Journal of Psychosomatic Research 69 (2010) 81 – 89

Measurements of sleepiness and fatigue

Azmeh Shahid[a,b,*], Jianhua Shen[a,c], Colin M. Shapiro[a,b,c]

[a]Department of Psychiatry, University Health Network, University of Toronto, Toronto, Canada
[b]Youthdale Treatment Center, University of Toronto, Toronto, Canada
[c]Sleep Research laboratory, Toronto, Canada

Received 17 July 2009; received in revised form 9 April 2010; accepted 9 April 2010

Abstract

Sleepiness and *fatigue* are terms commonly used in clinical practice and research. At times sleepiness and fatigue are used interchangeably; however, each of them has distinct implications for diagnosis and treatment. The objective of this article is to review the psychometric properties of the measurements of sleepiness and fatigue. Although there are objective and subject measures to evaluate sleepiness, only rating scales are available to assess fatigue. Further research should be directed toward exploring the potential mechanisms underlying the measurements of sleepiness and fatigue. Establishing objective assessing instruments to evaluate fatigue and clarifying the relationship between objective and subjective assessments of sleepiness are crucially needed.

Keywords: Fatigue; Measurement; Rating scale, Sleepiness

Introduction

Sleepiness and *fatigue* are terms commonly used in both clinical practice and research literature. Both sleepiness and fatigue are ubiquitous phenomena. Sleepiness and fatigue negatively effect daily functioning, and patients who have these feelings are distressed.

Although sleepiness and fatigue are two different and distinct entities, many patients and unfortunately many medical practitioners are unaware of the complexity and heterogeneity of these symptoms. This may be because that some patients use the terms tired, sleepy and fatigued interchangeably and it is difficult to tease apart whether the primary issue is fatigue or sleepiness. The two complaints have distinct implications for clinical diagnosis and treatment. These two symptoms are particularly common in the psychosomatic field. It is beyond the scope of this primarily methodological review to discuss the implications of the distinctions but the interested reader is referred to the following references [1–3].

Currently, there are objective methods measuring sleepiness. A number of rating scales are being used to subjectively assess sleepiness and fatigue. The objective of this review is to assess the psychometric properties of the objective and subjective measurements available on the topic of sleepiness and fatigue. A broader compilation of scales relevant to sleep medicine will be published in the book titled “One Hundred Sleep Related Scales” [4].

Sleepiness

Generally, sleepiness means an increased propensity to doze off or fall asleep; it may be related to a low arousal level [5]. Sleepiness is also defined as a tendency to fall asleep.

Sleepiness may be affected by different conditions, such as medical problems, psychiatric diseases and primary sleep disorders. When sleepiness occurs at an inappropriate time or

* Corresponding author. Toronto Western Hospital Bathurst St 7th floor Main Pavilion Rm 427 Toronto ON, Canada M5T 2S8. Tel.: +1 416 603 5273; fax: +1 416 603 5292.

E-mail addresses: azmeh.shahid@uhn.on.ca, azmehs@hotmail.com (A. Shahid).

0022-3999/10/$ – see front matter
doi:10.1016/j.jpsychores.2010.04.001

an atypical situation, it becomes pathological. For example, excessive daytime sleepiness (EDS) is an important feature of narcolepsy. Sleepiness may be influenced by mood, motivation, autonomic and physiological changes, and fatigue and sleep requirement of the individual [6]. Because of this, it is difficult to obtain an accurate prevalence of EDS. The reported range of the prevalence of EDS is between 3% and 22.6% in different studies [7].

Sleepiness is a multifactorial phenomenon, and it may possess ‘Trait’ and ‘State’ features [8]. Trait features are those could be influenced by individual personality and person specific features; they are usually long term and stable; while state features are short term. Therefore, authors have recommended that EDS should be measured using various rating techniques and it should be assessed on a sound empirical basis. Because sleepiness increases the risk of the occupational injuries and automobile accidents, it is important to assess and treat daytime sleepiness [9,10].

There are objective and subjective measures to assess excessive daytime sleepiness.

Objective measurements of sleepiness

Multiple Sleep Latency Test

The Multiple Sleep Latency Test (MSLT) was first described by Carskadon and Dement [11]. The underlying background of establishing the MSLT is that sleepiness is a physiological need for sleep, while increasing tendency to fall asleep indicates greater sleepiness. During the test, an individual is required to lie down in a dark room to fall asleep. An MSLT includes four or five sessions; each session lasting 20–30 min and performed at two hourly intervals. Sleep latency is measured by standard electrophysiological means and is defined as the time elapsed from lights out to the first epoch of any stage of sleep. The two important outcome variables are the mean sleep onset latency and the number of rapid eye movement sleep episodes that occur.

The validity and reliability of the MSLT have been documented in several clinical and experimental situations [12]. The inter-rater reliability and test–retest reliability are acceptable. One study found that over a time span of 4 and 14 months, the test–retest reliability of MSLT was good to excellent (r values were between 0.65 and 0.97) [13].

MSLT is a useful tool to assess sleepiness induced by various conditions, including acute and partial sleep deprivation, circadian rhythm disorders, disrupted sleep, sleep apnea, and narcolepsy, use of hypnotics and alcohol usage and idiopathic hypersomnolence. A mean sleep latency of 5 min or shorter on the MSLT represents severe pathological sleepiness [14–17] (please see Table 1).

The measurement cutoff points of the MSLT affect its sensitivity and specificity to assess EDS. The sensitivity and specificity were 80.9% and 89.8%, respectively, when the cutoff of 5 min was used; 94.5% and 73.3%, respectively, when the cutoff was raised to 8 min; and 52% and 98.3%, respectively, when the cutoff was 3 min or shorter. Clinically it is useful to think of severe, moderate and mild daytime sleepiness, but little effort has been made to validate these distinctions.

Although MSLT is a reliable, valid and accurate test, it may fail to measure some aspects of sleepiness [8]. For instance, MSLT may not separate an individual’s ability to fall asleep (sleep propensity) from his need to sleep. This may be because the MSLT ignores the effects of an “arousal factor” on sleepiness. An “arousal factor” is generated internally and it is independent of the environment [18].

Maintenance of Wakefulness Test

The major function of the Maintenance of Wakefulness Test (MWT) is to measure the strength of the arousal system. While some may view MWT as a useful tool in estimating daytime sleepiness we would not share this view. Although there are some correlations between sleepiness and decreased the level of alertness, sleepiness is not the reverse of alertness. Clinically, many patients who have significant sleepiness, have little impairment of alertness. During the process of doing the MWT, subjects are instructed to sit in a dimly lit room for 30 min and attempt to stay awake. The duration may be set at 20, 30 or 40 min. We have preferred 30 min as a compromise between a ceiling effect (with 20 min) and the onerous 40 min version. It has also facilitated performance tasks at the midpoint of successive tests. Sleep onset is defined as the first appeared three consecutive 30 second epochs of stage 1 sleep or the first epoch of any other sage of sleep [19]. Based on studies we and others have completed [19], using the 30 min protocol, a mean sleep latency of the MWT is determined by the mean value of the sleep onset latency on the four naps. A mean sleep latency between 19 and 24 min indicates mild impairment of alertness; between 13 and 18.9 min indicates moderate impairment of alertness; and <12.9 min indicates severe impairment (Table 1).

When the cutoff of the MWT was set at 12 min, the sensitivity was 84.3%, and the specificity was 98.4%. The MWT has been used to evaluate the response of pharmacological treatment in narcolepsy [20,21].

Subjective measures of sleepiness

Subjective measures of sleepiness mainly rely on rating scales. The majority of the rating scales are self-evaluated instruments. Using a rating scale to measure sleepiness is inexpensive, simple and less time-consuming [8]. Self-reported scales have some drawbacks. These include unintended bias and purposeful falsification.

Sleepiness rating scales broadly include two categories. The first category is used to assess “State” sleepiness, such as the Stanford Sleepiness Scale (SSS), Karolinska Sleepiness Scale (KSS), and Visual Analogue Scales. These scales are used to measure short term changes in sleepiness. Other tests measure a global level of sleepiness, most like a “Trait”

Table 1
Features of the instruments for measuring sleepiness

Scale Name	MSLT	MWT	KSS	ESS	SSS	SWAI	DSS
Reference	12, 13,16–18	19–21	32–34	18, 35–37	23–25,29,31	39–42	38,43,44
What is assessed (construct)	How quickly one falls asleep when asked to do so	Measures strength of the arousal system	Situational sleepiness	General level of sleepiness	Sleepiness in patients with sleep complaints	Multidimensional measurement of sleepiness	Originally developed as a part of the SWAI
Item number	PSG recording	PSG recording	9 items	8 items	7 levels	59 items	8 items
Scale type	Investigator administered	Investigator administered		4 points (0–3)	Likert	9 semi-continuous scale	4 points ordinal scale:
Target population	Subjects with daytime sleepiness	Subjects with impaired alertness			Patients with sleep complaints	Daytime sleepiness	Daytime sleepiness
Internal consistency (α)				0.88		0.89 (for EDS component)	0.71
Test–retest reliability (r)	0.65–0.97	–	–	0.822	–		–
Concurrent validity (r)	Gold standard		Tested with EEG variable			Tested with MSLT	Tested with MSLT
Cut off point	10–15 min, mild sleepiness or normal; 5–10 min moderate sleepiness; <5 min severe sleepiness	19–24 min, mild; 13–19 min, moderate;≤2.9 min severe impairment of alertness		Total score is between 0–24; 10 or over indicates abnormal or pathological sleepiness		1 indicates that the behavior was always present, 9 meant never present and 5 meant sometimes present.	Total score is from 0 to 24; scores 0–10: normal (no significant skew); ≥11: significant skewdness.

aspect of sleepiness. The Epworth Sleepiness Scale (ESS) and Sleep Wake Activity Inventory (SWAI) belong in this category.

Stanford Sleepiness Scale

The SSS was developed by Hoddes et al. [22]. It is one of the most widely used subjective sleepiness measuring instrument. This self-rating scale is used to quantify sleepiness levels in patients with sleep complaints (Table 1) The SSS is a Likert-type scale with seven vigilance levels. Subjects are asked to indicate which level best describes their current state. Studies have suggested that the SSS is sensitive to reveal sleepiness induced by sleep deprivation [23,24]; however, it is not so sensitive with patients suffering from sleep disorders and tending to deny sleepiness [25,26]. This may partially explain why the SSS is not sensitive to distinguish between sleep apnea subjects and normal sleepers, even though for patients with narcolepsy it has been a useful subjective sleepiness measurement scale [27].

John [28] had indicated that the SSS is not a valid measure to assess sleep propensity. He investigated the concurrent validity of the SSS with performance on mental tasks and evaluated whether the SSS shows any changes in sleepiness with sleep loss. The results indicated that the mean SSS ratings positively correlated (r=0.68) with those of performance on the Wilkinson test, which was developed by Wilkinson (1968), and it is used to measure performance following a reduction of sleep time by two hours. The SSS ratings had a lower correlation with the performance on a memory test (r=0.47). Following sleep deprivation, the mean SSS ratings were significantly lower than the corresponding ratings with no deprivation. This may be because that sleepiness is a multi-dimensional phenomenon [29]. Additionally, the SSS may significantly correlate with fatigue. This makes the usage of the scale complicated [30].

Karolinska Sleepiness Scale

The KSS was developed by Akerstedt and Gillberg [31]. This nine-point scale measures the subjective level of sleepiness at a particular time during the day. On this scale, subjects indicate which level best reflects the psycho-physical sate experienced in the last 10 min. The KSS is a measure of situational sleepiness. It is sensitive to fluctuations. Scores on the KSS increase with increased periods of wakefulness and strongly correlate with the time of the day [32] (Table 1).

Kaida et al. have investigated the validity of the KSS and found that the KSS was highly correlated to electroencephalography (EEG) and behavioral variables. These results indicate that the validity of the KSS is high [33]. However, because the scores of the KSS vary according to earlier sleep, time of day and other parameters, it is difficult to deduce its test–retest reliability.

The KSS is useful in assessing the changes in response to environmental factors, circadian rhythm, and effects of drugs. Because the KSS is not a measure of Trait sleepiness, it has not been widely used for clinical purposes.

Epworth Sleepiness Scale

The ESS was developed by Johns [34]. The background of establishing the ESS was driven from the observations of the nature and occurrence of daytime sleepiness. This simple self-administered scale is used to measure general level of daytime sleepiness.

The ESS has a set of eight situations commonly encountered in daily life. Some of them are very soporific, others less so. On a scale of 0–3, subjects rate how quickly they would fall asleep or doze off in each of the eight situations. The total score is between 0 and 24. A score of 10 or higher indicates abnormal or pathological sleepiness. The ESS tries to deal with the fact that people have different daily routines; some of these routines promote sleep while others are more activating (Table 1).

According to Johns [18], the ESS is the most discriminative test, so far, on average sleep propensity. One study by Johns showed that the ESS was the more discriminative test compared with MSLT and MWT. The ESS is sensitive to the changes produced by treatments for sleep apnea [34–36]. Although using the ESS we may discriminate normal from pathological subjects, we need to develop good objective tests to measure sleep propensity in different situations in which the subjects posture and activity inter alia is also taken into consideration [18].

Johns has investigated the reliability and internal consistency of the ESS. Over a period of 5 months, the Pearson correlation between the scores of two time points was 0.822, indicating that the test–retest reliability of the ESS was high. The ESS has a high level of internal consistency with a Cronbach's alpha of 0.88 (Table 1).

Sleep Wake Activity Inventory

The SWAI was developed by Rosenthal [37]. The purpose of developing the SWAI was to establish a measure to assess multidimensional components of sleepiness. The items on the SWAI were derived from clinical experience and an appraisal of earlier self-report scales [38–41]. The final version of the SWAI consists of 59 items. A Likert-type scale is used for each item, which includes a 1–9 semicontinuous scale. While performing the assessment, subjects are asked to circle the number that best describes them: 1 indicates that the behavior is always present, 9 means never present and 5 means sometimes present. Subjects are asked for their response over the last 7 days (Table 1).

A high Cronbach's alpha (0.89) is obtained for the factor of Excessive Daytime Sleepiness (EDS). Cronbach's alpha values of other factors are moderate: 0.72 for Psychic Distress, 0.76 for Social Desirability, 0.71 for Energy Level, 0.69 for Ability to Relax, and 0.69 for Sleep. The SWAI is the first self-report instrument to assess sleepiness that has been shown to be sensitive to various sleepiness levels. The MSLT was used for validating the SWAI, and the EDS factor of the SWAI showed a high predictive value of mean MSLT scores.

The EDS scores on the SWAI are helpful to differentiate pathological sleepiness from diagnostic gray zone and normal levels of sleepiness. The scores of the SWAI are generally improved following normalization of sleep disordered breathing. Further studies are needed to establish the reliability of the EDS factor in the absence of an intervening variable over time.

Daytime Sleepiness Scale and Nocturnal Sleep Onset Scale

The Daytime Sleepiness Scale (DSS) and Nocturnal Sleep Onset Scale (NSOS) are self-report measures originally part of the SWAI [37]. In the study performed by Johnson et al., the psychometric evaluation of the DSS and NSOS was assessed in a representative community sample [42].

A DSS has eight items. Items are rated on the base of the two week period preceding the interview. Each item has a four-point ordinal scale: never, rarely, sometimes and often. The score on the DSS is from 0 to 24. Scores 0–10 indicate "normal" (no significant skew), while scores of 11 or higher indicate significant positive skew (Table 1).

The revised eight-item DSS has a good internal consistency with a Cronbach's alpha of 0.71. The construct validity of the DSS is acceptable. There is a similarity on the thresholds of the MSLT and the DSS. It is generally agreed that the threshold for pathological sleepiness for MSLT <5 min describes 10–15% of normal healthy volunteers, while the proposed threshold in DSS identifies 17% of the population as potentially having abnormal levels of daytime sleepiness [42,43].

Revised NSOS includes two items. Rating of the items is based on the two week period preceding the interview. As in the DSS, each item of the NSOS has a four-point ordinal scale: never, rarely, sometimes and often. The NSOS has a reasonable internal consistency with a Cronbach's alpha of 0.58.

FACES

The FACES is a 50-item adjectival checklist which is useful in assessing both sleepiness and fatigue at the same time. It is discussed in more detail in the section on scales for fatigue in this article [44].

Toronto Sleepiness and Fatigue Scale

The Toronto Sleepiness and Fatigue Scale (TSFS) was developed by Shen et al. [45]. The objective of this study was to develop a questionnaire to measure sleepiness and fatigue concurrently. This 10-item questionnaire includes sleepiness (TSFS-S) and fatigue assessments (TSFS-F) related to a common stem. It was showed that the Pearson correlation coefficients (r) between scores on the TSFS and those of the validated scales measuring sleepiness and fatigue, including the Epworth Sleepiness Scale and Fatigue Severity scale were between 0.54 and 0.69 on Day 0 and between 0.46 and 0.71 on Day 7. Cronbach's alphas were between 0.84 and 0.87 for the TSFS-S and TSFS-F. Intraclass correlation coefficients were 0.87 for the TSFS-S and 0.80 for the TSFS-F. The TSFS has been shown to be a useful instrument to measure sleepiness and fatigue simultaneously and may be particularly useful in trying to discern if a treatment influences fatigue or sleepiness to a greater or lesser extent.

The problem hitherto is that these were separate scales but changes in a sleepiness scale would not be accurately compared to changes in a fatigue scale.

Fatigue

Fatigue is common in physical and psychiatric disorders [46]. Symptoms of fatigue are commonly reported in patients with depression, chronic fatigue syndrome, HIV, cancer. Fatigue may be a side effect of a number of medication treatments [47]. Generally, fatigue is not specific in its presentation and symptomatology. Fatigue may be induced by physical, physiological and psychological causes; it often presents as a feeling of tiredness and exhaustion [48].

Fatigue is the most common symptom reported to the physicians. The prevalence rate of fatigue in primary care ranges between 7% and 45% [49]. This large difference in the prevalence rate may be because of a lack of a working definition for fatigue and differences in measurement techniques.

Fatigue is a feeling of strain or exhaustion; it includes physiological fatigue and pathological fatigue. Physiological fatigue, or “normal fatigue,” is induced by daily activities it lasts a short period and is usually relived by rest [50]. Pathological fatigue is usually caused by a medical or emotional disorder or an intervention for a disease, and is more chronic in nature [51].

Due to lack of objective measures and assessment tools, fatigue is still poorly defined and measured [52]. Research is needed to develop “gold standard” to objectively measure fatigue. This will help to quantify and strategize the treatment of the patients who have fatigue.

There are many subjective rating scales which have been developed to measure fatigue. The following eleven scales are commonly used to evaluate fatigue. These scales provide a wide range of tools which are of interest both in the mental health field and in the somatic domain.

Brief Fatigue Scale

The Brief Fatigue Scale (BFI) was developed by Mendoza and colleagues [53] to assess the severity of fatigue and the impact of fatigue on daily functioning in patients with cancer and during cancer treatment. The BFI assesses the severity of fatigue and the impact of fatigue on daily functioning in past 24 h. It has nine items; each item is rated on an 11-point Likert scale type. A global fatigue score can be obtained by averaging all the items. The BFI has no cutoff score. The authors have stated that the scale is easily understood and translated.

Reported Cronbach’s alpha of the BFI ranges from 0.82 to 0.96, indicating that the internal consistency of the scale is high. The concurrent validity was established by comparing the scores of the BFI and those of the Profile of Mood Status (POMS)-Fatigue Subscale (POMS-F) and Functional Assessment of Cancer Therapy (FACT)-Fatigue Subscale (FACT-F). The results showed that correlation between the scores of the BFI and the POMS-F and those between the BFI and the FACT-F are highly significant (Table 2).

Although the BFI has been accepted as a reliable instrument to assess fatigue levels in cancer patients, it has not been validated in non-cancer patients. The test–retest reliability of the BFI has not been established.

Functional Assessment of Cancer Therapy

FACT was developed by Yellen et al. [54]. It includes a set of comprehensive questions used to measure health related quality of life in patients with chronic diseases, especially cancer. The general version of the FACT (FACT-G) has 34 items [54].

The FACT-F has 13 items, which has been validated and is used alone. The construct of the FACT-F is similar to those of the POMS-F, the POMS-Vigor Subscale and the Piper Fatigue Scale. The FACT-F is mainly used in cancer patients and has not been validated in other populations. The scale has high internal consistency with the Cronbach’ alpha of 0.93 and has excellent test–retest reliability with a correlation coefficient of 0.90 (Table 2).

Fatigue severity scale

The Fatigue Severity Scale (FSS) is a nine-item self-report questionnaire. While scoring the items, respondents have to mark from one to seven, where “1” indicates no fatigue and “7” indicates severe fatigue. High scores indicate more fatigue. The FSS measures the impact of fatigue on functioning and behavioral aspects, rather than the intensity of symptoms related to fatigue [55].

The FSS is one of the most widely used fatigue evaluating scales. It has high internal consistency with Cronbach’s alpha between 0.81 and 0.89 [46]. It also has good test–retest reliability with a correlation coefficient of 0.84 and is sensitive to changes with time and treatment. The FSS has shown good concurrent validity (with Visual Analogue Scale, r=0.68) [55] and is able to distinguish patients with different diagnoses, such as systemic lupus erythematosus, multiple sclerosis and chronic fatigue syndrome, shift workers, and depression [51,55]. The scale may predict changes in fatigue over time. It is capable of showing clinical improvement with treatment. It is applicable in both clinical and research settings.

Checklist Individual Strength

The Checklist Individual Strength (CIS) is a 24-item questionnaire designed to measure several aspects of fatigue. It was developed by Vercoulen [56] for hospital studies. Each item on the scale is scored on a seven-point Likert scale.

The CIS has four subscales: Subjective Experience, Concentration, Motivation and Physical Activity. It has

Table 2
Psychometric features of fatigue measurement scales

Scale Name	BFI	FACT-F	FSS	CIS	FAI	FIS	CFS	FSI	MFI	PFS	FACES
Reference	54	55	47,52,53,56	57	58	59	60	61, 62	63,64	51	45
What is assessed (construct)	Fatigue severity and impact on daily functioning in cancer patients and cancer treatment	Health related quality of life in patients with cancer and chronic diseases	Impact of fatigue on functioning	Fatigue in patients with chronic fatigue syndrome in hospital studies	Qualitative and quantitative aspects in patients with medical problems	Impact of fatigue on the quality of life	Severity of fatigue in patients with chronic fatigue syndrome	Severity, impact and duration of fatigue	Multidimensional features of fatigue	In cancer patient in research	Characterize different aspects of sleepiness states
Item number	9	13	9	24	29	40	14	–	20	41	50
Scale type	Self-report 11–point Likert	Rater administered	Self-report	–	Self-report Likert	–	Responses on a continuum	Multidimensional	Self-report	Visual analogue scale	Self-report
Target population	Cancer patients	Cancer patients	Fatigued Subjects	Chronic fatigue syndrome	Medical disorder patients		Chronic fatigue syndrome	Women with breast cancer	Cancer patients	Cancer research	Fatigued patients
Internal consistency	0.82–0.96	0.93	0.81–0.89	0.96	0.70–0.92	0.87	0.88–0.90	0.94	0.84	High	Good
Test–retest reliability	–	0.90	0.84	–	0.29–0.69	–	–	Weak		–	–
Concurrent validity	Tested with POMS-F, FACT-F	POMS-Fatigue and Vigor Subscales and PFS		–	Extended version of FSS	–	–	–	Compared within and between groups	Moderate	–

been validated for patients with chronic fatigue syndrome (CFS). The CIS has a good internal consistency with a Cronbach's alpha of 0.96. It is able to discriminate fatigue levels in patients with multiple sclerosis, CFS and healthy controls. The test–retest reliability of the CIS has not been demonstrated (Table 2).

Fatigue Assessment Inventory

The Fatigue Assessment Inventory (FAI) has 29 items. It was developed by Schwartz [57] to assess both qualitative and quantitative aspects of fatigue in medical patients in whom fatigue is the major symptom (Table 2).

The FAI has four subscales namely: Fatigue Severity, Situation Specificity, Psychological Consequences and Response to Rest/Sleep. In general, the FAI has good psychometric properties. It shares some items with those of the FSS and has good convergent validity. The FAI has a high correlation with the FSS (r=0.98). The test–retest reliability is moderate with correlation coefficients between 0.29 and 0.69. The internal consistency is reasonably high with Cronbach's alpha between 0.70 and 0.92. The FAI has been validated in out-patients in neurological and rheumatologic clinics. It is able to distinguish healthy subjects from patients.

Fatigue Impact Scale

The Fatigue Impact Scale (FIS) was developed by Fisk [58] to assess the impact of fatigue on quality of life and to improve understanding of the effects of fatigue. The scale has 40 items and is intended to assess the impact of fatigue on cognitive, physical and psyschosocial functions.

The FIS has a high internal consistency with a Cronbach's alpha of 0.87. It moderately correlates with the Sickness Impact Profile (r=0.51). The FIS has been validated in patients with hypertension and multiple sclerosis (Table 2).

Chalder Fatigue Scale

The Chadler Fatigue Scale was developed to measure severity of fatigue in patients with chronic fatigue syndrome [59]. This 14-item scale has questions related to physical and mental fatigue. The Chadler Fatigue Scale can be scored by using a Likert scale; each item has four options: better than usual, no more than usual, and worse than usual and much worse than usual.

The Chadler Fatigue Scale has been found to be reliable. The reported Cronbach's alpha values of the scale are between 0.88 and 0.90. The Chadler Fatigue Scale is a valid estimator of change and is used to assess the symptom severity (Table 2). The scale is useful in detecting fatigue in epidemiological studies. However, it is recommended that, clinically, the scale should not be used alone, but as an adjunct measuring instrument.

Fatigue Symptom Inventory

The Fatigue Symptom Inventory (FSI) was developed by Hann et al. [60]. This multidimensional scale is used to assess the severity, impact and duration of fatigue. The initial standardization of the inventory was performed in women who were diagnosed with breast cancer, and who were undergoing treatment for cancer, as well as those who did not have cancer. The FSI has a good internal consistency with a Cronbach's alpha of 0.94. It is widely used in both male and female cancer patients, although its test–retest reliability is reported to be weak [61] (Table 2).

Multidimensional Fatigue Inventory

The Multidimensional Fatigue Inventory (MFI) was developed by Smets et al. [62]. This 20-item self-report instrument has five subscales: General Fatigue, Physical Fatigue, Mental Fatigue, Reduced Motivation and Reduced Activity. The MFI is one of the more comprehensive measures of fatigue used in cancer patients and it needs further development in order to be used in clinical settings [63]. The MFI has shown good internal consistency with a Cronbach's alpha of 0.84 (Table 2). It has been tested for psychometric properties in cancer patients receiving radiotherapy, patients with CFS, psychological and medical students, army recruits and junior physicians. The convergent validity was established by comparison of the MFI with a Visual Analogue Scale measuring fatigue. The correlations of between and within group comparisons are generally acceptable ($0.22<r<0.78$).

Piper Fatigue Scale

The Piper Fatigue Scale (PFS) was developed by Piper [50] and it was originally used in cancer patients for the purpose of research. The PFS is a 41-item Visual Analogue Scale representing the temporal, intensity, affective, and sensory dimensions of fatigue. The validation test was performed in 42 patients. The concurrent validity of the scale was moderate and the internal consistency was high (Table 2). However, the validity of dimensional structures has not been tested.

Visual Analogue Scale for Fatigue

The Visual Analogue Scale for Fatigue (VAS-F) is an 18-item scale to measure fatigue and energy levels. The energy subscale has 5 items and the Fatigue subscale has 13 items. The VAS-F is easily understood and administered. Performing the VAS-F requires little in the way of reading skills and time. The VAS-F has good psychometric properties. The internal consistency of the scale and subscales are good with the values of Cronbach's alpha between 0.91 and 0.96 [64]. The limitations are that subjects should have good motor and visual abilities and some subjects feel hesitant in using the extreme ends of the 100-mm lines. In general our view would be that VAS are useful for longitudinal but not for cross-sectional purposes [65].

FACES

The FACES is a 50-item adjectival checklist designed to characterize different aspects of sleep and fatigue states [44]. The abbreviation of FACES represents its five subscales: Fatigue, Energy, Consciousness, Energized and Sleepiness. The fatigue subscale has 15 items and the sleepiness subscale includes 10 items. It has good convergent and discriminate validities, indicating that the FACES is a promising self-report instrument for the measurement of sleepiness, fatigue and related subjective experiences at the same time (Table 2).

Conclusion

Sleepiness and fatigue are commonly seen in clinical settings, as well as in the general population. Accurately assessing sleepiness and fatigue is crucial for clinical understanding of patients and for research. There are objective and subjective instruments to measure sleepiness. However, only rating scales are available for measuring fatigue. In the listed 11 scales measuring fatigue, the top three commonly used are the FSS, CFS and FIS. During any assessment (clinical purposes or for research) one should be mindful and clear concerning the reasons for using a specific instrument recognizing that each measuring instrument has its own psychometric properties.

Further research should be directed toward exploring the potential mechanisms underlying the measurements of sleepiness and fatigue, clarifying the relationship between objective and subjective assessments of sleepiness and establishing objective assessing instruments to evaluate fatigue.

One key issue in research is the need to tease out the effects of sleepiness versus fatigue. Using a sleepiness and separate fatigue scale is not adequate as changes in one may not parallel equitable to changes in the other. As rated above, a scale that measures both states concomitantly is highly desirable [6]. To some extent the FACES Checklist [44] and TSFS [45] accomplish this.

References

[1] Peck DF, Shapiro CM. Guidelines for the construction, selection and interpretation of measurement devices. In: Peck DF, Shapiro CM, editors. Measuring human problems. a practical guide. England: John Wiley and Sons, 1990. pp. 1–12.
[2] Shapiro CM, Rathouse K. Sleep Disorders. In: Peck DF, Shapiro CM, editors. Measuring Human Problems. A practical guide. England: John Wiley and Sons, 1990. pp. 375–91.
[3] Shapiro CM, Ohayon MM, Huterer N, Grunstein R. Fighting fatigue and sleepiness. Canada: Joli Joco publication, 2006. pp. 1–96.
[4] Shahid A, Wilkingson K, Marcu S, Shapiro CM. One hundred sleep related scales, 2010. Springer, New York. [in press].
[5] Curcio G, Casagrande M, Bertini M. Sleepiness: evaluating and quantifying methods. Int J Psychophysiol 2001;41:251–63.
[6] Shen J, Barbera J, Shapiro CM. Distinguishing sleepiness and fatigue: focus on definition and measurement. Sleep Med Rev 2006;10:63–76.
[7] Ruggles K, Hausman N. Evaluation of excessive daytime sleepiness. WMJ 2003;102:21–4.
[8] Cluydts R, DeValck E, Verstraeten E, Theys P. Daytime sleepiness and its evaluation. Sleep Med Rev 2002;6:83–96.
[9] Masa JF, Rubio M, Findley LJ. Habitually sleepy drivers have a high frequency of automobile crashes associated with respiratory disorders during sleep. Am J Respir Crit Care Med 2000;162:1407–12.
[10] Melamed S, Oksenberg A. Excessive daytime sleepiness and risk of occupational injuries in non-shift daytime workers. Sleep 2002;25: 315–22.
[11] Carskadon MA, Dement WC. Sleepiness and sleep state on a 90-min schedule. Psychophysiol 1977;14:127–33.
[12] Roth T, Roehrs T, Carskadon M, Dement W. Daytime sleepiness and alertness. In: Kryger M, Roth T, Dement W, editors. Principles and practice of sleep medicine. Philadelphia: Saunders, 1989. pp. 14–23.
[13] Zwyhuizen-Doorenbos A, Roehrs T, Scaefer M, Roth T. Test-retest reliability of the MSLT. Sleep 1988;11:562–5.
[14] Carskadon MA, Dement WC, Mitler MM, Roth T, Westbrook PR, Keenan S. Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. Sleep 1986;9:519–24.
[15] Hanly PJ, Shapiro CM. Sleep solutions: excessive daytime sleepiness. Solutions sommeil: hypersomnolence diurne, 4. St. Laurent: Kommunicom Publications, 1992. pp. 1–28.
[16] Thorpy MJ. The clinical use of the Multiple Sleep Latency Test. The Standards of Practice Committee of the American Sleep Disorders Association. Sleep 1992;15:268–76.
[17] Roehrs T, Roth T. Multiple Sleep Latency Test: technical aspects and normal values. Clin Neurophysiol 1992;9:63–7.
[18] Johns MW. Sensitivity and specificity of the multiple sleep latency test (MSLT), the Maintenance Of Wakefulness Test and the Epworth Sleepiness Scale: failure of the MSLT as a gold standard. J Sleep Res 2000;9:5–11.
[19] Doghramji K, Mitler MM, Sangal RB, Shapiro C, Taylor S, Walsleben J, et al. A normative study of the maintenance of wakefulness test (MWT). Electroencephalogr Clin Neurophysiol 1997;103:554–62.
[20] Mitler MM, Gujavarty KS, Browman CP. Maintenance of wakefulness test: a polysomnographic technique for evaluating treatment efficacy in patients with excessive somnolence. Electroenceph Clin Neurophysiol 1982;53:658–61.
[21] Mitler MM, Gujavarty KS, Sampson MG, Browman CP. Multiple daytime nap approaches to evaluating the sleepy patient. Sleep 1982;5 (Suppl 2):S119–127.
[22] Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: a new approach. Psychophysiology 1973;10:431–6.
[23] Hoddes E, Zarcone V, Dement WC. The history and use of the Stanford Sleepiness Scale. Psychophysiology 1972;9:150.
[24] Herscovitch J, Broughton R. Sensitivity of the Stanford sleepiness scale to the effects of cumulative partial sleep deprivation and recovery oversleeping. Sleep 1981;4:83–91.
[25] Carskadon MA, Dement WC. Cumulative effects of sleep restriction on daytime sleepiness. Psychophysiology 1979;18:107–13.
[26] Roth T, Hartze KM, Zorick F, Conway W. Multiple naps and the evaluation of daytime sleepiness in patients with upper airway sleep apnea. Sleep 1980;3:425–39.
[27] Broughton RJ, Aguirre M. Differences between REM and NREM sleepiness measured by event-related potentials (P300, CNV), MSLT and subjective estimate in narcolepsy–cataplexy. Electroencephalogr Clin Neurophysiol 1987;67:317–26.
[28] Johns MW. Daytime sleepiness, snoring, and obstructive sleep apnea. The Epworth Sleepiness Scale. Chest 1993;103:30–6.

[29] MacLean AW, Fekken GC, Saskin P, Knowles JB. Psychometric evaluation of the Stanford Sleepiness Scale. J Sleep Res 1992;1:35–9.
[30] Avni N, Avni I, Barenboim E, Azaria B, Zadok D, Kohen-Raz R, et al. Brief posturographic test as an indicator of fatigue. Psychiatry Clin Neurosci 2006;60:340–6.
[31] Akerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. Int J Neurosci 1990;52:29–37.
[32] Kecklund G, Akerstedt T. Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. Ergonomics 1993;36: 1007–17.
[33] Kaida M, Takahashi T, Åkerstedt A, Nakata Y, Otsuka T, Haratani K, et al. Validation of the Karolinska sleepiness scale against performance and EEG variables. Clin Neurophysiol 2006;117:1574–81.
[34] Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep 1991;14:540–5.
[35] Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep 1992;15:376–81.
[36] Johns MW. Sleepiness in different situations measured by the Epworth Sleepiness Scale. Sleep 1994;17:703–10.
[37] Rosenthal L, Roehrs TA, Roth T. The Sleep-Wake Activity Inventory: a self-report measure of daytime sleepiness. Biol Psychiatry 1993;34: 810–20.
[38] Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193–213.
[39] Cloninger RC. A systemic method for clinical description and classification of personality variants. Arch Gen Psychiatry 1987;44: 573–88.
[40] Miles L. A sleep questionnaire. In: Gulleminault C, editor. Sleeping and waking disorders. Indications and techniques. Melo Park (Calif): Addison-Wesley, 1982. pp. 383–413.
[41] Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54:1063–70.
[42] Johnson RO, Breslau N, Roth T, Roehrs T, Rosenthal L. Psychometric evaluation of daytime sleepiness and nocturnal sleep onset scales in a representative community sample. Biol Psychiatry 1999;45:764–70.
[43] Levine B, Roehrs T, Zorick F, Roth T. Daytime sleepiness in young adults. Sleep 1988;11:39–46.
[44] Shapiro CM, Flanigan M, Fleming JA, Morehouse R, Moscovitch A, Plamondon J, et al. Development of an adjective checklist to measure five FACES of fatigue and sleepiness. Data from a national survey of insomniacs. J Psychosom Res 2002;52:467–73.
[45] Shen J, Chatoo K, Streiner DL, Chung SA, Huterer N, Shapiro CM. A novel questionnaire measuring sleepiness and fatigue concurrently, 2010 [in press].
[46] Loge JH, Ekeberg O, Kaasa S. Fatigue in the general Norwegian population: normative data and association. J Psychom Res 1998;45: 53–65.
[47] Smets EM, Visser MR, Willems-Groot AF, Garssen B, Schuster-Uitterhoeve AL, de Haes JC. Fatigue and radiotherapy: (B) experience in patients 9 months following treatment. Br J Cancer 1998;78:907–12.
[48] Kennedy HG. Fatigue and fatigability. Br J Psychiatry 1988;153:1–5.
[49] Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. J Epid Com health 1992;46:92–7.
[50] Piper B, Lindsey A, Dodd M, Ferketich S, Paul S, Weller S. Development of an instrument to measure the subjective dimension of fatigue. In: Funk S, Tournquist E, Champagne M, Copp L, Weise R, editors. Key aspects of comfort: management of pain, fatigue and nausea. New York: Springer, 1989. pp. 199–208.
[51] Hossain JL, Reinish LW, Kayumov L, Bhuyia P, Shapiro CM. Underlying sleep pathology may cause chronic high fatigue in shift workers. J Sleep Res 2003;12:223–30.
[52] Taylor RR, Jason LA, Torres A. Fatigue Rating Scale: an empirical comparison. Psychol Med 2000;30:849–56.
[53] Mendoza TR, Wand XS, Cleeland CS, Morrissey M, Johnson BA, Wendt JK, et al. The rapid assessment of fatigue severity in cancer patients: use of Brief Fatigue inventory. Cancer 1999;85:1186–96.
[54] Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with functional assessment of cancer therapy (FACT) measurement system. J Pain Symptom Manage 1997;13:63–74.
[55] Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989;46:1121–3.
[56] Vercoulen JHMM, Swanink CMA, Fennis JFM, Galama JMD, vander Meer JWM, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. J Psychosom Res 1994;38:383–92.
[57] Schwartz JE, Janford L, Krupp LB. The assessment of fatigue: a new measurement. J Psychom Res 1993;37:753–62.
[58] Fisk JD, Ritvo PG, Ross L, Haase DA, Marrie TJ, Schlech WF. Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. Clin Infect Dis 1994;18(suppl 1):s79–s83.
[59] Chadler T, Berelowitz G, Pawlikowska T, Watts l, Wessely S, Wright D, et al. Development of a fatigue scale. J Psychomsom Res 1993;37: 147–53.
[60] Hann DM, Jacobsen PB, Azzarello LM, Martin SC, Curran SL, Fields KK, et al. Measurement of fatigue in cancer patients: development and validation of fatigue symptom inventory. Qual Life Res 1998;7: 301–10.
[61] Hann DM, Denniston MM, Baker F. Measurement of fatigue in cancer patients: further validation of fatigue symptom inventory. Qual Life Res 2000;9:847–54.
[62] Smets EMA, Garssen B, Bonk B, De Haes JCJM. The multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Research 1995;39:315–25.
[63] Stein KD, Martin SC, Hann DM, Jacobsen PB. A multidimensional measure of fatigue for use with cancer patients. Cancer Pract 1998;6: 143–52.
[64] Lee KA, Hicks G, Ninomurcia G. Validity and reliability of a scale to assess fatigue. Psychiatry Res 1991;36:291–8.
[65] Maxwell C. Sensitivity and accuracy of the visual analogue scale: a psycho-physical classroom experiment. Br J Clin Pharmac 1978;6: 15–24.