Neuroendocrinology Letters Volume 44 No. 5 2023
ISSN: 0172-780X; ISSN-L: 0172-780X; Electronic/Online ISSN: 2354-4716
Web of Knowledge / Web of Science: Neuroendocrinol Lett
Pub Med / Medline: Neuro Endocrinol Lett

REVIEW ARTICLE

Pupillography in contemporary sleep medicine – A narrative review.

Jan Mizera[1], Milan Sova[1,2], Samuel Genzor[1,3], Tomas Krejci[4], Jaromir Vachutka[4], Jakub Vanek[5], Pavol Pobeha[6], Jan Prasko[5,7,8,9]

1 Department of Tuberculosis and Respiratory Diseases, University Hospital and Faculty of Medicine and Dentistry of Palacky University Olomouc, Czech Republic.
2 Clinic of pulmonary diseases and tuberculosis, The University Hospital and Faculty of Medicine and Dentistry of Masaryk University Brno, Czech Republic.
3 Centre for Digital Health, University Hospital and Faculty of Medicine and Dentistry of Palacky University Olomouc, Czech Republic.
4 Department of Medical Biophysics, Faculty of Medicine and Dentistry of Palacky University Olomouc, Czech Republic.
5 Department of Psychiatry, University Hospital and Faculty of Medicine and Dentistry of Palacky University Olomouc, Czech Republic.
6 Department of Respiratory Diseases and Tuberculosis, University Hospital and Faculty of Medicine and Dentistry, Pavol Jozef Safarik University, Kosice, Slovak Republic.
7 Jessenia Inc. Rehabilitation Hospital Beroun, Akeso Holding, Czech Republic.
8 Department of Psychological Sciences, Faculty of Social Sciences and Health Care, Constantine the Philosopher University in Nitra, Slovak Republic.
9 Department of Psychotherapy, Institute for Postgraduate Training in Health Care, Prague, Czech Republic.

Correspondence to: prof. Jan Prasko, PhD
Department of Psychiatry, Faculty of Medicine and Dentistry, Palacky University Olomouc, I. P. Pavlova 6, 77520 Olomouc, Czech Republic
TEL: +420 603 414 930, E-MAIL: praskojan@seznam.cz

Submitted: 2023-03-04 *Accepted: 2023-06-26* *Published online: 2023-06-26*

Key words: **Pupillography; sleepiness; sleep medicine; Pupillographic Sleepiness Test; PST**

Neuroendocrinol Lett 2023; **44**(5):297–308 **PMID:** 37524319 NEL440523R02

Abstract

Excessive daytime sleepiness (EDS) is a common symptom of sleep disorders such as narcolepsy, obstructive sleep apnea, and hypersomnia. The most common tools for assessing EDS are various specialized questionnaires such as Epworth Sleepiness Scale (ESS) and Stanford Sleepiness Scale (SSS). However, the scores obtained from self-rating questionnaires do not seem to measure physiological sleepiness but rather a more complex phenomenon of subjective sleepiness modulated by other factors such as motivation, expectation, and capability of self-perception. The golden standard for measuring physiological sleepiness and assessing EDS is the Multiple Sleep Latency Test (MSLT). However, MSLT is very time consuming and requires trained personnel and expensive equipment. Different method modifications are employed in various medical and industrial fields for different purposes. The infrared pupillography in darkness has the potential to measure objective physiological sleepiness, especially the Pupillographic Sleepiness Test (PST), which is the method of choice for pupillographic measurement of daytime sleepiness. The method has also been employed in several specific sleep disorders, outlining possible future usage. This narrative review summarizes the current state of knowledge on the relevance and usefulness of pupillography in sleep medicine.

INTRODUCTION

Traditionally, sleepiness has been defined as a physiological inability to stay awake (Axelsson *et al.* 2020). Pivik (1991) defines sleepiness as a variable, relatively comprehensive and actively regulated reoccurring process influenced by arousal levels and sleep drive. Attention and sleepiness are complex phenomena. A model of attention by Posner & Rafal (1987) hypothesizes five levels of attention: tonic central nervous activation (basic level), phasic central nervous activation, selective attention (selection of relevant stimuli), divided attention (rapid, automatic, controlled processing of information) and vigilance (nonspecific readiness or capacity to keep a high or higher level of attention over a prolonged period of time, in a situation where stimuli requiring a response are rare and occur by chance).3 The latter three levels of attention are under conscious control.

Prevalence of some degree of daytime sleepiness is estimated to be around 4% to 20%, with 5% of the adult population suffering from excessive daytime sleepiness (EDS) (Carskadon 1993; Sauter *et al.* 2007).

EDS is defined as the inability to stay awake and alert during the major waking episodes of the day, with sleep occurring unintentionally or at inappropriate times almost daily for at least three months (Kraemer *et al.* 2000). Some sleep disorders are known to be accompanied by excessive daytime sleepiness (narcolepsy, obstructive sleep apnea syndrome, etc.) (Kraemer *et al.* 2000; Sauter *et al.* 2007).

Daytime sleepiness increases the risk of driving accidents and work-related incidents and is linked to worse overall health and poor occupational performance (Horne & Reyner 1995; Philip *et al.* 2001; Lockley *et al.* 2004; Barger *et al.* 2006). Assessment of daytime sleepiness plays a crucial role in sleep research, sleep medicine, and travel and work safety (Regen *et al.* 2013).

Methods of assessing sleepiness

There are four ways to measure daytime sleepiness: behavioural observation, laboratory test performance, introspection (subjective sleepiness, also termed manifest sleepiness), and physiological parameters (Multiple Sleep Latency Test) (Carskadon 1993). The subjective self-rating questionnaires and the physiology-based methods are mostly used in clinical practice.

Subjective self-rating scales

The first attempts to assess sleepiness consisted of free-form self-reports (Mitler & Miller 1996). Later, standardized questionnaires were developed. One of the most widely used questionnaires to evaluate self-perceived sleepiness is the Epworth Sleepiness Scale (ESS), in which the subject is asked to rate on a 0 to 3 scale (never, slight, moderate and high) the chance they would fall asleep in 8 described situations (Johns 1991). The ESS can statistically differentiate between normal subjects and sleep-disordered patients with EDS. ESS correlates with EEG-derived sleep latency and reflects on the improvements associated with the therapy in OSA patients (Johns 1992).

Another frequently used self-rating scale is the Stanford Sleepiness Scale (SSS) (Hoddes *et al.* 1972). In SSS, the subject chooses between 7 statements describing their level of sleepiness on a Likert-type scale ranging from 1 – "feeling active and vital, alert, wide awake" to 7 – "almost in reverie, sleep onset soon lost the struggle to remain awake". SSS can be easily repeated many times daily, correlates with standard performance measures, and reflects the effects of sleep loss (Mitler & Miller 1996).

However, most subjects cannot objectively evaluate physiological sleepiness (Carskadon *et al.* 1986). Some studies described glaring discordance between SSS scores and objective signs of sleepiness, such as snoring and eyelid closure in OSA patients. This discordance may be attributed to the loss of an objective frame for normal alertness, self-denial, or the fact that subjective and behavioural sleepiness indicators reflect different things.

Other options include the Visual Analog Scale (VAS) or different, more sporadically used scales, such as the Inappropriate Sleep Composite Score (ISCS) (Hobson *et al.* 2002). Some disease-specific questionnaires also contain items evaluating daytime sleepiness. One such questionnaire is the Berlin questionnaire, a useful tool for screening the general population for the presence of OSA, with a high negative predictive value. Four of the ten questionnaire items screen for daytime sleepiness symptoms (Thurtell *et al.* 2011).

Multiple Sleep Latency Test and Maintenance of Wakefulness Test

The Multiple Sleep Latency Test (MSLT), developed by sleep medicine pioneers Carskadon *et al.* (1986), is the gold standard for objectively measuring physiologic sleepiness. In its original version, the latency to the first occurrence of a 30-second epoch of stage 1 sleep on electroencephalogram (EEG) is measured in 4 to 5 20-minute nap opportunities under defined sleep-promoting conditions. Mean latency to sleep stage 1 shorter than 5 min is considered pathological, and a latency of 10 minutes is considered normal. Some authors, depending on the design of their studies, used various modifications of MSLT with a varying number (such as the Two-Nap Sleepiness Test) and duration (e.g., MSLT15) of nap opportunities or measuring latencies to deeper sleep stages and considered shorter episodes of sleep patterns of EEG as significant to account for micro-sleep (Kraemer *et al.* 2000; Suzuki *et al.* 2000). However, OSA is not a standard diagnosis for using MSLT, and this test is mainly validated for narcolepsy and idiopathic hypersomnia.

The Maintenance of Wakefulness Test (MWT) is a modification of MSLT, where the subject sits and is

asked to try to stay awake as long as possible in a series of 4 to 5 20–40-minute nap opportunities (Mitler *et al.* 1982). This test seems more appropriate than the MSLT for OSA drivers to assess the ability to stay awake in certain situations (not falling asleep while driving).

MSLT and MWT require expertise, are time and resource-consuming, and take up to 22 hours to perform in a sleep lab in their standard unmodified form (Merritt *et al.* 2004).

History of pupillography

In the late 1940s, Lowenstein & Lowenfeld (1958) pioneered using infrared (IR) pupillography to record and evaluate pupil size and behaviour under various conditions. First, the measurements were performed manually on a film. An important step forward was made with the development of a photoelectric pupillograph. The method was revitalized in the 1970s with the advent of the digital era and the possibility of continuous videotaping, leading to an almost exponential increase in researchers and publications on pupillography (Kelbsch *et al.* 2019). The surge of studies with various experimental designs resulted in the need for methodological standardization. The preparations for this important step began at the 32nd International Pupil Colloquium 2017 in Morges, Switzerland, and resulted in the publishing of the "Standards in pupillography" in 2019 (Kelbsch *et al.* 2019).

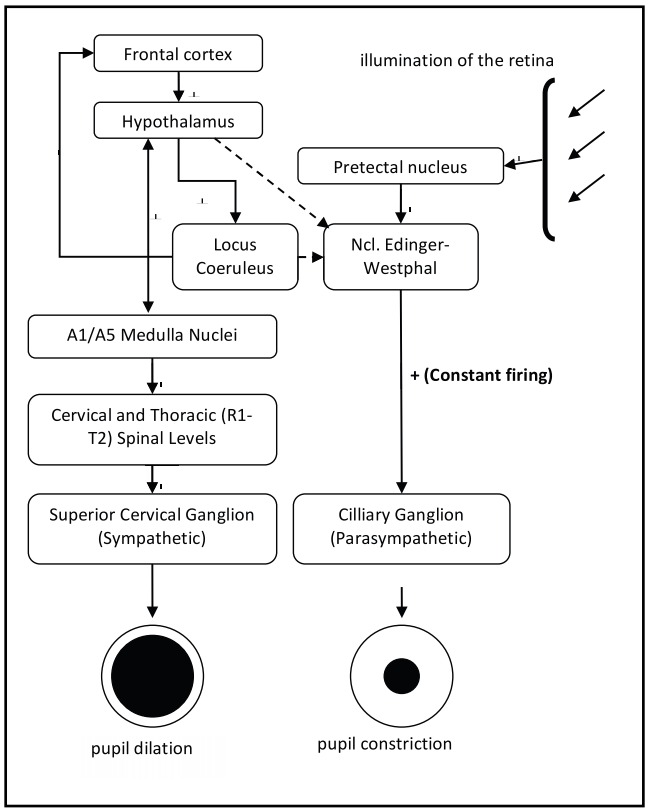

Fig. 1. Pupil size regulation (based on Szabadi & Bradshaw 1996)

Technology

In modern-day pupillography, the video recording of the pupil is usually obtained with an infrared camera in darkness, with or without the use of various stimuli employed to elicit a pupillary response. Both systems for recording pupil diameter from a distance and integrated wearable head-mounted systems are available. The standards in pupillography specify the minimum recommended spatial and temporal resolution for a pupillographic system, as well as the recommended stimulus characteristics and test conditions for the various use cases of pupillography (Kelbsch *et al.* 2019). Usually, the pupil is recorded at 25 to 60 Hz, with a spatial resolution of 0.01 to 0.05mm, with a minimum 8-bit (preferably a 10- or 12-bit) analogue-to-digital sampling. The image processing software detects the pupil on each frame and calculates its diameter or area (Kelbsch *et al.* 2019). Artefacts caused by blinking, rapid eye movements, etc., are removed automatically or manually (Ludtke *et al.* 1998). Nowak *et al.* (2014) published a design paper detailing their binocular experimental pupillography recording with a time resolution of 75Hz and spatial resolution better than 0.02 mm (Nowak *et al.* 2014).

Basic physiology of the pupil

The pupil's diameter is determined by a mutually opposite action of the two smooth muscles of the iris – musculus constrictor pupillae and musculus dilatator pupillae.

The constrictor muscle receives constant firing from the nucleus Edinger-Westfall (ncl. EW), leading to pupil constriction. Ncl. EW also receives projections from the retina via the pretectal nucleus (Szabadi & Bradshaw 1996). The latter pathway is responsible for the pupillary light reflex and near response (Lowenstein *et al.* 1963).

The dilator muscle is innervated by the sympathetic branch of the autonomic nervous system via the medullary nuclei A1/A5, Cervical and Thoracic spinal levels (T1-2) and the Superior Cervical Ganglion. The activity of medullary nuclei A1/A5 and Locus Coeruleus (LC), the major source of the neurotransmitter norepinephrine, is modulated by cortical activity predominantly in the frontal lobe, which is down-propagated to the hypothalamus and the reticular activating system (RAS), where it is integrated with the feedback loop originating from the medullary nuclei and LC (Figure 1). Both the hypothalamus and LC have an inhibitory pathway to ncl. EW, leading to the inhibition of parasympathetic activation of the constrictor pupillae (Szabadi & Bradshaw 1996).

The pupil's diameter is thus a result of a dynamic equilibrium between the action of the two opposing muscles of the iris, which in turn reflects sympathetic and parasympathetic activity (Wilhelm *et al.* 2001). Bumke states that pupil diameter is affected by every

mental effort, exertion of attention, mental image, distress, and every sensory stimulus (Bumke 1911).

The article aims to answer following questions:
1) Which pupillographic techniques are possibly applicable in clinical praxis?
2) How reliable is pupillography in the measurement of sleepiness?
3) Do pupillography measures correlate with the results of other methods (especially multiple sleep latency test)?
4) In which sleep disorders may pupillography be used to measure daytime sleepiness?

METHOD

The authors conducted a database search via PubMed (PM) and Web of Science (WoS) with the following restrictions: articles written in English and date of publication unrestricted. The primary search term was "pupillography" OR "pupillometry". This search retrieved 1241 results and 1720 results in PM and WoS, respectively. A secondary search term was applied ("fatigue" OR "Obstructive Sleep Apnea" OR "sleep" OR "sleepiness"), which reduced the number of potentially relevant items to 141. The following inclusion criteria were considered: (1) published in peer-reviewed journals, (2) human studies, and (3) related topics. The exclusion criteria were: (1) pediatric population, (2) commentaries (3) abstracts from conferences. This resulted in 41 items being nominated for review. After obtaining and studying the full texts, 17 papers were excluded for not matching the scope of this review (e.g., measuring task-related fatigue, study scenario unrelated to sleepiness and sleep medicine). References of the remaining 24 studies were screened for relevance based on the same inclusion criteria as the primary articles leading to the inclusion of 32 further studies. Four additional papers written in German were added, violating the language criterion. This is justified by their high relevance to the subejct and their direct relationship (primary author) to other works written in English. A professional English translation was obtained for all four articles. A total of 60 papers were thus included in the review (Figure 2).

RESULTS

Pupillographic techniques and their clinical applications

There are several pupillographic techniques with different clinical applications. Pupillography may be used to record and characterize the pupillary light reflex (PLR), a change in pupil diameter elicited by a defined light stimulus. This technique was successfully employed in ophthalmology: Pupillographic PLR can accurately measure afferent pupillary defects (automated swinging flashlight test) (Wilhelm *et al.* 2001). Pupil campimetry attempted to measure visual field defects, as the pupillomotor field shows a similar profile to the visual field (Kardon *et al.*1991). Pupil dilation lag has been useful in diagnosing Horner Syndrome (Pilley & Thompson 1975).

The measurement of Post-Illumination Pupil Response (PIPR), a sustained pupil constriction after light offset mediated by intrinsically photosensitive Retinal Ganglion Cells (ipRGCs), was employed in ageing research and the detection and monitoring of glaucoma, diabetic retinopathy, age-related macular degeneration and ischemic optic neuropathy.

In psychology and psychiatry, pupil diameter changes and alterations in PLR and PIPR were observed in reaction to cognitive effort and emotional stimuli (Kelbsch *et al.* 2019).

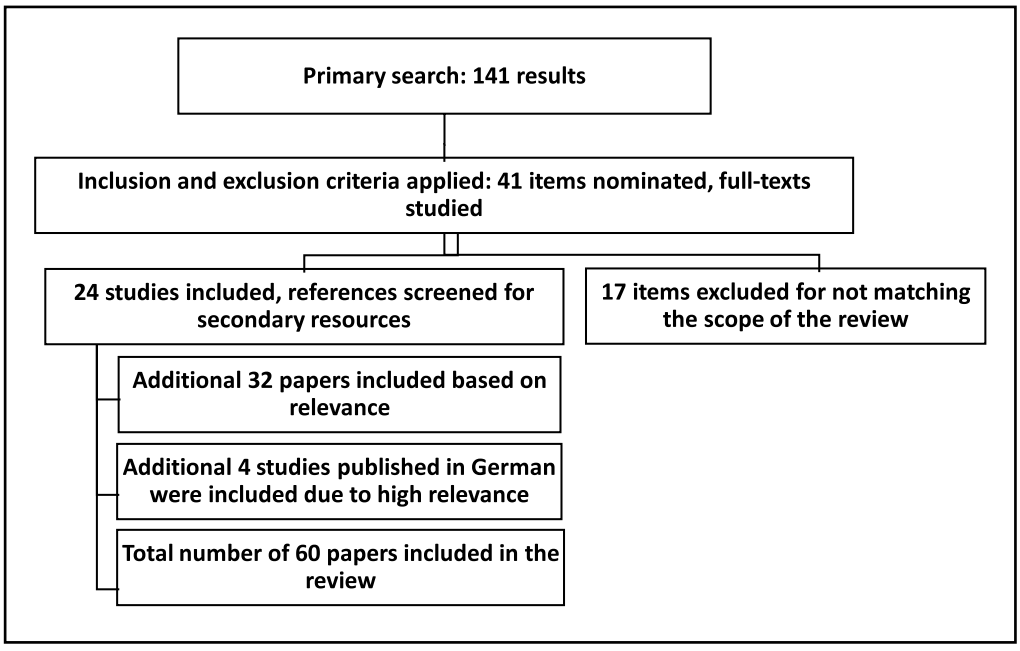

Fig. 2. Selection of papers for the systematic review

Colour pupillography is currently on the rise, promising new insights into retinal physiology, pupil circuitry, and further understanding of diseases such as glaucoma, retinitis pigmentosa, age-related macular degeneration, diabetes, or hereditary optic neuropathy (Kelbsch *et al.* 2019).

Nowak & Kasprzak (2008) examined a 12-second dark pupillography recording at 90 Hz (the authors called the protocol "fast pupillometry"), with the recording being analyzed by short-time Fourier Transform (STFT) with attention being drawn not only to the slow pupillary oscillations < 0.8 Hz but also to their harmonics > 1Hz up to 20 Hz. The authors suggest power instability factor (PIF) and frequency instability factor (FIF) as novel parameters describing spontaneous pupillary oscillations.

Pupillographic measurement of sleepiness

In the 1950s, Lowenstein & Lowenfeld (1951; 1952a; 1952b; 1958) studied the relationship between pupil behaviour and alertness. They described the typical pupil behaviour in darkness, with a stable and large pupil in alert subjects and oscillating pupil diameter with a tendency to smaller overall diameter in sleepy subjects. The slow oscillations happen within 4 to 40 seconds, with an amplitude of up to 0.5 mm in young, healthy, non-sleep-deprived subjects. These slow pupillary oscillations are exaggerated in sleep-deprived subjects (Lowenstein *et al.* 1963). Lowenstein named these slow pupillary oscillations in sleepy subjects "fatigue waves" (nowadays, the term "sleepiness waves" is preferred, as there is a distinction being drawn between sleepiness and fatigue) (Lowenstein *et al.* 1963; Kelbsch *et al.* 2019).

The interest in the use of pupillography in sleep medicine was rather limited between 1963-1993. This was mainly due to the lack of technology that could effectively deal with the many artefacts in the recordings, especially in drowsy subjects (Wilhelm *et al.* 1998a). The interest in pupillographic assessment of sleepiness was reconsidered during the 20th Pupil Colloquium in Iowa City in 1993 (Wilhelm *et al.* 1998a). In the 1990s, the collective of authors around a mathematician and physicist Dr Holger Lüdtke, together with professors Helmut and Barbara Wilhelm (a neuro-ophthalmologist and neurologist, respectively), developed the Pupillographic Sleepiness Test (PST), which employed a novel approach to the processing of the recordings of slow pupillary oscillations (Wilhelm *et al.* 1998b). The evidence supporting the use of PST accumulated over the years, leading to the PST being the recommended pupillographic technique for measuring physiological sleepiness in the Standards in Pupillography published in 2019 (Kelbsch *et al.* 2019).

Alertness Level Testing (ALT) - a variant of a dark pupillography by McLaren *et al.* (1992; 1995) using different outcome parameters to describe the dynamics of spontaneous pupillary oscillations. The test seems to have been abandoned due to the success of PST.

In neurology and ophthalmology, pupillary light reflex (PLR) measures afferent and efferent visual pathways. Colour pupillography allows for selective testing of different photosensitive receptors of the retina (Kelbsch *et al.* 2019). However, several studies also explored the utility of PLR-derived metrics as potential measures of sleepiness with conflicting results. Lowenstein *et al.* (1963) reported a decreased pupil response to light stimulation in an extremely sleepy subject. According to Wilhelm *et al.* (2001), several authors reported inconsistent results regarding the alterations of PLR in clinically sleepy subjects between 1951 and 1994. Ranzijin & lack (1997) concluded that alterations in PLR metric due to sleepiness could not be observed under most circumstances.

Non-pupillometric eye-based features such as eye blinks, eyelid droopiness, and eye movements and their usefulness for measuring task-related fatigue are discussed in more detail in a recently published systematic review by Bafna & Hansen (2021).

Pupillographic Sleepiness Test (PST)

The PST works on infrared pupillographic recording in near-complete darkness. Nowadays, fully integrated head-mounted systems for recording and evaluating PST are commercially available. The minimum standards recommend abstinence from caffeine for 8-10 hours and abstinence from nicotine products one-hour before measurement (Kelbsch *et al.* 2010). Medication (including topical medication such as eye drops) possibly affecting pupil size should be avoided or at least documented. Immediately before the measurement, the subject should rest in a sedentary position for 10 minutes to eliminate the effect of physical activity on pupil diameter.

The measuring device is then set up; this usually includes the subject putting on a pair of goggles. The goggles may consist of IR filters permissive only to the IR wavelengths, with the IR cameras being placed at a specified distance from the subject, whose head must be supported by a chin and forehead rest. The other type of goggles is fully integrated, with one or more IR cameras and IR light sources being a part of the head-mounted device. In either instance, the standards in pupillometry recommend that the measuring take place in a room as dark as possible (3 cd/m^3 or less) as the goggles may not fit every face shape ideally and thus not be completely lightproof (Kelbsch *et al.* 2019). It is further important to rule out acoustic stimuli, as these influence pupil behaviour. This may be achieved in a silent room with sound-dampening or noise-cancelling headphones (Kelbsch *et al.* 2019).

This preparatory phase is followed by a short period of dark adaptation – the exact length varies by author, ranging from 90 seconds to 15 minutes (Reimann *et al.* 2009; Yamamoto *et al.* 2013). Wilhelm *et al.* (2001).

suggested that complete dark adaptation is unnecessary to measure sleepiness waves. Longer adaptation times of up to 30 minutes are required for pupillographic measurements, such as pupillary light response. After the dark adaptation period, the pupil is recorded for 11 minutes at 25 Hz, with a spatial resolution of 0.01 to 0.05 mm, depending on the apparatus. The subject fixes a dim red light as previously instructed during this time.

The pupil is automatically detected in each frame using the first Purkinje reflex (Wilhelm *et al.* 2009). The pupil margins are then electronically detected, and the horizontal diameter of the pupil is automatically measured in mm. Artefacts such as eye blinks and movements are automatically detected by an algorithm based on a threshold for maximum instant physiological change in pupil diameter. The missing frames are replaced by linear interpolation (Ludtke *et al.* 1998). The 11-minute recording is evaluated in 8 segments by 81.92 seconds (2048 frames). Mean Pupillary Unrest Index (PUI) is calculated for each of the eight segments and the entire recording. PUI is calculated as the mean pupillary diameter (PD) in 16 consecutive frames minus the mean PD in the following 16 consecutive frames (Ludtke *et al.* 1998) PUI is one of the three main outcomes of the PST test. Another parameter is the power of low frequencies <= 0.8 Hz obtained by applying the Fast Fourier Transform (FFT) to the time series of recorded pupillary diameters. The other output variables of the PST are baseline PD (mean PD in the first 81.92-second segment), PD (average PD over the entire testing period), and relative PUI (RPUI, which is calculated as PUI divided by the baseline diameter). The higher the PUI and the higher the power of low frequency bands, the higher the sleepiness in the subject.

The ability of PST to objectively and reliably measure daytime sleepiness has been verified in several sleep deprivation studies. Wilhelm *et al.* (1998b) conducted repeated PST measurements in 13 healthy subjects during all-night forced wakefulness. Their results showed progressive PD reduction and an increase in the metrics of slow papillary oscillations (PUI and power in the frequency band 0.1 - 0.8 Hz). The pupillographic sleepiness did not correlate significantly with the subjective sleepiness (SSS) measure, even though this also steadily increased during night waking. Wilhelm *et al.* (2001) also studied time-of-day variation in PST measures, which included a night of forced wakefulness during the 30-hour protocol (08:00 to 14:00 the following day). Again, 13 healthy young adults were enrolled, with PST and subjective assessment of sleepiness (SSS and VAS scales) repeated every 2 hours. The results showed a significant intra-individual time-of-day variation of PD, PUI, and power in frequencies <= 0.8 Hz. Maximum PUI (signifying maximum sleepiness) was measured at 09:00 and 23:00, with a statistically non-significant peak at 15:00. The subjective self-rating scales scored lowest in the morning, with a continuous increase throughout the experiment.

Compared to the values recorded during the day, PUI and the power of the low frequency spectrum during the night of forced wakefulness were significantly higher, and PD was significantly lower. Furthermore, Wilhelm *et al.* (2009) applied pupillography in the occupational setting to gauge daytime sleepiness in 34 neurology residents following a night of undisturbed sleep at home, compared to the morning following a night shift. After the night shift, the results showed significantly higher lnPUI (natural logarithm of PUI). The differences in lnPUI between night duty and day work correlated with the differences in SSS and VAS ($p = 0.02$). A similar study by Reimann *et al.* compared PST and Paced Auditory Serial Addition Test (PASAT) outcomes in 38 neurology residents with different work schedules (night shift, 24-hour on-call duty, sleep at home). Their results have confirmed significantly higher PUI following a night shift or an on-call duty compared to sleep at home. The difference in PUI between the types of night duty was insignificant. SSS scores correlated with the amount of sleep deprivation, significantly correlating with PUI in the total sample but not in the individual subgroups. The authors hypothesize that the differences between the groups may have been attenuated due to the chronic sleep deprivation in some of the subjects. PASAT scores did not significantly differ between the groups, suggesting that sleep deprivation during the night shift did not significantly influence the capacity to perform short-term cognitive tasks (Reimann *et al.* 2009). Regen *et al.* (2013) conducted a sleep deprivation study on 24 young adults. The subjects had to endure 40 hours of sustained wakefulness, with repeated PST, tympanic temperature measurements and self-evaluation on SSS and VAS scales. Waking EEG was recorded continuously throughout the experiment using a portable device. The results showed amplification of delta, theta and alpha-1 bands with growing sleep deprivation consistent with previous literature remarks. PUI exhibited a continuous increase with a growing sleep deficit, confirming the results of Wilhelm *et al.* (2001). The close association of PUI and distinct changes in waking EEG were novel findings. Surprisingly, PUI decreased in the afternoon of day two despite cumulating sleep debt, but it did not reach the levels found at the beginning of the experiment. This likely reflects the strong circadian modulation of PUI outlined in the previous time-of-day variation studies fighting the homeostatic sleep pressure caused by sleep deprivation. In this study, PUI correlated with both SSS and VAS.

As the PST and its main outcome parameter PUI proved to be able to measure daytime sleepiness, a set of normal values was required. The normal values for lnPUI were provided by Wilhelm *et al.* (2001b), who examined PST during the morning hours in a cohort of 191 men and 158 women between 20 and 60 years

of age. Eggert *et al.* (2012) administered PST in the morning and the afternoon to a cohort of 239 healthy subjects aged 20 to 79. Unreliable PST data were more frequent in the subjects older than 60, possibly due to dry eyes, inadequate testing conditions or technical limitations.

The short-term reproducibility and variability of PST were examined by Wilhelm *et al.* (2015) in a group of 13 healthy young adults. Each subject underwent PST measurement and subjectively rated their sleepiness on SSS and VAS scales at 09:00, 11:00 and 13:00 on three consecutive days. The inter-individual differences of lnPUI were highly significant ($p < 0.0001$). The results confirmed the time-of-day effect on PUI ($p < 0.001$) found in the previously published time-of-day variation studies. The Intra-class Correlation Coefficient (ICC) of 73.1% and R^2 of 75.44% indicate a good short-term reproducibility of lnPUI measurements. SSS and VAS also displayed significant inter-individual and time-of-day variability, but the ICC and R^2 were significantly lower (38,83% and 52,04%, respectively). The intra- and inter-individual variability of lnPUI did not alter substantially when recorded at different times on different days. The authors draw a comparison with the repeatability of MSLT as reported by Zwyghuizen-Doorenbos *et al.* (1988), who used a 4-nap MSLT (with nap opportunities at 10:00, 12:00, 14:00 and 16:00) measured twice in each of the 14 healthy subjects, with a span of 4 to 14 months between the measurements. These authors reported very high reliability and consistency of the MLST test consisting of 3 or more naps ($r = 0.97$, $p < 0.001$). Wilhelm *et al.* (2015) argue that considering only the two-morning measurements out of four in the Zwyghuizen-Doorenbos *et al.* (1988) MSLT study brings the hypothetical repeatability of MSLT down to 0.64, which is similar to that reported for lnPUI measured twice in the morning. The long-term repeatability reported by Lüdtke *et al.* (2000), who measured morning PST twice with a gap of three months in 38 healthy men, is similar, with a correlation coefficient of 0.64 and a coefficient of repeatability of 0.76 for lnPUI ($p < 0.0001$).

Pupillographic Sleepiness Test is the protocol recommended by the Standards in Pupillography for measuring daytime sleepiness. The standards also encourage further experimental studies (Kelbsch *et al.* 2019). Even with the acceptance of PST as a standard for pupillographic measurement of sleepiness, experiments continue with different pupillographic techniques for measuring sleepiness.

Pupillography versus Multiple Sleep Latency Test

Before the development of PST, Newman & Broughton (1991) aimed to quantify excessive daytime sleepiness (EDS) in 10 narcolepsy-cataplexy patients and ten age- and sex-matched controls with the use of pupillography and MSLT. The authors employed a robust 2-day test protocol. The subjects filled in the SSS questionnaire every 30 minutes on both days. On the first day, the subjects went through 24-hour electrophysiological monitoring at home using a portable device. On the second day, the subjects were introduced to the sleep laboratory, where they underwent a 5-nap MSLT, with each nap opportunity preceded by pupillographic measurement. The authors employed a custom and rather complex pupillographic protocol spanning over 18 minutes and including periods of relaxation, ten brief and precisely defined visual stimuli to elicit Pupil Light Response (PLR), and also a series of 11 precisely defined auditory stimuli with the inclusion of an extraordinary auditory stimulus to elicit Orientational Reflex (OR). With this robust protocol, authors managed to acquire a variety of parameters for statistical analysis, including SSS scores, sleep latencies, EEG sleep stages and their latencies, various metrics characterizing pupil diameter and its oscillations, pupil light reflex magnitude and dynamics, and orientational reflex parameters. The authors then proceeded with data transformations to correct for the previous night's sleep effect, followed by thorough statistical scrutiny (Newman & Broughton 1991). Out of all the pupillographic variables obtained, only the number of pupillary oscillations proved to be a promising measure of EDS. The frequency of spontaneous oscillations correlated with MSLT latencies in the controls but not in narcoleptics. No significant correlations were found between spontaneous oscillations and SSS scores. The authors suggest that the real utility of pupillary oscillations for measuring daytime sleepiness may be increased by employing a sleep-promoting test protocol rather than the attention-demanding protocol used in the study (Newman & Broughton 1991).

Kraemer *et al.* (2000) investigated time-of-day variations of several physiological parameters (modified MSLT/MWT derived sleep latencies, pupillographic PD and its coefficient of variation - COVPD), subjective sleepiness scales (SSS, VAS), and a battery of performance tests (Number Facility Test, Visualization Test, Critical Flicker Fusion Test, Reaction time testing). Despite the limited number of healthy, non-sleep-deprived subjects (n=12), the authors managed to generate a fairly robust dataset due to the longitudinal design of the study, with physiological and performance tests repeated every 2 hours between 7-23:00, and self-rating scales filled in hourly, with one MSLT or MWT (randomized) session at the end of the test day. Every subject attended 2 test days with a gap of several weeks. The subjects also underwent polysomnography the night before each test day to assess the previous night's effects on the physiological measures. The study results have shown pupillometric variables and sleep latencies to have a similar circadian rhythm suggesting maximum alertness at 07:00, with a sharp, steady decrease till 09:00, followed by a relatively constant slow reduction to a minimum at around 17:00-19:00, with subsequent increase in

alertness until its second peak at around 21:00. The pupillographic variables correlated with sleep latencies and subjective self-ratings, but not with the parameters of performance.

A year later, Danker-Hopfe *et al.* (2001) performed a similar study on 12 healthy subjects, with repeated measurements of MSLT15 (modified MSLT) and pupillography, using PD, COVPD, and this time also PUI and the square root of power within the frequency band 0.1–0.8 Hz as outcome variables. COVPD was calculated from mean PDs for nine consecutive 1-minute recordings. The testing protocol was similar to the previous study, with the test battery being performed every 2 hours from 7:00 to 23:00. Subjects also filled in SSS hourly. Both the MSLT15 latencies and pupillary oscillations as measured by PUI, COVPD and power ≤ 0.8 Hz followed a similar time-of-day variation with a peak after getting up in the morning and another peak between 19 and 21:00. A post-lunch dip could not be observed in the present study. These results show that MSLT15 and pupillography measure the same dimension of sleepiness. In contrast, SSS shows an almost opposite pattern of time-of-day variation, with decreasing subjective sleepiness during the morning hours and a minimum around noon followed by steadily increasing subjective sleepiness until the evening hours. Despite the small study sample, some correlation coefficients between MSLT latencies and pupillographic variables were statistically significant.

Prasad *et al.* (2011) studied the utility of PUI as a measure of daytime sleepiness in a group of 20 narcoleptic patients (both with and without cataplexy) and 56 healthy controls. After night polysomnography, the test protocol was carried out, featuring 4 MSLT nap opportunities followed by PST at 09:00, 11:00, 13:00 and 15:00. Visual Analog Scale was used to gauge subjective sleepiness before each test session, and ESS was filled in before one randomly selected session. The subjects also performed a 10-minute Psychomotor Vigilance Test (PVT) at 9:30 and 13:30 with median reaction time (MRT) and the frequency of lapses (RT > 500ms) as the outcome variables. When performing statistical analysis on the whole dataset (76 subjects), all subjective and objective measures of sleepiness were significantly correlated with one another but not with the performance measurements. However, considering the two groups separately, only PUI to SL, PUI to VAS and ESS to VAS were significantly correlated ($p = 0.04$, $p = 0.04$ and $p = 0.01$, respectively) in the control group. At the same time, none of the objective or subjective measures of sleepiness displayed a significant correlation in the narcoleptics group. In the narcoleptics group, VAS correlated significantly with PVT MRT ($p = 0.02$). The authors concluded that although PUI distinguished pathological sleepiness in narcoleptics and is an accurate indicator of sleep propensity in controls, it cannot substitute for MSLT in diagnosing narcolepsy (Prasad *et al.* 2011). The possible role of PUI as an indicator of therapeutic effect in narcoleptics should be examined in future studies.

McLaren *et al.* (2002) published a study in 49 patients scheduled for a diagnostic MSLT and 33 healthy controls, combining their ALT protocol (15-minute dark pupillography protocol recorded at 30Hz) with a combination of the original ALT outcome variables and an adapted version of the PST outcomes PUI and power within frequency range 0.1 – 0.8 Hz. The authors did not find a significant correlation between pupil diameter and the parameters describing its low-frequency oscillations (McLaren *et al.* 2002). Merritt *et al.* (2004) attribute this to the faults in study design (inclusion of elderly subjects, the reduced timespan of 2 hours between the four nap opportunities, the possibility of experimental fatigue, and the absence of MSLT measurement in the control group).

More recently, Yamamoto *et al.* (2013) performed a study comparing PST outcome variables with a two-nap sleepiness test (TNST, modified version of MSLT) derived sleep latency (SL) in 45 sleep disordered patients (with diagnoses including sleep apnea syndrome, hypersomnia, narcolepsy, idiopathic hypersomnia, behaviorally induced insufficient sleep syndrome, and depression). The subjects also filled in ESS questionnaires before each testing session performed at 10:00 and 12:00 within one testing day. Mean PUI, RPUI and SL values obtained in the two testing sessions were used for data analysis. The results show that PUI and RPUI correlated significantly with SL, but the correlation of PUI was much stronger ($p < 0.01$ versus $p < 0.05$, respectively). ESS score did not correlate to PUI, RPUI, or SL.

Pupillography in specific sleep disorders

Before PST was developed, Lichtenstein *et al.* (1992; 1994) performed two experimental pupillographic studies on 34 and 30 patients with insomnia, and 29 and 30 self-reported insomnia patients, respectively. The protocol for both studies was identical: The subjects underwent four testing sessions from early morning till bedtime. In each test session, they filled in a Sleep Questionnaire (SQ), SSS, and Events Affecting Sleep (EAS) questionnaires, followed by pupillographic measurements (Lichstein & Fischer 1985). After a 3-minute dark adaptation, a 10-minute pupillographic recording was obtained. The outcome variables used for the statistical analysis were: Pupil diameter (PD) measured vertically (sampled every 15 seconds and averaged per minute); Oscillations (OS) expressed as a number of small oscillations (0.5–1.5mm) and large oscillations (> 1.5 mm) per minute; Eye blinks (EY) defined as rapid eye closures with eyelid covering 50% or more of the pupil area, followed by eye reopening within 2 seconds. Pupil diameter (PD) separated insomniacs in every minute of every recording by an average margin of about 0.3 mm in the first study and 0.5 mm in the second study. The PD was consistently smaller in

insomnia patients. However, collapsing across minutes and sessions revealed that the two plotted normal distributions for insomnia and non-insomnia groups were strikingly similar, with a 95.4% overlap in the second study. Both small and large oscillations did not significantly correlate with the patient group. As a secondary outcome, SSS scores could only statistically significantly discriminate between the two subject groups at midday and not in the other three measurements. The authors acknowledge several serious pitfalls of their study, such as the age difference between the two groups and the fact that insomnia in the test subjects was not objectively verified by polysomnography, perhaps leading to the inclusion of subjects whose complaints of disturbed sleep may have been better described by the "insomnoid" model ("insomnoids" are individuals who manage to meet their biological sleep need despite complaints of disturbed sleep) (Lichstein & Fischer 1985). The authors also point out that even MSLT is not great at discriminating insomnia, as the test requires sleep behaviour in patients who have trouble falling asleep.

Wilhelm *et al.* (1998; 1999) performed two studies on patients suffering from obstructive sleep apnea (OSA), reporting significantly lower PUI after two nights and three months of nasal Continuous Positive Airway Pressure (nCPAP) therapy in 37 and 35 male OSA patients, respectively. PUI differed significantly pre- versus post-treatment, with a mean difference of 21% in the first study (Wilhelm *et al.* 1998). The results were similar in the second study (Wilhelm *et al.* 1999). PST measures of sleepiness did not significantly correlate to the apnea-hypopnea index (AHI), but PUI was significantly higher in patients with AHI > 55 in the first study (Wilhelm *et al.* 1998). These studies have shown the possible usefulness of PUI as a measure of treatment efficacy, especially in severe OSA patients, as those typically present with daytime sleepiness as one of the leading symptoms

DISCUSSION

This review summarises and discusses current knowledge about pupillography and its use in sleep medicine. According to the presented data, pupillography seems to be an easy, inexpensive and relatively reliable method for sleepiness evaluation. Unfortunately, the method is nowadays not widely used, and we are missing normal values for different population groups, races, genders, elderly (60+), etc. However, large studies focusing on the specific use of pupillography in sleep disorders are still missing. The following part of the discussion will be divided into paragraphs answering the study questions.

1) Which pupillographic techniques are possibly applicable in clinical praxis?

According to the available literature, several pupillographic techniques have different clinical applications. Methods evaluating PLR measure afferent pupillary defects and may be used to diagnose Horner Syndrome. On the other side, the measurement PIPR (a sustained pupil constriction after light offset) have applications in ageing research and the detection and monitoring of glaucoma, diabetic retinopathy, age-related macular degeneration and ischemic optic neuropathy. Pupil diameter changes and alterations in PLR and PIPR may have possible clinical use also in psychology and psychiatry, as they were observed in reaction to cognitive effort and emotional stimuli (Kelbsch *et al.* 2019). For possible clinical use of sleepiness measurement, computer processing of the recorded signal (e.g., STFT) is needed. Spontaneous pupil oscillations seem to be the most important correlate of excessive sleepiness (Kraemer *et al.* 2000; Nowak & Kasprzak 2008; Prasad *et al.* 2011).

2) How reliable is pupillography in the measurement of sleepiness?

The key to the reliability of any measurement are well defined normal values. In pupillography, we have several medium-sized studies with defined protocols which we can use as reference values. Each study restricted caffeine use (at least 8-10 hours before the study) (Kelbsch *et al.* 2019). Certain medications (including eye drops) possibly affecting pupil size should be avoided. The room should be dark and completely quiet, and the device-patient interface should be pleasant to the subject (Kelbsch *et al.* 2019). In the preparatory phase, patient adaptation should be given some time (differs in studies from 90 seconds to 15 minutes) (Yamamoto *et al.* 2013; Reimann *et al.* 2009). Several studies were identified aiming to set the normal values for lnPUI. Firstly, Wilhelm *et al.* (2001a) examined PST during the morning hours in a cohort of 191 men and 158 women between 20 and 60. Eggert *et al.* (2012) administered PST in the morning and the afternoon to a cohort of 239 healthy subjects aged 20 to 79. Unreliable PST data were more frequent in the subjects older than 60, possibly due to dry eyes, inadequate testing conditions or technical limitations.

The ability of pupillography to objectively and reliably measure daytime sleepiness has been verified in several sleep deprivation studies. Wilhelm *et al.* (1998b) conducted repeated measurements in 13 healthy subjects during forced wakefulness for the whole night. The results showed progressive PD reduction and an increase in the metrics of slow papillary oscillations. Interestingly, no correlation was found between the pupillographic and subjective measures of sleepiness. Reimann *et al.* (2009) compared PST and Paced Auditory Serial Addition Test (PASAT) outcomes in 38 neurology residents with different work schedules (night shift, 24-hour on-call duty, sleep at home). The PUI was consistently higher in more sleep-deprived individuals. However, the subjective sleepiness score

correlated in this study with both – PUI and the length of the sleep deprivation.

The PUI value correlates with the day's time, as Wilhelm *et al.* (2015) show. Each of the 13 subjects underwent PST measurement and subjectively rated their sleepiness on SSS and VAS scales at 09:00, 11:00 and 13:00 on three consecutive days. The Intra-class Correlation Coefficient (ICC) of 73.1% and R^2 of 75.44% are considered good short-term reproducibility of lnPUI measurements. Similarly, the daytime variation of PUI was found in other studies (Zwyghuizen-Doorenbos *et al.* 1988; Regen *et al.* 2013).

3) Do pupillography measures correlate with the results of other methods (especially multiple sleep latency test)?

The study of Newman *et al.* (1991) correlated results of pupillography measurements, SSS and MSLT in 10 narcolepsy-cataplexy individuals and healthy controls. Interestingly, spontaneous pupil oscillations correlated closely with the results of MSLT only in healthy controls. Moreover, no correlations were found between spontaneous oscillations and the SSS score. The two studies of Kreamer *et al.* (2000) and Danker-Hopfe *et al.* (2001) demonstrated correlations between pupillography measures and MSLT/MWT (or modified MSLT protocol – MSLT15) to evaluate time-of-day variations. The study concluded that the pupillographic variables correlated closely with sleep latencies and subjective self-ratings but not with the performance parameters. Surprisingly, subjective sleepiness showed a different pattern of time-of-day variation, with the lowest values in the morning (when MSLT and PUI reached maximum). The study of Prasad *et al.* (2011) tested the PUI's utility in evaluating sleepiness in healthy individuals and narcoleptic patients. The results of MSLT, subjective evaluation of sleep (ESS and VAS) and PUI were closely correlated. However, the results did not correlate with the performance measurements (PVT).

4) In which sleep disorders may pupillography be used to measure daytime sleepiness?

Two studies by Lichtenstein *et al.* (1992; 1994) tested the utility of pupillography in the sleepiness measurement in insomniacs compared with healthy controls. The PD was consistently smaller in insomnia patients. However, there were no clear differences between the groups; only the SSS score at midday was significantly higher in the insomniacs. Unfortunately, both studies were very small, and the insomnia was diagnosed only by the patient's history, not verified by polysomnography.

Wilhelm *et al.* (1998; 1999) performed two studies on males suffering from obstructive sleep apnea (OSA), reporting significantly lower PUI after two nights and three months of CPAP therapy. PUI decreased consistently by 21% in both studies. PST measures did not correlate to the AHI, but PUI was significantly higher in patients with AHI > 55 in the first study (Wilhelm *et al.* 1998). Therefore, we suggest that there might be possible clinical use of pupillography measurement of daytime sleepiness in severe OSA patients.

Yamamoto *et al.* (2013) tested the utility of PST compared with TNST in patients with sleep disorders (including sleep apnea syndrome, hypersomnia, narcolepsy, idiopathic hypersomnia, behaviorally induced insufficient sleep syndrome, and depression). The study revealed that PUI and RPUI correlated significantly with SL, but the correlation of PUI was much stronger ($p < 0.01$ versus $p < 0.05$, respectively). ESS score did not correlate to PUI, RPUI, or SL.

Prasad *et al.* (2011), in their study of narcoleptic individuals, concluded that although PUI distinguished pathological sleepiness in narcoleptics and is also a very accurate indicator of sleep propensity in controls, it cannot fully substitute for MSLT in diagnosing narcolepsy. However, there might be a possible role of PUI as an indicator of the therapeutic effect of stimulants in narcoleptic patients, which should be tested in further studies.

CONCLUSION

PST is a validated qualitative measure of daytime sleepiness in healthy sleep-deprived subjects and patients with sleep disorders. Measuring daytime sleepiness may be useful for diagnosing and treating various sleep disorders. Although PST cannot fully substitute for MSLT in the experimental setting, its ease of administration compared to the latter makes it a welcome option for assessing physiological sleepiness. However, it seems to be the more reliable correlate of sleepiness than subjective evaluation (e.g., ESS, SSS or VAS). Other methods of pupillographic measurement of sleepiness provided conflicting results. More time-efficient, convenient, and perhaps even more accurate methods may be developed in the future.

DEDICATION

Supported by Ministry of Health, Czech Republic – conceptual development of research organization (FNOL, 00098892)

REFERENCES

1 Axelsson J, Ingre M, Kecklund G, Lekander M, Wright KP, Sundelin T (2020). Sleepiness as motivation: a potential mechanism for how sleep deprivation affects behavior. Sleep. **43**(6): zsz291.
2 Bafna T & Hansen JP (2021). Mental fatigue measurement using eye metrics: A systematic literature review. Psychophysiology. **58**(6): e13828.
3 Barger LK, Ayas NT, Cade BE, Cronin JW, Rosner B, Speizer FE, et al. (2006). Impact of extended-duration shifts on medical errors, adverse events, and attentional failures. PLoS Med. **3**(12): e487.
4 Bumke O (1911). Die Pupillenstorungen, bei Geistes- und Nervenkrankheiten. (Physiologie und Pathologie der Irisbewegungen.) Jena: Fischer.

5 Carskadon MA (1993). Evaluation of excessive daytime sleepiness. Neurophysiol Clin. **23**(1): 91–100.
6 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**(4): 519–524.
7 Danker-Hopfe H, Kraemer S, Dorn H, Schmidt A, Ehlert I, Herrmann WM (2001). Time-of-day variations in different measures of sleepiness (MSLT, pupillography, and SSS) and their interrelations. Psychophysiology. **38**(5): 828–835.
8 Eggert T, Sauter C, Popp R, Zeitlhofer J, Danker-Hopfe H (2012). "Vigilance" of the German Society for Sleep Research and Sleep Medicine (DGSM). The Pupillographic SleepinessTest in adults: effect of age, gender, and time of day on pupillometric variables. Am J Hum Biol. **24**(6): 820–828.
9 Hobson DE, Lang AE, Martin WRW, Razmy A, Rivest J, Fleming J (2002). Excessive daytime sleepiness and sudden-onset sleep in Parkinson disease: a survey by the Canadian Movement Disorders Group. 2002; **287**: 455–463.
10 Hoddes E, Dement WC, Zarcone V (1972). The development and use of the Stanford Sleepiness Scale (SSS). Psychophysiology. **9**: (150).
11 Horne JA & Reyner LA (1995). Sleep-related vehicle accidents. BMJ **310**: 565–567.
12 Johns MW (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. **14**(6): 540–5.
13 Johns MW (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep. **15**: 376–381.
14 Kardon RH, Kirkali PA, Thompson HS (1991). Automated pupil perimetry. Pupil field mapping in patients and normal subjects. Ophthalmology. **98**(4): 485–495.
15 Kelbsch C, Strasser T, Chen Y, Feigl B, Gamlin PD, Kardon R, et al. (2019). Standards in Pupillography. Front Neurol.10:129. Erratum in: Front Neurol. **10**: 371.
16 Kraemer S, Danker-Hopfe H, Dorn H, Schmidt A, Ehlert I, Herrmann WM (2000). Time-of-day variations of indicators of attention: performance, physiologic parameters, and self-assessment of sleepiness. Biol Psychiatry. **48**(11): 1069–1080.
17 Lichstein KL & Fischer SM (1985). Insomnia: In: Hersen M & Bellack AS (Eds) Handbook of clinical behavior therapy with adults (pp, 319–352). New York: Plenum Press.
18 Lichstein KL & Johnson RS (1994). Pupillometric discrimination of insomniacs. Behav Res Ther. **32**(1): 123–129.
19 Lichstein KL, Johnson RS, Sen Gupta S, O'Loughlin DL, Dykstra TA (1992). Are insomniacs sleepy during the day? A pupillometric assessment. Behav Res Ther. **30**(3): 283–292.
20 Lockley SW, Cronin JW, Evans EE, Cade BE, Lee CJ, Landrigan CP, et al. (2004). Effect of reducing interns' weekly work hours on sleep and attentional failures. N Engl J Med. **351**(18): 1829–1837.
21 Lowenstein O & Loewenfeld IE (1951). Types of central autonomic innervation and fatigue; pupillographic studies. AMA Arch Neurol Psychiatry. **66**: 580–599.
22 Lowenstein O & Loewenfeld IE (1952a). Disintegration of central autonomic regulation during fatigue and its reintegration by psychosensory controlling mechanisms. I. Disintegration; pupillographic studies. J Nerv Ment Dis. **115**: 1–21.
23 Lowenstein O & Loewenfeld IE (1952b). Disintegration of central autonomic regulation during fatigue and its reintegration by psychosensory controlling mechanisms. II. Reintegration; pupillographic studies. J Nerv Ment Dis. **115**: 121–145.
24 Lowenstein O & Loewenfeld IE (1958). Electronic pupillography: a new instrument and some clinical applications. AMA Arch Ophthalmol. **59**: 352–363.
25 Lowenstein O, Feinberg R, Löwenstein IE (1963). Pupillary movements during acute and chronic fatigue: a new test for the objective evaluation of tiredness. Invest Ophthalmol Vis Sci. **2**: 138–157.
26 Ludtke H, Wilhelm B, Adler M, Schaeffel F, Wilhelm H (1998). Mathematical procedures in data recording and processing of pupillary fatigue waves. Vision Res. **38**(19): 2889–2896.
27 Lüdtke H, Körner A, Wilhelm B, Wilhelm H (2000). Reproduzierbarkeit des pupillographischen Schläfrigkeitstests bei gesunden Männern Reproducibility of the pupillographic sleepiness test in healthy men. Somnologie. **4**: 170–117.
28 McLaren JW, Erie JC, Brubaker RF (1992). Computerized analysis of pupillograms in studies of alertness. Invest Ophthalmol Vis Sci. **33**: 671–676.
29 McLaren JW, Fjerstad WH, Ness AB, Graham MD, Brubaker RF (1995). New video pupillometer. Opt Eng. **34**(3): 676–683.
30 McLaren JW, Hauri PJ, Lin SC, Harris CD (2002). Pupillometry in clinically sleepy patients. Sleep Med. **3**(4): 347–352.
31 Merritt SL, Schnyders HC, Patel M, Basner RC, O'Neill W (2004). Pupil staging and EEG measurement of sleepiness. Int J Psychophysiol. **52**(1): 97–112.
32 Mitler MM & Miller JC (1996). Methods of testing for sleepiness corrected. Behav Med. **21**(4): 171–183.
33 Mitler MM, Gujavarty KS, Browman CP (1982). Maintenance of Wakefulness Test: a polysomnographic technique for evaluating treatment efficacy in patients with excessive somnolence. Electroen Clin Neurophys. **53**: 658–661.
34 Newman J & Broughton R (1991). Pupillometric assessment of excessive daytime sleepiness in narcolepsy-cataplexy. Sleep. **14**(2): 121–129.
35 Nowak W & Kasprzak HT (2008). Time-frequency analysis of spontaneous fluctuation of the pupil size of the human eye. Optica Applicata. **38**(2): 469–480.
36 Nowak W, Żarowska A, Szul-Pietrzak E, Misiuk-Hojło M (2014). System and measurement method for binocular pupillometry to tudy pupil size variability. Biomed Eng Online. **13**: 69.
37 Philip P, Taillard J, Niedhammer I, Guilleminault C, Bioulac B (2001). Is there a link between subjective daytime somnolence and sickness absenteeism? A study in a working population. J Sleep Res. **10**: 111–115.
38 Pilley SF & Thompson HS (1975). Pupillary "dilatation lag" in Horner's syndrome. Br J Ophthalmol. **5**(12): 731–735.
39 Pivik RT (1991). The several qualities of sleepiness: psychophysiological considerations. In: Monk, T.H. (Ed.), Sleep, Sleepiness and Performance. Wiley, New York: 3–37.
40 Posner M & Rafal R (1987). Cognitive theories of attention and the rehabilitation of attentional deficits. In: Meier, M., Benton, A., Diller, L. (Eds.), Neuropsychological Rehabilitation. Churchill Livingstone, Edinburgh: 182–201.
41 Prasad B, Choi YK, Weaver TE, Carley DW (2011). Pupillometric assessment of sleepiness in narcolepsy. Front Psychiatry. **2**: 35.
42 Ranzijn R & Lack L (1997). The pupillary light reflex cannot be used to measure sleepiness. Psychophysiology. **34**: 17–22.
43 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**(9): 902–912.
44 Reimann M, Manz R, Prieur S, Reichmann H, Ziemssen T (2009). Education research: cognitive performance is preserved in sleep-deprived neurology residents. Neurology. **73**(21): e99–e103.
45 Sauter C, Popp R, Danker-Hopfe H, Büttner A, Wilhelm B, Binder R, et al. (2007). Normative values of the German Epworth sleepiness scale. Results from a multicenter study. Somnologie. **11**: 272–278.
46 Suzuki H, Moriguchi K, Matsuura M, Kojima T, Matsuda T, Noda Y, et al. (2000). Two nap sleep test: an easy objective sleepiness test. Psychiatry Clin Neurosci. **54**(3): 285–286.
47 Szabadi & Bradshaw CM (1996). Autonomic pharmacology of a2-adrenoceptors. J Psychopharmacol. **10**(Suppl. 3): 6–18.
48 Thurtell MJ, Bruce BB, Rye DB, Newman NJ, Biousse V (2011). The Berlin questionnaire screens for obstructive sleep apnea in idiopathic intracranial hypertension. J Neuroophthalmol. **31**(4): 316–319.
49 Wilhelm B, Bittner E, Hofmann A, Koerner A, Peters T, Lüdtke H, et al. (2015). Short-term reproducibility and variability of the pupillographic sleepiness test. Am J Hum Biol. **27**(6): 862–866.
50 Wilhelm B, Giedke H, Lüdtke H, Bittner E, Hofmann A, Wilhelm H (2001). Daytime variations in central nervous system activation measured by a pupillographic sleepiness test. J Sleep Res. **10**(1): 1–7.
51 Wilhelm B, Lüdtke H, Peters T, Schmid R, Wilhelm H, Zrenner E (2001). Automatisierter Swinging-flashlight-Test bei Patienten mit Sehnervenerkrankungen [Automated swinging flashlight test in patients with optic nerve diseases]. Klin Monbl Augenheilkd. **218**(1): 21–25.

52 Wilhelm B, Neugebauer P, Lüdtke H, Hohenstein E, Ederle K, Wilhelm H (1999). Pupillographischer Schläfrigkeitstest zur Therapiekontrolle beim Schlafapnoe-Syndrom nach drei Monaten nächtlicher Beatmung. Somnologie. **3**(2): 53–56.
53 Wilhelm B, Rühle KH, Widmaier D, Lüdtke H, Wilhelm H (1998). Objektivierung von Schweregrad und Therapieerfolg beim obstruktiven Schlafapnoe-Syndrom mit dem pupillographischen Schläfrigkeitstest. Somnologie. **2**: 51–57.
54 Wilhelm B, Wilhelm H, Lüdtke H, Streicher P, Adler M (1998b). Pupillographic assessment of sleepiness in sleep-deprived healthy subjects. Sleep. **21**(3): 258–265.
55 Wilhelm BJ, Körner A, Heldmaier K, Mol IK, Wilhelm H, Lüdtke H (2001b). Normwerte des pupillographischen Schläfrigkeitstests für Frauen und Männer zwischen 20 und 60 Jahren Normal values of pupillographic sleepiness tests for men and women aged 20 to 60 years. Somnologie. **5**: 115–120.
56 Wilhelm BJ, Widmann A, Durst W, Heine C, Otto G (2009). Objective and quantitative analysis of daytime sleepiness in physicians after night duties. Int J Psychophysiol. **72**(3): 307–313.
57 Wilhelm H, Lüdtke H, Wilhelm B (1998a). Pupillographic sleepiness testing in hypersomniacs and normals. Graefes Arch Clin Exp Ophthalmol. **236**(10): 725–729.
58 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**(5): 361–367.
59 Zwyghuizen-Doorenbos A, Roehrs T, Schaefer M, Roth T (1988). Test-retest reliability of the MSLT. Sleep. **11**: 562–565.