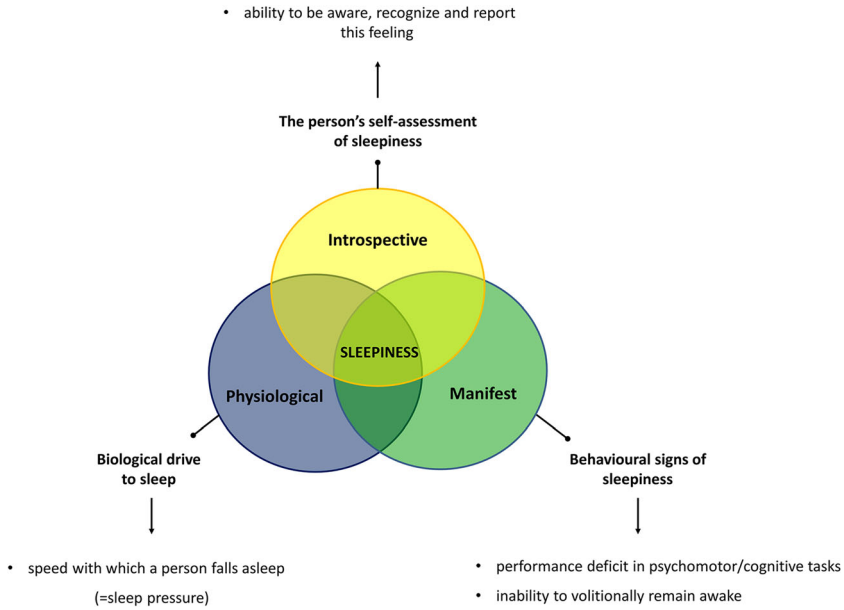

Fig. 1 Representation of the model for sleepiness characterization proposed by Carskadon and Dement [7], which addresses the available tools for EDS evaluation in three main groups analysing the symptom form three different standpoints

falls asleep is representative of sleepiness intensity [21]. Notably, the latter is influenced not only by the homeostatic drive but also by the circadian one, resulting in a bimodal distribution of sleepiness during 24/h, which includes a major nocturnal sleepiness crest and a secondary mid-afternoon peak [22, 23]. These peaks are separated by the so-called "forbidden zone" for sleep in which sleep pressure is reduced [22, 23]. Physicians should be aware of these ultradiurnal changes of sleepiness for a correct interpretation of tests measuring physiological sleepiness.

Multiple sleep latency test

The most studied measure to assess physiological sleepiness is the Multiple Sleep Latency Test (MSLT) [24]. The MSLT is a validated and objective tool measuring the ability or tendency to fall asleep under standardized conditions [25, 26]. This test was developed on the a priori assumption that the shorter the sleep latency, the higher the degree of sleepiness. In clinical practice, the MSLT has indication for the evaluation of patients with suspected narcolepsy to confirm the diagnosis and to differentiate narcolepsy from idiopathic hypersomnias, whereas it is not routinely used to evaluate sleepiness associated with other medical conditions or sleep disorders such as sleep-related breathing disorders, insomnia, and circadian rhythm sleep-wake disorders [25]. Notably, the MSLT is the only measure of sleepiness also contributing to the diagnosis of a specific sleep disorder that is narcolepsy.

The MSLT should be performed in a sleep laboratory according to a standardized protocol (for extensive methodological information see [25]). To rule out acute sleep deprivation, the night before MSLT administration, it is mandatory to perform a polysomnography documenting at least 6 h of sleep. The test consists of four or five nap opportunities performed at 2-h intervals; in each nap, the patient lies in a dark and quiet room and should try to fall asleep. During each session, the patient undergoes continuous neurophysiologic monitoring, including central (C3-A2, C4-A1) and occipital (O1-A2, O2-A1) electroencephalographic derivations, left and right electrooculogram, superficial electromyogram of mylohyoid muscle and electrocardiogram, which allow us to evaluate the sleep onset and the presence of sleep onset rapid eye movement sleep periods (SOREMPs). Sleep latency is defined as time between lights-off and sleep onset while SOREMPs are defined as the emergence of rapid eye movement sleep within 15 min from sleep onset [24]. The MSLT has to be conducted by an experienced technologist who in the absence of sleep interrupts the test after 20 min, whereas if the subject falls asleep, the test continues for a further 15 min from sleep onset in order to evaluate the presence of a SOREMP. In clinical practice, a mean sleep latency below 8 min is considered indicative of EDS [2]. According to the ICSD-3, the recording of a SOREMP on polysomnography the night before the test has the same diagnostic value for narcolepsy of those recorded during the MSLT [2].

Together with its application as a diagnostic tool, the very high interrater and intrarater reliability for both mean sleep latency and SOREMPs [27] as well as the high test-retest reliability [28] (although recently disputed in the longitudinal evaluation of central hypersomnias different from narcolepsy type 1 [29]) certainly contributed to the success of this test.

Nevertheless, the MSLT has also some weaknesses limiting sweeping applications. First, it is expensive and time consuming, therefore unsuitable for EDS screening in large populations. Second, it is highly sensitive to sleep deprivation [30] and consequently also the debt of few hours of sleep may induce healthy subjects to fall into pathological results. Third, it has a floor effect for severe degrees of sleepiness. Fourth, its application to children lacks normative data. Fifth, some subjects with circadian misalignment such as shift workers could need to perform the test outside the usual testing hours (8 a.m.–6 p.m.), but in alternative hours, MSLT has not been validated.

Other measures of physiological sleepiness

In the research setting, other tests evaluating a variety of physiologic variables that are known to be influenced by vigilance changes have also been developed. Given the well-known modifications of the EEG power spectrum between wakefulness and sleep [31], which are more pronounced in sleep-deprived subjects [32], the analysis of EEG signals has been proposed as a useful tool to evaluate sleepiness [33, 34]. In line with this evidence, self-perceived sleepiness has been associated with an increase of theta power spectrum [35, 36] and a concurrent decrease of alpha [36]. Nevertheless, the lack of normative data and standardized protocols, as well as the significant inter-individual variability, limits its broad application in clinical practice.

The sleep-wake transition is also characterized by a reduced eye-blink activity and an increased density of slow eye movement [37]. Ocular activity and eyelid drop have been evaluated by computational analysis of electrooculogram [38] or by visual assessment of the percentage of eye closures [39]. The former showed more promising preliminary results for application in the clinical setting, while the latter provided the basis for the development of non-contact video systems allowing the detection of eyelid closures while driving which represents an interesting tool for motor vehicle accident prevention in patients with EDS.

Physiological sleepiness can also be measured by analysing spontaneous pupil oscillations while the patient is in a dark room (pupillography) [40]. In a state of full wakefulness, the dark environment should induce mydriasis. However, when the subject is sleepy, the predominance of parasympathetic over sympathetic autonomic nervous system activity induces both constriction of pupillary diameter and pupil size instability [40]. Pupillographic metrics showed a good correlation with sleep latency on MSLT [41] and resting EEG activity [42]. Normative data are available for subjects with age ranging from 20 to 60 years [43], while elderly shows most frequently unreliable results [44]. Given this limit, the test is far from clinical application due to the lack of well-defined clinical cut-points distinguishing physiological from pathological sleepiness.

Manifest sleepiness

Manifest sleepiness refers to identifiable and measurable symptoms, signs, and behaviours revealing that a person has a decreased alertness [7]. It represents the external manifestation of the complex balance between sleep and wakefulness biological mechanisms, when the former prevails on the latter. The assessment of manifest sleepiness has a pivotal relevance for the evaluation of individual safety, since this aspect of sleepiness may impact on the ability to volitionally remain awake and on performances in cognitive or psychomotor tasks.

Maintenance of wakefulness test

The Maintenance of Wakefulness Test (MWT) is a validated and objective measure of the ability to stay awake for a defined time in an environment with low levels of stimulation [25, 26]. It is recommended for the objective assessment of the individual's ability to remain awake when the inability would constitute a public or personal safety issue and for the evaluation of response to treatment in subjects affected by conditions associated with EDS [25, 26]. Accordingly, the MWT is not a diagnostic tool. It is an in-laboratory standardized test that, paralleling the MSLT construct, consists of four sessions performed with 2-h intervals under the same neurophysiologic monitoring (EEG, EOG, EMG, and EKG). Although a variety of protocols and interpretations have been proposed [45, 46], the 40-min protocol is recommended [25]. At variance with the MSLT, the patient is seated (not lying) on bed with the back and head supported by a bolster pillow and should try not to fall asleep for as long as possible. Moreover, although during the test the room should be shielded from external light, a low intensity indirect light source (0.1–0.13 lux), positioned slightly behind the patient, is allowed and clinicians have to decide whether to perform a night polysomnography (not obligatory) the night before the test. The test is performed under the supervision of a trained sleep technician and ends after 40 min if no sleep occurs, otherwise after the appearance of three consecutive epochs of N1 stage sleep or one epoch of any other sleep stage. The correct interruption of each test session is fundamental in order not to dissipate the homeostatic sleep drive. Sleep latency is defined by the interval between lights-off and the first epoch of unequivocal sleep. Since several stimuli may impact sleep latency, before

MWT administration, patients are instructed not to consume stimulants such as tobacco and coffee between each test session, and to avoid extreme self-activating behaviours to improve alertness (e.g. slapping, talking) during each session. According to normative data for the 40-min protocol [25, 45], a mean sleep latency below 8 min is considered pathological, while values greater than 30 min indicate normal alertness. Sleep latencies between 8 and 30 min are of uncertain significance, often labelled as "borderline". To date, the lack of a well-defined cut-off defining normality is probably the main limit of MWT, especially when used to assess sleepiness for the driving license. Indeed, although not explicitly pathological, borderline results cannot be considered normal and either the patient's evaluation should include alternative measures of sleepiness or the MWT should be repeated, despite the uncertain test-retest reliability [45, 47]. Another limit is the evidence of a ceiling effect in patients with normal alertness; in these subjects, MWT sessions end after 40 min and the sleep latency is consequently calculated (i.e. 40 min), although they could still remain awake for a long time [25, 26]. Although the MWT has been increasingly used in the evaluation of workers with known high-risk of professional accidents related to EDS (e.g. commercial drivers, shift workers), clinicians should be aware that the test has limitations in predicting performance in real-life situations, which is also influenced by multiple environmental factors that cannot be accounted for in the laboratory setting. A further criticism of MWT is the poor correlation with other sleepiness measures [48, 49].

Performance tests

Beside the MWT, a variety of performance tests have also been developed to indirectly assess sleepiness, or better alertness, based on the ability to respond to tasks with different levels of complexity. Notably, to be reliable, they require that tested subjects have a normal intellectual profile and physical integrity. For this reason and for the substantial lack of normative data, performance tests are mainly used for research purposes. Similarly to MWT, the Oxford Sleep Resistance (OSLER) test consists of four 40-min sessions during which the patient should try to respond to a visual signal given every 3 s by pressing a button [50]. Neurophysiologic monitoring during the test is not performed. The test ends at the 40th minute or when the subject misses responses to seven consecutive stimuli (i.e. 21 s) and then "sleep latency" is calculated. This tool has been validated against the MWT [51] and is sensitive to vigilance modification following treatment in patients with obstructive sleep apnea [52]. The main limit of the OSLER test is once again the lack of normative data supporting clinical cut-points. Moreover, the test cannot be administered to subjects with motor and visual deficits, which may interfere with pushing the button or recognizing visual stimuli.

The psychomotor vigilance test (PVT) consists of a simple reaction time in response to visual or acoustic stimuli given randomly for 10 min by pressing a button. Shorter duration protocols also have been extensively studied [53, 54]. The test gives information about both the accuracy, defined by the number of lapses, and the speed of responses. A prolonged latency of responses to stimuli and multiple lapses have been associated with decreased vigilance in sleep-deprived subjects [55] and in patients with sleep-breathing disorders [56, 57]. The construct and interpretation of the Sustained Attention to Response Task (SART) are very similar to PVT but for a more complex go/no-go task (e.g. to press a button when a number (1–9) appears on a screen, except when that number is a 3) [58]. The test has been used in patients affected by a variety of sleep disorders characterized by hypersomnolence, including narcolepsy [59], idiopathic hypersomnia, and obstructive sleep apnea syndrome, and demonstrated a good correlation with the vigilance impairment irrespective of the cause [60].

Finally, to objectify sleepiness when driving, researchers focused their interest on driving simulation tasks and real-road driving tests. A great variety of driving simulators have been independently developed, and data from scientific literature can only in part be comparatively analysed due to the lack of a standardized simulator type and setting and different test duration [61]. Analogously, the parameters analysed also differ among the studies, with near-miss accidents and inappropriate line crossings being the most frequently evaluated [61]. In sleep-deprived subjects sleepiness affected both simulated and real driving, but for slower reaction times and more pronounced line crossings in the simulated condition [62]. Driving simulators showed a good correlation with the MWT [63–67] and a moderate association with real-road driving performance [61], but on the basis of current evidence, they are not able to reliably predict real-life accidents, especially when results are evaluated on an individual level [61].

Conclusive remarks

Sleepiness is a multifaceted physiological state and, although a variety of tools are available for EDS evaluation, to date, there is not a single objective test or subjective scale that, alone, can assess sleepiness complexity. Given that different tools investigate various dimensions of sleepiness, the clinical evaluation has a pivotal role in addressing the

Table 2 Summary of the most diffuse tools for sleepiness evaluation and their indications in clinical practice

Test	Normative data	Clinical indications
Introspective sleepiness		
SSS	No	Research only
KSS	Limited	Research only
ESS	Yes	Self-assessment of sleepiness perception, screening for EDS in the outpatient setting
SWAI	Yes	Self-assessment of sleepiness perception
THAT	Yes	Self-assessment of alertness
Physiological sleepiness		
MSLT	Yes	Diagnosis of narcolepsy, differential diagnosis between narcolepsy and idiopathic hypersomnia
Pupillography	Limited	Research only
EOG	No	Research only
EEG	No	Research only
Manifest sleepiness		
MWT	Yes	To test the volitional ability to remain awake in soporific situation in response to treatment or in "fitness for duty" evaluation in workers at high-risk for EDS
OSLER	No	Research only
PVT	Limited	Research only
SART	No	Research only
Driving simulation	No	Research only

SSS, Stanford sleepiness scale; *KSS*, Karolinska sleepiness scale; *ESS*, Epworth sleepiness scale; *MSLT*, multiple sleep latency test; *EOG*, electrooculogram; *EEG*, electroencephalography; *MWT*, maintenance of wakefulness test; *OSLER*, Oxford Sleep Resistance test; *PVT*, psychomotor vigilance test; *SART*, sustained attention to response task; *EDS*, excessive daytime sleepiness

diagnostic work-up and the choice of the most advisable test(s). Often, to address the clinical challenge, it is mandatory to perform multiple tests and scales. Moreover, most tests have no practical usefulness outside the research setting (Table 2) and all the subjective and objective measures of sleepiness have weaknesses and limitations. In this regard, it is remarkable that the results of objective tests are strongly influenced by individual volition and possibly by intentional lack of cooperation, whereas the main limitation to the interpretation of results and subsequent translation into clinical decisions is the poor correlation between in-laboratory measures and real-life situations. Future research should aim to develop inexpensive, less time-consuming objective tools for EDS evaluation and to demonstrate the relationship between the results obtained and the real-life performance. From the standpoint of occupational safety, initial evidence indicates that performance tests are advisable in the evaluation of populations with a high risk of injuries due to EDS, such as professional drivers. Paralleling these advisable advances on clinical evaluation of EDS, the development of accurate non-contact video monitoring systems detecting eyelid closure while operating a motor vehicle (with both professional and non-professional purposes) is expected to improve individual and public safety reducing the risk of EDS-related accidents.

Acknowledgments The authors wish to thank Cecilia Baroncini for the English revision.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Research involving human participants This article does not contain any studies with human participants performed by any of the authors.

Informed consent Not applicable.

References

1. Borbély AA (1982) A two process model of sleep regulation. Hum Neurobiol 1:195–204
2. American Academy of Sleep Medicine (2014) The international classification of sleep disorders, third edition (ICSD-3). American Academy of Sleep Medicine, Darien

3. Young TB (2004) Epidemiology of daytime sleepiness: definitions, symptomatology, and prevalence. J Clin Psychiatry 65(Suppl 16): 12–16
4. Häkkänen H, Summala H (2000) Sleepiness at work among commercial truck drivers. Sleep 23:49–57
5. Sullivan SS, Kushida CA (2008) Multiple sleep latency test and maintenance of wakefulness test. Chest 134:854–861. https://doi.org/10.1378/chest.08-0822
6. Pizza F (2014) Sleepiness assessment. In: Sleepiness and human impact assessment. Springer-Verlag Mailand, pp 313–324
7. Carskadon MA, Dement WC (1987) Daytime sleepiness: quantification of a behavioral state. Neurosci Biobehav Rev 11:307–317
8. Akerstedt T, Gillberg M (1990) Subjective and objective sleepiness in the active individual. Int J Neurosci 52:29–37
9. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC (1973) Quantification of sleepiness: a new approach. Psychophysiology 10:431–436
10. Johns MW (1991) A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 14:540–545
11. Johns MW (2015) The assessment of sleepiness in children and adolescents. Sleep Biol Rhythm:97
12. Janssen KC, Phillipson S, O'Connor J, Johns MW (2017) Validation of the Epworth Sleepiness Scale for children and adolescents using Rasch analysis. Sleep Med 33:30–35. https://doi.org/10.1016/j.sleep.2017.01.014
13. Kendzerska TB, Smith PM, Brignardello-Petersen R, Leung RS, Tomlinson GA (2014) Evaluation of the measurement properties of the Epworth sleepiness scale: a systematic review. Sleep Med Rev 18:321–331. https://doi.org/10.1016/j.smrv.2013.08.002
14. Johns MW (1992) Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep 15:376–381
15. Johns MW (1994) Sleepiness in different situations measured by the Epworth Sleepiness Scale. Sleep 17:703–710
16. Baiardi S, La Morgia C, Sciamanna L et al (2018) Is the Epworth Sleepiness Scale a useful tool for screening excessive daytime sleepiness in commercial drivers? Accid Anal Prev 110:187–189. https://doi.org/10.1016/j.aap.2017.10.008
17. Nguyen ATD, Baltzan MA, Small D, Wolkove N, Guillon S, Palayew M (2006) Clinical reproducibility of the Epworth sleepiness scale. J Clin Sleep Med 2:170–174
18. Damiani MF, Quaranta VN, Falcone VA, Gadaleta F, Maiellari M, Ranieri T, Fanfulla F, Carratù P, Resta O (2013) The Epworth Sleepiness Scale: conventional self vs physician administration. Chest 143:1569–1575. https://doi.org/10.1378/chest.12-2174
19. Rosenthal L, Roehrs TA, Roth T (1993) The sleep-wake activity inventory: a self-report measure of daytime sleepiness. Biol Psychiatry 34:810–820
20. Shapiro CM, Auch C, Reimer M, Kayumov L, Heslegrave R, Huterer N, Driver H, Devins GM (2006) A new approach to the construct of alertness. J Psychosom Res 60:595–603. https://doi.org/10.1016/j.jpsychores.2006.04.012
21. Borbély AA, Daan S, Wirz-Justice A, Deboer T (2016) The two-process model of sleep regulation: a reappraisal. J Sleep Res 25: 131–143. https://doi.org/10.1111/jsr.12371
22. Lavie P (1986) Ultrashort sleep-waking schedule. III. "Gates" and "forbidden zones" for sleep. Electroencephalogr Clin Neurophysiol 63:414–425
23. Dijk DJ, Czeisler CA (1995) Contribution of the circadian pacemaker and the sleep homeostat to sleep propensity, sleep structure, electroencephalographic slow waves, and sleep spindle activity in humans. J Neurosci 15:3526–3538
24. Carskadon MA, Dement WC, Mitler MM, Roth T, Westbrook PR, Keenan S (1986) Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. Sleep 9:519–524
25. Littner MR, Kushida C, Wise M, Davila DG, Morgenthaler T, Lee-Chiong T, Hirshkowitz M, Daniel LL, Bailey D, Berry RB, Kapen S, Kramer M, Standards of Practice Committee of the American Academy of Sleep Medicine (2005) Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. Sleep 28:113–121
26. Arand D, Bonnet M, Hurwitz T, Mitler M, Rosa R, Sangal RB (2005) The clinical use of the MSLT and MWT. Sleep 28:123–144
27. Drake CL, Rice MF, Roehrs TA, Rosenthal L, Guido P, Roth T (2000) Scoring reliability of the multiple sleep latency test in a clinical population. Sleep 23:911–913
28. Zwyghuizen-Doorenbos A, Roehrs T, Schaefer M, Roth T (1988) Test-retest reliability of the MSLT. Sleep 11:562–565
29. Lopez R, Doukkali A, Barateau L, Evangelista E, Chenini S, Jaussent I, Dauvilliers Y (2017) Test-retest reliability of the multiple sleep latency test in central disorders of hypersomnolence. Sleep 40. https://doi.org/10.1093/sleep/zsx164
30. Bonnet MH, Arand DL (1995) 24-hour metabolic rate in insomniacs and matched normal sleepers. Sleep 18:581–588
31. De Gennaro L, Ferrara M, Bertini M (2001) The boundary between wakefulness and sleep: quantitative electroencephalographic changes during the sleep onset period. Neuroscience 107:1–11
32. Borbély AA, Baumann F, Brandeis D, Strauch I, Lehmann D (1981) Sleep deprivation: effect on sleep stages and EEG power density in man. Electroencephalogr Clin Neurophysiol 51:483–495
33. Putilov AA, Donskaya OG (2014) Alpha attenuation soon after closing the eyes as an objective indicator of sleepiness. Clin Exp Pharmacol Physiol 41:956–964. https://doi.org/10.1111/1440-1681.12311
34. Putilov AA, Donskaya OG, Verevkin EG (2017) Generalizability of frequency weighting curve for extraction of spectral drowsy component from the EEG signals recorded in eyes-closed condition. Clin EEG Neurosci 48:259–269. https://doi.org/10.1177/1550059416673271
35. Cajochen C, Brunner DP, Kräuchi K et al (1995) Power density in theta/alpha frequencies of the waking EEG progressively increases during sustained wakefulness. Sleep 18:890–894
36. Strijkstra AM, Beersma DGM, Drayer B, Halbesma N, Daan S (2003) Subjective sleepiness correlates negatively with global alpha (8-12 Hz) and positively with central frontal theta (4-8 Hz) frequencies in the human resting awake electroencephalogram. Neurosci Lett 340:17–20
37. Torsvall L, Akerstedt T (1988) Extreme sleepiness: quantification of EOG and spectral EEG parameters. Int J Neurosci 38:435–441
38. Fabbri M, Pizza F, Magosso E, Ursino M, Contardi S, Cirignotta F, Provini F, Montagna P (2010) Automatic slow eye movement (SEM) detection of sleep onset in patients with obstructive sleep apnea syndrome (OSAS): comparison between multiple sleep latency test (MSLT) and maintenance of wakefulness test (MWT). Sleep Med 11:253–257. https://doi.org/10.1016/j.sleep.2009.05.020
39. Sommer D, Golz M (2010) Evaluation of PERCLOS based current fatigue monitoring technologies. Conf Proc IEEE Eng Med Biol Soc 2010:4456–4459. https://doi.org/10.1109/IEMBS.2010.5625960
40. Wilhelm B, Giedke H, Lüdtke H et al (2001) Daytime variations in central nervous system activation measured by a pupillographic sleepiness test. J Sleep Res 10:1–7
41. Yamamoto K, Kobayashi F, Hori R, Arita A, Sasanabe R, Shiomi T (2013) Association between pupillometric sleepiness measures and sleep latency derived by MSLT in clinically sleepy patients. Environ Health Prev Med 18:361–367. https://doi.org/10.1007/s12199-013-0331-0
42. Regen F, Dorn H, Danker-Hopfe H (2013) Association between pupillary unrest index and waking electroencephalogram activity in sleep-deprived healthy adults. Sleep Med 14:902–912. https://doi.org/10.1016/j.sleep.2013.02.003

43. Wilhelm B, Korner A, Heldmaier K, Moll K, Wilhelm H, Ludtke H (2001) Normwerte des pupillographischen Schlafrigkeitstests für Frauen und Männer zwischen 20 und 60. Jahren Somnologie 5: 115–120
44. Eggert T, Sauter C, Popp R, Zeitlhofer J, Danker-Hopfe H, on behalf of the working group "Vigilance" of the German Society for Sleep Research and Sleep Medicine (DGSM) (2012) The pupillographic sleepiness test in adults: effect of age, gender, and time of day on pupillometric variables. Am J Hum Biol 24:820–828. https://doi.org/10.1002/ajhb.22326
45. Doghramji K, Mitler MM, Sangal RB, Shapiro C, Taylor S, Walsleben J, Belisle C, Erman MK, Hayduk R, Hosn R, B. O'Malley E, Sangal JAM, Schutte SL, Youakim JM (1997) A normative study of the maintenance of wakefulness test (MWT). Electroencephalogr Clin Neurophysiol 103:554–562
46. Arzi L, Shreter R, El-Ad B et al (2009) Forty- versus 20-minute trials of the maintenance of wakefulness test regimen for licensing of drivers. J Clin Sleep Med 5:57–62
47. Mitler MM, Doghramji K, Shapiro C (2000) The maintenance of wakefulness test: normative data by age. J Psychosom Res 49:363–365
48. Sangal RB, Thomas L, Mitler MM (1992) Maintenance of wakefulness test and multiple sleep latency test. Measurement of different abilities in patients with sleep disorders. Chest 101:898–902
49. Sangal RB, Mitler MM, Sangal JM (1999) 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 110:2131–2135
50. Bennett LS, Stradling JR, Davies RJ (1997) A behavioural test to assess daytime sleepiness in obstructive sleep apnoea. J Sleep Res 6:142–145
51. Krieger AC, Ayappa I, Norman RG, Rapoport DM, Walsleben J (2004) Comparison of the maintenance of wakefulness test (MWT) to a modified behavioral test (OSLER) in the evaluation of daytime sleepiness. J Sleep Res 13:407–411. https://doi.org/10.1111/j.1365-2869.2004.00417.x
52. Alakuijala A, Maasilta P, Bachour A (2014) The Oxford Sleep Resistance test (OSLER) and the Multiple Unprepared Reaction Time Test (MURT) detect vigilance modifications in sleep apnea patients. J Clin Sleep Med 10:1075–1082. https://doi.org/10.5664/jcsm.4104
53. Loh S, Lamond N, Dorrian J, Roach G, Dawson D (2004) The validity of psychomotor vigilance tasks of less than 10-minute duration. Behav Res Methods Instrum Comput 36:339–346
54. Basner M, Dinges DF (2011) Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. Sleep 34:581–591
55. Dinges DF, Pack F, Williams K, Gillen KA, Powell JW, Ott GE, Aptowicz C, Pack AI (1997) Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. Sleep 20:267–277
56. Sforza E, Haba-Rubio J, De Bilbao F et al (2004) Performance vigilance task and sleepiness in patients with sleep-disordered breathing. Eur Respir J 24:279–285
57. Li Y, Vgontzas A, Kritikou I, Fernandez-Mendoza J, Basta M, Pejovic S, Gaines J, Bixler EO (2017) Psychomotor vigilance test and its association with daytime sleepiness and inflammation in sleep apnea: clinical implications. J Clin Sleep Med 13:1049–1056. https://doi.org/10.5664/jcsm.6720
58. Fronczek R, Middelkoop HAM, van Dijk JG, Lammers GJ (2006) Focusing on vigilance instead of sleepiness in the assessment of narcolepsy: high sensitivity of the Sustained Attention to Response Task (SART). Sleep 29:187–191
59. van der Heide A, van Schie MKM, Lammers GJ, Dauvilliers Y, Arnulf I, Mayer G, Bassetti CL, Ding CL, Lehert P, van Dijk JG (2015) Comparing treatment effect measurements in narcolepsy: the sustained attention to response task, Epworth sleepiness scale and maintenance of wakefulness test. Sleep 38:1051–1058. https://doi.org/10.5665/sleep.4810
60. Van Schie MKM, Thijs RD, Fronczek R et al (2012) Sustained attention to response task (SART) shows impaired vigilance in a spectrum of disorders of excessive daytime sleepiness. J Sleep Res 21:390–395. https://doi.org/10.1111/j.1365-2869.2011.00979.x
61. Schreier DR, Banks C, Mathis J (2018) Driving simulators in the clinical assessment of fitness to drive in sleepy individuals: a systematic review. Sleep Med Rev 38:86–100. https://doi.org/10.1016/j.smrv.2017.04.004
62. Philip P, Sagaspe P, Taillard J, Valtat C, Moore N, Åkerstedt T, Charles A, Bioulac B (2005) Fatigue, sleepiness, and performance in simulated versus real driving conditions. Sleep 28:1511–1516
63. Banks S, Catcheside P, Lack LC, Grunstein RR, McEvoy RD (2005) The maintenance of wakefulness test and driving simulator performance. Sleep 28:1381–1385
64. Sagaspe P, Taillard J, Chaumet G, Guilleminault C, Coste O, Moore N, Bioulac B, Philip P (2007) Maintenance of wakefulness test as a predictor of driving performance in patients with untreated obstructive sleep apnea. Sleep 30:327–330
65. Philip P, Sagaspe P, Taillard J, Chaumet G, Bayon V, Coste O, Bioulac B, Guilleminault C (2008) Maintenance of Wakefulness Test, obstructive sleep apnea syndrome, and driving risk. Ann Neurol 64:410–416. https://doi.org/10.1002/ana.21448
66. Pizza F, Contardi S, Mondini S, Trentin L, Cirignotta F (2009) Daytime sleepiness and driving performance in patients with obstructive sleep apnea: comparison of the MSLT, the MWT, and a simulated driving task. Sleep 32:382–391
67. Philip P, Chaufton C, Taillard J, Sagaspe P, Léger D, Raimondi M, Vakulin A, Capelli A (2013) Maintenance of Wakefulness Test scores and driving performance in sleep disorder patients and controls. Int J Psychophysiol 89:195–202. https://doi.org/10.1016/j.ijpsycho.2013.05.013