

Document downloaded from:

<http://hdl.handle.net/10251/104017>

This paper must be cited as:

Ibáñez Del Valle, V.; Silva, J.; Cauli, O. (2017). A Survey on Sleep Questionnaires and Diaries. *Sleep Medicine*. 42:90-96. doi:10.1016/j.sleep.2017.08.026



The final publication is available at

<https://doi.org/10.1016/j.sleep.2017.08.026>

Copyright Elsevier

Additional Information



Sleep Medicine

## A Survey on Sleep Questionnaires and Diaries

Vanessa Ibáñez<sup>a\*</sup>, Omar Cauli<sup>a</sup>, Josep Silva<sup>b</sup>

<sup>a</sup>Universitat de València, Av. de Blasco Ibáñez, 13, 46010 València, Spain

<sup>b</sup>Universitat Politècnica de València, Camino de Vera sn, 46022 València, Spain

---

### Abstract

Sleep assessment is a fundamental part of health evaluation. In fact, many diseases such as obesity, diabetes, or hypertension, as well as psychiatric, neurological, and cardiovascular diseases produce sleep disorders that are often used as indicators, diagnosis symptoms, or even as predictors, e.g., for depression. For this reason, many efforts have been devoted to design methods to control and report about the sleep quality. Two of the most used sleep assessment tools are sleep questionnaires and sleep diaries. Both methods have a very low cost, they are easy to administer, they do not require a sleep centre (such as, e.g., polysomnography), they can be self-administered, and most important, their accuracy is relatively high, as it has been shown in recent studies. In this survey, we systematically revise and compare these tools. We examine the evolution over time of sleep questionnaires and diaries, comparing also their structure and usage. We also review the validation studies and comparatives done in previous studies. This allows us to compare the relative sensitivity and specificity of these methods. Modern sleep diaries come in the form of an app. Therefore, we also present the most advanced and used apps, and discuss their advantages over classical paper diaries.

Keywords: Sleep questionnaires; sleep diaries; sleep detection methods; Sleep quality assessment

---

---

\* Corresponding author. Tel.: +34-657-022-083; fax: +34-963-877-359.

E-mail address: [vanessa.ibanez@ucv.es](mailto:vanessa.ibanez@ucv.es).

## 1. Introduction

Health care systems invest a lot of money into the creation of sleep centres and sleep units in hospitals to evaluate the sleep. Two of the most important methods to assess the sleep are sleep questionnaires and sleep diaries. Along the years, there have appeared many different tools of this kind, but it is still controversial to decide what sleep questionnaire to use. In fact, there is a live debate that has promoted several validation and comparison studies of questionnaires, and even the design of different variations of the same questionnaire. Specially in the last years, new methods have emerged with the appearance of new technologies such as mobile apps.

In this survey, we systematically review the state of the art related to sleep questionnaires and sleep diaries. This review is not limited to the presentation of sleep assessment tools, but also to the evaluation and comparison of these tools. The main criticism about sleep questionnaires and diaries is that they are mostly subjective, and this can affect their accuracy and reliability compared with other objective methods such as polysomnography. However, we present and discuss different studies that report relatively high sensitivity and specificity (up to 97.6% sensitivity and 96% specificity). But even if the precision is high, there is a fundamental limitation of diaries and questionnaires: they are unable to distinguish all the sleep stages (*Awake / N1 / N2 / N3 / REM*), being limited to just reporting about awake or asleep states. Therefore, it is important to contextualize the usefulness of these tools, to profit their specific advantages, and to complement them with other methods.

The main motivation of this survey is to produce an up to date classification of sleep assessment methods based on questionnaires and diaries, and to include in the classification the new technologies such as mobile apps. There have been in the past different reviews of sleep assessment methods, but most of them are outdated (see, e.g., [Lomeli H.A., 2008; Jessica M.K., 2012; C. Winter, 2014]), or they are partial, or only focus on a specific subset of methods (e.g., sleep questionnaires [Silva G.E., 2011; El-Sayed I.H., 2012; Firat H., 2012; Pataka A., 2014; Adams R., 2015; Singh J. and Mims N., 2015], mobile apps [Lee J. and Finkelstein J., 2015; Ong A.A. and Gillespie M.B., 2016], contact sleep detection methods [Kolla B.P. et al., 2016; Green E., 2017; Maslakovic M., 2017], etc.).

The survey has been written to broaden the range of people that may be interested in. For each sleep detection method, it provides a table with the most used concrete public and market products. All technical terms have been conveniently introduced and explained. In all cases, explanations are supported and complemented with adequate references.

The rest of the paper has been structured as follows: First, in Section 2, we present the state of the art associated with sleep questionnaires (Section 2.1), their validation and comparison (Section 2.2), and sleep diaries (Section 2.3). Then, in Section 3, we run a discussion about the accuracy and validity of the methods presented. Finally, in Section 4 we conclude.

## 2. Sleep Assessment with Questionnaires and Diaries

The most advanced method for the sleep assessment in terms of accuracy is the *polysomnogram* (PSG) [Robertson B. et al., 2014, Pandi-Perumal S.R. et al., 2014]. It monitors different body functions during sleep, which are recorded for their later study using different channels. Among the variables studied we have heart rate and rhythm, brain activity, breathing rate and rhythm, flow level of air through the mouth and the nose, eye movement, snoring oxygen and carbon dioxide blood levels, body muscle movements, chest and belly movement, etc. A PSG uses advanced technology such as electroencephalograms, electrocardiograms, etc. that cannot be used at home. This advantage make these methods to be extremely precise, and able to distinguish between, e.g., sleep phases. For this reason, these methods have been often used as the gold standard for sleep evaluation (see, e.g., [Silva G.E. et al., 2011, El-Sayed I.H., 2012, Firat H. et al., 2012, Luo J. et al., 2014, Adams R. et al., 2015, Silva G.E. et al., 2016]). Of course, the use of this exclusive technology comes with a

cost: these methods are expensive, time-consuming, require professional assistance and, often, they can only be done for a reduced period of time (e.g., one or two days). But, additionally, there is another important functional disadvantage: The assessment made by these methods is done in a context that is not the usual sleep context of the patient (i.e., a sleep clinic or a hospital) and, thus, not a normal sleep situation is measured. These problems are often solved by administering a sleep questionnaire and or diary.

### 2.1. Sleep Questionnaires

One of the most used tools for the preliminary evaluation of sleep has traditionally been and it still continues to be the questionnaires. Questionnaires are often the first diagnostic test used in primary care, and they provide a general (quantitative) measure about the subjective quality of sleep. Of course, because questionnaires are based on subjective reports, they can be influenced by the same sources of bias and inaccuracy as any other such reports. Even though this fact, many questionnaires have been validated through large statistical studies and they are a commonly used tool in almost all sleep centres.

The main advantage of questionnaires is that they do not require any device or any sleep context of the patient. They are fast, they can be done at any moment and, normally, they do not require professional assistance. In fact, questionnaires are also accessible to the general public through web systems (see, e.g., the Sleep Disorders questionnaires at <http://skylinefamilypractice.net/IMH/IMHSFP.htm>) or mobile apps (e.g., the FOSQ questionnaire app, publicly available at Google play). A self-administered questionnaire may help to alert people (e.g., with sleep apnea) about the need for proper diagnosis.

In this section, we review and compare the most extended and influential questionnaires used along the last 30 years. There are hundreds of other different questionnaires to evaluate sleep. In fact, many hospitals and sleep centres have developed their own, or have parameterized one of the questionnaires presented here. The most important questionnaires are presented in Table 1. We have collected all of them and made them publicly available for the first time at: <http://www.dsic.upv.es/~jsilva/sleep/>

*Pittsburgh Sleep Quality Index (PSQI)* [Buysse D.J. et al., 1989]: It is used to measure the quality and patterns of sleep in adults. It differentiates “poor” from “good” sleep quality by measuring seven areas: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medications, and daytime dysfunction over the last month.

*Epworth Sleepiness Scale (ESS)* [Johns M.W., 1991]: It is a self-administered questionnaire based on retrospective reports of the likelihood of dozing off or falling asleep in a variety of different situations. Respondents are asked to rate their usual chances of dozing off or falling asleep while engaged in eight different activities. The ESS measures a general level of daytime sleepiness, or their average sleep propensity in daily life.

*Insomnia Severity Index (ISI)* [Morin C.M., 1993]: It was designed to assess the nature, severity, and impact of insomnia; and to monitor treatment response in adults. It has seven questions. The seven answers are added up to get a total score.

*Sleep Disorders Questionnaire (SDQ)* [Douglass A.B. et al., 1994]: It is a long questionnaire with 175 items that measures sleep disturbance and usual sleep habits during the past month only. It was designed as a tool for the identification of those at high risk for possessing a sleep disorder. The developers also created a smaller, 45-item version, to assess four common sleep disorders: sleep apnea, narcolepsy, psychiatric sleep disorders, and periodic limb movement disorder.

Table 1. Questionnaires for the detection of sleep disorders

Sleep Questionnaire	Year	City	Reference	Structure
<i>Pittsburgh Sleep Quality Index (PSQI)</i>	1989	Pittsburgh (USA)	Buysse D.J. et al., 1989	24 items 4 points scale
<i>Epworth Sleepiness Scale (ESS)</i>	1991	Melbourne (Australia)	Johns M.W., 1991	8 items 4 points scale
<i>Insomnia Severity Index (ISI)</i>	1993	Quebec (Canada)	Morin C.M., 1993	7 items 5-points scale
<i>Sleep Disorders Questionnaire (SDQ)</i>	1994	Michigan (USA)	Douglass A.B. et al., 1994	175 items 5-points scale
<i>Sleep apnea clinical score (SACS)</i>	1994	Alberta (Canada)	Flemons W.W. et al., 1994	4 items 100-point scale
<i>Functional Outcomes of Sleep Questionnaire (FOSQ)</i>	1997	Philadelphia (USA)	Weaver T.E. et al., 1997	30 items 4-5 points scale
<i>Calgary Sleep Apnea Quality of Life Index (SAQLI)</i>	1998	Calgary (Canada)	Flemons W.W., Reimer M.A., 1998	35 items 7-point scale
<i>Oviedo Sleep Questionnaire (OSQ)</i>	1998	Oviedo (USA)	Bobes J. et al., 1998	15 items 4-7 points scale
<i>Berlin Questionnaire (BQ)</i>	1999	Berlin (Germany)	Netzer N.C. et al., 1999	10 items 2-5 points scale
<i>Athens Sleep Questionnaire (ASQ)</i>	2000	Athens (Greece)	Soldatos C.R. et al., 2000	8 items 4-point scale
<i>Self-efficacy in Sleep apnea (SEMSA)</i>	2003	Philadelphia (USA)	Weaver T.E. et al., 2003	26 items 4-point scale
<i>STOP Questionnaire (SQ)</i>	2008	Toronto (Canada)	Chung, F. et al., 2008	4 items 2 points scale
<i>STOP-BANG Questionnaire (SBQ)</i>	2008	Toronto (Canada)	Chung, F. et al., 2008	8 items 2 points scale
<i>Bergen Insomnia Scale (BIS)</i>	2008	Bergen (Norway)	Pallesen S. et al., 2008	6 items 8-point scale
<i>FOSQ-10</i>	2009	Pittsburgh (USA)	Chasens, E.R. et al., 2009	10 items 4 points scale
<i>Simple Four Variables (SFV)</i>	2009	Kyoto (Japan)	Takegami M. et al., 2009	4 items 2-6 points scale
<i>OSA50</i>	2011	Adelaida (Australia)	Chai-Coetzer C.L. et al., 2011	4 items 3-4 points scale

*Sleep apnea clinical score (SACS)* [Flemons W.W. et al., 1994]: It incorporates information on neck circumference, hypertension, habitual snoring, and nocturnal gasping or choking to generate a score ranging from 0 to 100.

*Functional Outcomes of Sleep Questionnaire (FOSQ)* [Weaver T.E. et al., 1997]: It is a disease specific quality of life questionnaire to determine functional status in adults. Measures are designed to assess the impact of disorders of excessive sleepiness on multiple activities of everyday living and the extent to which these abilities are improved by effective treatment.

*Calgary Sleep Apnea Quality of Life Index (SAQLI)* [Flemons W.W., Reimer M.A., 1998]: It is a condition-specific health-related quality of life measure designed as an interview-administered scale. It evaluates four domains of quality of life associated with sleep apnea: daily functioning, social interactions, emotional functioning, and symptoms. A fifth domain—treatment-related symptoms—was developed for use with individuals currently undergoing therapeutic intervention.

*Oviedo Sleep Questionnaire (OSQ)* [Bobes J. et al., 1998]: It is a semi-structured interview for making the DSM-IV and ED-10 diagnoses of insomnia and hypersomnia in the last month. Initially designed to target depressed patients, it contains three subscales: an analogic-visual scale concerning subjective satisfaction with sleep, a scale concerning insomnia, and a scale concerning hypersomnia. It also provides additional information to the clinician (parasomnias and use of some kind of help for sleeping).

*Berlin Questionnaire (BQ)* [Netzer N.C. et al., 1999]: It is a sleep apnea screening questionnaire used to quickly identify the risk (low to high) of sleep disordered breathing. It consists of 10 items relating to snoring, nonrestorative sleep, sleepiness while driving, apneas during sleep, hypertension, and body mass index.

*Athens Sleep Questionnaire (ASQ)* [Soldatos C.R. et al., 2000]: This scale is intended to record the own patient's assessment of any sleep difficulty they have experienced. It is measured by assessing eight factors amongst which first five factors are related to nocturnal sleep and three factors are related to daytime dysfunction.

*Self-efficacy in Sleep apnea (SEMSA)* [Weaver T.E. et al., 2003]: It provides a disease-specific measure of pre-treatment expectancies regarding obstructive sleep apnea and continuous positive airway pressure treatment in adults. It was designed to assess adherence-related cognitions.

*STOP Questionnaire (SQ)* [Chung, F. et al., 2008]: STOP stands for Snoring, Tiredness, Observed apnea and high blood Pressure. It is a screening tool for obstructive sleep apnea.

*STOP-BANG Questionnaire (SBQ)* [Chung, F. et al., 2008]: BANG stands for Body mass index, Age, Neck circumference, Gender. It extends SQ to incorporate information about these four demographics items.

*Bergen Insomnia Scale (BIS)* [Pallesen S. et al., 2008]: It is composed of six items, of which the first three pertain to sleep onset, maintenance, and early morning wakening insomnia. The last three items refer to not feeling adequately rested, experiencing daytime impairment, and being dissatisfied with current sleep.

*FOSQ-10* [Chasens, E.R. et al., 2009]: The own authors of FOSQ noticed that, in some situations, it contains too many questions (30). Hence, they developed a shorter version with only 10 questions (thus, named FOSQ-10) that may be more easily implemented in clinical practice. It is used in research and clinical practice to measure the impact of daytime sleepiness on activities of daily living.

*Simple Four Variables (SFV)* [Takegami M. et al., 2009]: It is composed of four questions about sex, blood pressure level, body mass index, and self-reported snoring. Its objective was to produce a simple questionnaire that used information routinely checked during daily clinical practice.

*OSA50* [Chai-Coetzer C.L. et al., 2011]: OSA50 stands for Obesity, Snoring, Apneas, aged over 50. It was designed to address the growing burden of disease and long waiting lists for sleep services. It is used in primary care consisting of an easy to administer screening questionnaire with four questions whose score ranges from 0 to 10.

The above list does not pretend to be exhaustive. We have omitted many questionnaires that are very restricted due to copyright, that are less extended, or that, although they are more or less extended, they have

not been validated. This is for instance the case of the *Graduated Apnea Screening Protocol (GASP)*. The questionnaires selected are the most important scales and have been validated with studies in different countries.

## 2.2. Comparison and evaluation of questionnaires

The question whether one scale is better than other emerges naturally. Clearly, having less questions and a smaller range of answers makes a questionnaire easier to administer, but it reduces the amount of information collected, and thus its effectivity. When comparing two questionnaires, there are different statistical values that can be used (for instance, the area under the ROC curve), but there are two that are especially useful:

- *Sensitivity* (true positive rate). It is calculated as the number of identified positives divided by the total number of positives (e.g., the percentage of sick people who are correctly identified as sick).
- *Specificity* (true negative rate). It is calculated as the number of identified negatives divided by the total number of negatives (e.g., the percentage of healthy people who are correctly identified as healthy).

Clearly, these variables are confronted. For instance, a method that always classifies as positive (e.g., having a sleep disease), will have a 100% sensitivity and a 0% specificity, and the opposite happens with a method that always classifies as negative. Often, identification methods have to reach a compromise between both variables. In general, high specificities may be more useful in excluding low-risk patients, while avoiding false positives. However, sleep clinicians may prefer to use screening tools with high sensitivities in order to avoid missing cases that may lead to adverse health consequences and increased healthcare costs.

There have been different studies to compare some of the questionnaires proposed. These studies usually compare questionnaires with a similar length, such as for instance, those questionnaires with only 4 items (e.g., SQ, SFV, OSA50, etc.). The most relevant comparisons are the following:

*ESS vs. SQ vs. SBQ vs. SFV [Silva G.E. et al., 2011]*: This study compared four different questionnaires for screening obstructive sleep apnea (OSA): STOP-Bang questionnaire, Epworth sleepiness scales, Berlin questionnaire, STOP questionnaire, and Simple Four Variable. 4,770 patients were asked to fill in the questionnaires before completing a polysomnogram, which was used as gold standard. The SBQ had higher sensitivity to predict moderate-to-severe (87.0%) and severe (70.4%) sleep-disordered breathing (SDB), while the SFV had higher specificity to predict moderate-to-severe and severe SDB (93.2% for both).

*ESS vs. BQ vs. SQ vs. SBQ [El-Sayed I.H., 2012]*: This study compared four different questionnaires for screening OSA: STOP-Bang questionnaire, Epworth sleepiness scales, Berlin questionnaire, STOP questionnaire. 234 patients were asked to fill in the SBQ, ESS, BQ, and SQ questionnaires before overnight polysomnography. The questionnaires were compared using PSG as gold standard. After statistical analysis, they concluded that the SBQ, BQ, and SQ had the highest sensitivity to predict OSA (97.55%, 95.07%, and 91.67%, respectively), moderate-to-severe OSA (97.74%, 95.48%, and 94.35%, respectively) and severe OSA (98.65%, 97.3%, and 95.95%, respectively), but with a very low specificity for OSA patients (26.32%, 25%, and 25%, respectively), moderate-to-severe OSA patients (3.7%, 7.41%, and 25.93%, respectively) and severe OSA patients (5.36%, 10.71%, and 19.64%, respectively), while the ESS had the highest specificity to predict OSA, moderate-to-severe OSA and severe OSA (75%, 48.15%, and 46.43%, respectively) but with the lowest sensitivity (72.55%, 75.71%, and 79.73%, respectively). Hence, the sensitivity of BQ, SQ, and SBQ was very high yet, their low specificity results in increased false positives and failure of exclusion of individuals at low risk.

Table 2. Studies that compare sleep assessment questionnaires

Study	Sample	Questionnaires evaluated	Results
<i>Silva G.E. et al., 2011</i>	4770	<i>ESS vs. SQ vs. SBQ vs. SFV</i>	Best sensitivity: SBQ (87.0%) Best specificity: SFV (93.2%)
<i>El-Sayed I.H., 2012</i>	234	<i>ESS vs. BQ vs. SQ vs. SBQ</i>	Best sensitivity: SBQ (97.55%) Best specificity: ESS (75.0%)
<i>Firat H. et al., 2012</i>	90	<i>BQ vs. SQ vs. SBQ vs. OSA50</i>	Best sensitivity: SBQ (87%) Best specificity: SBQ (76.0%)
<i>Luo J. et al., 2014</i>	212	<i>ESS vs. BQ vs. SQ vs. SBQ</i>	Best sensitivity: SBQ (94.9%) Best specificity: SFV (50.0%)
<i>Pataka A. et al., 2014</i>	1853	<i>ESS vs. BQ vs. SQ vs. SBQ vs. SFV</i>	Best sensitivity: SB (97.6%) Best specificity: SFV (74.4%)
<i>Adams R. et al., 2015</i>	543	<i>SQ vs. SBQ vs. OSA50</i>	Best sensitivity: OSA50 + oximetry (73.0%) Best specificity: OSA50 + oximetry (96.0%)
<i>Silva G.E. et al., 2016</i>	884	<i>FOSQ vs. SAQLI</i>	Best sensitivity: Not reported Best specificity: Not reported

*BQ vs. SQ vs. SBQ vs. OSA50 [Firat H. et al., 2012]:* This study compared the STOP-Bang questionnaire, Berlin questionnaire, STOP questionnaire, and OSA50 in the context of high-risk bus drivers for OSA. The sample was composed of 90 highway bus drivers that underwent polysomnography after they completed the four questionnaires. SBQ had the highest sensitivity (87%) and the highest specificity (76%) in identifying high-risk for OSA.

*ESS vs. BQ vs. SQ vs. SBQ [Luo J. et al., 2014]:* This study aimed to compare four questionnaires in screening obstructive sleep apnea hypopnea syndrome (OSAHS) patients: STOP-Bang questionnaire, Epworth sleepiness scales, Berlin questionnaire, and STOP questionnaire. 212 patients were asked to fill in the SBQ, ESS, BQ, and SQ questionnaires before overnight polysomnography. The questionnaires were compared using PSG as gold standard. After statistical analysis, they concluded that SBQ is superior to ESS, BQ, and SQ. In particular, the sensitivities of SBQ score  $\geq 3$  with apnea hypopnea index (AHI)  $\geq 5/h$ , AHI  $\geq 15/h$ , and AHI  $\geq 30/h$  as gold standards were 94.9%, 96.5%, and 97.7%, respectively. The specificities were 50.0%, 28.6%, and 17.9%, respectively.

*ESS vs. BQ vs. SQ vs. SBQ vs. SFV [Pataka A. et al., 2014]:* This study compared five different questionnaires for assessing OSA: STOP-Bang questionnaire, Epworth sleepiness scales, Berlin questionnaire, STOP questionnaire, and Simple Four Variables. The main aim was to combine these questionnaires to produce an improved one. The sample was composed of 1853 (74.4% male) patients. After statistical analysis, they concluded that SB had the highest sensitivity (97.6%), but the lowest specificity (12.7%) for AHI  $\geq 15$ . The SFV  $\geq 14$  had the highest specificity (74.4%) followed by ESS (67%). BQ had good sensitivity (87%), worse specificity (33%) than SFV and ESS but better than SQ (13%) and SB (12.7%). The combination of questionnaires did not improve their predictive value.

*SQ vs. SBQ vs. OSA50 [Adams R. et al., 2015]:* This study compared the STOP-Bang questionnaire, STOP questionnaire, and OSA50. The sample was composed of 543 participants that underwent polysomnography after they completed the three questionnaires. The study concluded that there were no significant differences between the predictions of the three questionnaires, being OSA50 the one that produced a slightly higher sensitivity. The best results were achieved with a two-step diagnostic model: OSA50 followed by oximetry, which produced a 73% sensitivity and 96% specificity.



*FOSQ vs. SAQLI [Silva G.E. et al., 2016]*: The study compared the Calgary Sleep Apnea Quality of Life Index (SAQLI) and the Functional Outcomes Sleep Questionnaire (FOSQ) to assess the quality of life index (QoL). The sample was formed of 884 participants that completed the FOSQ and SAQLI, and in-home polysomnograms. When analyzed according to gender, no significant differences were found in males. However, females with severe OSA had significantly lower mean scores for the SAQLI and FOSQ compared to females with no OSA and those with mild to moderate OSA. Hence, the impact of OSA on QoL differs between genders with a larger effect on females. The FOSQ and SAQLI were correlated producing a correlation  $r=.66$ ,  $p<.001$ .

All the above studies are summarized in Table 2.

### 2.3. Sleep Diaries

A sleep diary is a tool that allows one to self-assess the sleep. Sleep diaries are easy to use and only take some minutes each day to complete. The importance of this tool as a source of information regarding the overall conditions of sleep have motivated the creation of many different sleep diaries. Almost all of them coincide in the following information that should be gathered every day:

- Time to go to bed and time to go out of the bed
- How long taken to fall asleep
- Number of hours slept
- Number of awakenings and their times, durations, and causes (e.g., nightmare, noise...)
- Feeling before and after sleep
- Activities before going to sleep
- Presence in the day of a nap, exercise, medication, caffeine, or alcohol

However, more exhaustive or specific sleep diaries can also contain some of the following information:

- Time the patient intended to wake up, and time when he actually woke up
- Cause to woke up (e.g., spontaneously, an alarm clock, or another (specified) disturbance)
- Short explanation about how the person felt during the day (often on a scale from 1 to 5)
- Time and type/heaviness of evening meal
- Level of stress before bedtime (often on a scale from 1 to 5) and the cause
- Time the patient tried to fall asleep and time the patient thinks he actually fell asleep
- Activity during the aforementioned two moments (meditation, eyes closed...)
- Subjective perception of the quality of sleep (often on a scale from 1 to 5) and the cause
- Short explanation about any recalled good or bad dreams

This information can help the own patient or their health care providers to evaluate the sleep, the patient's bedroom environment, sleep habits, etc.; and to identify patterns and practices that foster—or hinder—a good sleep. In Table 3, we list some of the most representative sleep diaries. We have collected all of them and made them publicly available at: <http://www.dsic.upv.es/~jsilva/sleep/>

Sleep diaries can also be administered in the form of an app. In fact, a recent study compared the use of paper and electronic sleep diaries [Tonetti L. et al., 2016] and concluded that their diagnosis power is similar, but also that electronic sleep diaries provide several benefits in comparison to the paper sleep diary:

- Reducing the time for data entry
- Automatic scoring
- Avoiding the “parking lot syndrome” (patients retrospectively complete several days at the same time)

- Automatic recording of the time when the diary was filled in (e.g., immediately after waking up)

In Table 4, we list the main sleep diary apps according to the number of reviews in App Store and Google Play.

Table 3. Sleep diaries for the detection of sleep problems

Sleep Diary	Number of Questions				Scale
	Before sleep	About sleep	After sleep	Total	
<i>Pittsburgh Sleep Diary (PSD)</i>	12	9	2	23	6 points scale
<i>Consensus Sleep Diary (CSD)</i>	10	7	3	20	5 points scale
<i>National Sleep Foundation (NSF)</i>	10	3	2	15	3 points scale
<i>Get Self Help Sleep Diary (GSH)</i>	10	2	2	14	11 points scale
<i>National Heart, Lung, and Blood Institute (NHLBI)</i>	7	4	1	12	3-4 points scale
<i>NPS MedicineWise Sleep Diary (NPS)</i>	7	2	2	11	3 points scale
<i>Loughborough Sleep Research Center (LSRC)</i>	2	5	1	8	5 points scale

Table 4. Sleep Diary apps (prices and reviews are taken from Google Play)

App Name	Developer	Price	Average review	Number of reviews
<i>Sleep Diary Pro</i>	Froggyware	2.49\$	4.2 out of 5	421
<i>Healthy Sleep Diary</i>	Fruct	0\$	3.9 out of 5	209
<i>Sleep Diary Lite</i>	Froggyware	0\$	3.8 out of 5	2,249

The oldest sleep diary we are aware of is the Pittsburgh Sleep Diary [Monk T.H. et al., 1994]. Since it was proposed, hundreds of other sleep diaries have appeared—only in our repository, we have dozens of sleep diaries from different organizations, hospitals, and sleep centres—. Due to the existence of so many different sleep diaries, comparisons across studies is often difficult and unfair because their scales can vary a lot. To solve this problem, the experts that attended the Pittsburgh Assessment Conference and other insomnia researchers studied and compared a collection of sleep diaries with the aim of proposing a new improved sleep diary. The result was the “Consensus Sleep Diary” [C.E. Carney et al., 2012] (see Table 3).

### 3. A critical view of accuracy and validation

Each sleep detection method has its own reliability and precision. While both, questionnaires and sleep diaries, are mostly subjective, they have an essential difference in the way that the data is collected. Questionnaires are filled in just one moment, and thus (i) its information depends on the memory of the patient about his sleep over a week or a month; and (ii) it summarizes information about a period, losing the variability among different days. In contrast, sleep diaries (i) record the subjective feeling of the patient every day, and (ii) the sleep is described just after waking up, thus potentially being more accurate and not depending on the patient’s memory. This makes the amount of information and the accuracy of sleep diaries to be objectively superior to that of questionnaires.

There exist several studies [Silva G.E. et al., 2011, El-Sayed I.H., 2012, Firat H. et al., 2012, Luo J. et al., 2014, Pataka A. et al., 2014, Adams R. et al., 2015, Silva G.E. et al., 2016] devoted to compare the accuracy of different questionnaires (see Section 2.1 for a detailed comparison and evaluation of sleep questionnaires). All of them use PSG as the gold standard. Depending on the considered study (they use samples targeting different

population kinds), sensitivity varies between 73.0% and 97.6%, while specificity varies between 50% and 96% (see Table 2).

Sleep diaries have also been evaluated in [Jungquist C.R. et al., 2015] and [Tonetti L. et al., 2016]. Both studies (see Table 5 for details) compare the accuracy of paper diaries against electronic diaries, and they both used actigraphy as the gold standard. Their conclusion is very similar: paper and electronic diaries collect statistically the same data. However, they coincide to highlight some valuable advantages of electronic diaries such as improved data collection and analysis and the ability to time-stamp the entry. In the comparison of sleep diaries against actigraphy there is a poor agreement between estimates of sleep quantity and sleep quality (specially about the number of awakenings, where the subjective patients' perception of awakenings is significantly lower).

Table 5. Studies that compare the accuracy of electronic sleep diaries and paper sleep diaries

Reference	Paper diary	Electronic Diary	Actigraph	Period	Sample
Jungquist C.R. et al., 2015	Ad hoc diary	PRO-diary (wrist device)	PRO-diary	14 days	35
Tonetti L. et al., 2016	Consensus Sleep Diary	Ad hoc app (tablet)	Actiwatch AW64	7 days	15

#### 4. Conclusions

Sleep is fundamental to life, and its assessment provides essential information about health. In this survey, we have reviewed and classified two of the more important methods to detect the sleep: sleep questionnaires, sleep diaries. For each method, we have reviewed, compared and discussed the state of the art (both, the literature and the current state of the practice, providing up-to-date reviews of apps).

One conclusion of this study is that all sleep assessment methods have advantages and disadvantages, thus, they should be combined and adapted to the specific needs. For instance, PSG is the method that provides the more accurate and complete information about sleep. It even allows us to precisely identify the sleep stages. But it is expensive, and requires special hardware and medical assistance. Hence, it can only be administered for very small periods of time, and the sleep is assessed out of the normal context (e.g., a hospital). When patients are hooked up to various machines with up to 20 electrodes placed on their body and head, and in a lab with cameras rather than their own bed, they understandably might not sleep normally.

On the opposite side, we have sleep diaries and questionnaires. They are mostly subjective and provide discretized approximations to the real sleep quality. But they report about the feeling of the patient about his own sleep, they are cheap, can be administered at home for long periods, and contain useful complementary information about, e.g., habits.

Given the inherent subjectivity of the sleep questionnaires, one could think that they are unreliable. Contrarily, we have shown (see Table 2) that their sensitivity is high (often above 90%, and between 73% and 97.7% in all the discussed studies). Specificity is not so high, ranging between 50% and 96%.

#### Abbreviations

ASQ: Athens sleep questionnaire; AWI: awakening index; BQ: Berlin questionnaire; BIS: Bergen insomnia scale; CSD: consensus sleep diary; ESS: Epworth sleepiness scale; FOSQ: functional outcomes of sleep questionnaire; GASP: graduated apnea screening protocol; GSH: get self help sleep diary; ISI: insomnia severity index; LSRC: Loughborough sleep research center; NHLBI: national heart, lung, and blood institute;

NPS: MedicineWise sleep diary; NSF: national sleep foundation; OSA50: obesity, snoring, apneas, aged over 50; OSAHS: sleep apnea hypopnea syndrome; OSQ: Oviedo sleep questionnaire; PSD: Pittsburgh sleep diary; PSG: polysomnogram; PSQI: Pittsburgh sleep quality index; QoL: quality of life index; ROC: receiver operating characteristic; SACS: sleep apnea clinical score; SAQLI: Calgary sleep apnea quality of life index; SBQ: STOP-BANG questionnaire; SDB: sleep-disordered breathing; SDQ: sleep disorders questionnaire; SEMSA: self-efficacy in sleep apnea; SFV: simple four variables; SQ: STOP questionnaire.

## References

- Adams R., Appleton S., Vakulin A., Martin S.A., Grant J.F., Antic N.A., Catcheside P., Taylor A.W., Wittert G.A., McEvoy R.D., 2015. Comparison of the OSA50, STOP and STOP-BANG Questionnaires for Identification of Obstructive Sleep Apnea in a Male Population Cohort. *American Journal of Respiratory and Critical Care Medicine*, 2015, 191:A5601.
- Bobes J., González M.P., Vallejo J., Sáiz J., Gibert J., Ayuso J.L., Rico F., 1998. Oviedo Sleep Questionnaire (OSQ): A new semistructured Interview for sleep disorders. *European Neuropsychopharmacology*, 1998 November, 8(2):S162.
- Buysse D.J., Reynolds C.F., Monk T.H., Berman S.R., Kupfer D.J., 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.*, 1989 May, 28(2):193-213.
- Carney C.E., Buysse D.J., Ancoli-Israel S., Edinger J.D., Krystal A.D., Lichstein K.L., Morin C.M., 2012. The Consensus Sleep Diary: Standardizing Prospective Sleep Self-Monitoring. *Sleep*. 2012 February; 35(2):287-302.
- Chai-Coetzer C.L., Antic N.A., Rowland L.S., Catcheside P.G., Esterman A., Reed R.L., Williams H., Dunn S., McEvoy R.D., 2011. A simplified model of screening questionnaire and home monitoring for obstructive sleep apnea in primary care. *Thorax*, 2011 March, 66(3):213-9.
- Chasens, E. R., Ratcliffe, S. J., Weaver, T. E., 2009. Development of the FOSQ-10: a short version of the functional outcomes of sleep questionnaire. *Sleep*, 32(7):915-919.
- Chung, F., Yegneswaran, B., Liao, P., Chung, S., Vairavanathan, S., Islam, S., Khajehdehi, A., Shapiro C., 2008. STOP questionnaire. A tool to screen patients for obstructive sleep apnea. *Anesthesiology*, 108(5):812-21.
- Chung F., Subramanyam R., Liao P., Sasaki E., Shapiro C. and Sun Y., 2012. High STOP-Bang score indicates a high probability of obstructive sleep apnea. *British Journal of Anesthesia*, 108(5):768-75.
- Dougllass A.B., Bornstein R., Nino-Murcia G., Keenan S., Miles L., Zarcone V.P. Jr, Guilleminault C., Dement W.C., 1994. The Sleep Disorders Questionnaire. I: Creation and multivariate structure of SDQ. *Sleep*. 1994 March, 17(2):160-7.
- El-Sayed I.H., 2012. Comparison of four sleep questionnaires for screening obstructive sleep apnea. *Egyptian Journal of Chest Diseases and Tuberculosis*, 2012 October, 61(4): 433-441.
- Firat H., Yuceege M., Demir A., Ardic S., 2012. Comparison of four established questionnaires to identify highway bus drivers at risk for obstructive sleep apnea in Turkey. *Sleep and Biological Rhythms*. 2012 July, 10(3): 231-236.
- Flemons W.W., Whitelaw W.A., Brant R., Remmers J.E., 1994. Likelihood ratios for a sleep apnea clinical prediction rule. *American Journal of Respiratory and Critical Care Medicine*, 1994 November, 150(5):1279-1285.
- Flemons W.W., Reimer M.A., 1998. Development of a disease-specific health-related quality of life questionnaire for sleep apnea. *American Journal of Respiratory and Critical Care Medicine*, 1998 August, 158(2):494-503.
- Green E., 2017. The 10 Best Sleep Trackers in 2017. [www.nosleeplessnights.com](http://www.nosleeplessnights.com/best-sleep-tracker/), 2017 January 6, <http://www.nosleeplessnights.com/best-sleep-tracker/>
- Jessica M.K., Strecker R.E., Bianchi M.T., 2012. Recent developments in home sleep-monitoring devices. *ISRN Neurology*, 2012:768794.
- Johns M.W., 1991. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*, 1991 December, 14(6):540-5.
- Jungquist C.R., Pender J.J., Klingman K.J., Mund J., 2015. Validation of Capturing Sleep Diary Data via a Wrist-Worn Device. *Sleep Disorders*, 2015, Article ID 758937.
- Kolla B.P., Mansukhani S., Mansukhani M.P., 2016. Consumer sleep tracking devices: a review of mechanisms, validity and utility. *Expert Review of Medical Devices*. 2016 May, 13(5):497-506.
- Lomeli H.A., Pérez-Olmos I., Talero-Gutiérrez C., Moreno C.B., González-Reyes R., Palacios L., de la Peña F., Muñoz-Delgado J., 2008. Sleep evaluation scales and questionnaires: a review. *Actas Españolas de Psiquiatría*. 2008 February;36(1):50-9.
- Luo J., Huang R., Zhong X., Xiao Y., Zhou J., 2014. STOP-Bang questionnaire is superior to Epworth sleepiness scales, Berlin questionnaire, and STOP questionnaire in screening obstructive sleep apnea hypopnea syndrome patients. *Chinese Medical Journal*, 2014, 127(17):3065-70.
- Maslakovic M., 2017. Ten gadgets for advanced sleep monitoring. [www.gadgetsandwearables.com](http://gadgetsandwearables.com/2017/01/09/the-best-sleep-trackers/), 2017 January 9, <http://gadgetsandwearables.com/2017/01/09/the-best-sleep-trackers/>
- Monk T.H., Reynolds C.F., Kupfer D.J., Buysse D.J., Coble P.A., Hayes A.J., MacHen M.A., Petrie S.R., Ritenour A.M., 1994. The Pittsburgh Sleep Diary. *Journal of Sleep Research*, 1994 June, 3(2):111-120.
- Morin C.M., 1993. *Insomnia: psychological assessment and management*. New York: Guilford Press, 1993.
- Netzer N.C., Stoohs R.A., Netzer C.M., Clark K., Strohl K.P., 1999. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. *Annals of internal medicine*, 1999, 131(7):485-91.

- Ong A.A., Gillespie M.B., 2016. Overview of smartphone applications for sleep analysis. *World Journal of Otorhinolaryngology-Head and Neck Surgery*, 2016 March, 2(1):45–49.
- Robertson B., Marshall B., Carno M.A., 2014. *Polysomnography for the Sleep Technologist*. Elsevier, ISBN: 9780323100199.
- Pallesen S., Bjorvatn B., Nordhus I.H., Sivertsen B., Hjørnevik M., Morin C.M., 2008. A new scale for measuring insomnia: the Bergen Insomnia Scale. *Perceptual and Motor Skills*. 2008 December, 107(3):691-706.
- Pandi-Perumal S.R., Spence D.W., BaHammam A.S., 2014. Polysomnography: An Overview. *Primary Care Sleep Medicine*, 2014 July, 29-42.
- Pataka A., Daskalopoulou E., Kalamaras G., Fekete Passa K., Argyropoulou P., 2014. Evaluation of five different questionnaires for assessing sleep apnea syndrome in a sleep clinic. *Sleep Medicine*, 2014 July, 15(7):776-81.
- Silva G.E., Vana K.D., Goodwin J.L., Sherrill D.L., Quan S.F., 2011. Identification of patients with sleep disordered breathing: comparing the four-variable screening tool, STOP, STOP-Bang, and Epworth Sleepiness Scales. *Journal of Clinical Sleep Medicine*, 2011 October, 7(5):467-72.
- Silva G.E., Goodwin J.L., Vana K.D., Quan S.F., 2016. Obstructive sleep apnea and quality of life: comparison of the SAQLI, FOSQ, and SF-36 questionnaires. *Southwest Journal on Pulmonary Critical Care*. 2016,13(3):137-49.
- Singh J., Mims N., 2015. Screening Tools for the Obstructive Sleep Apnea for the Cardiovascular Clinician. *American College of Cardiology*, 2015 Jul 14.
- Soldatos C.R., Dikeos D.G., Paparrigopoulos T.J., 2000. Athens Insomnia Scale: validation of an instrument based on ICD-10 criteria. *Journal of Psychosomatic Research*, 2000, 48:555-560.
- Takegami M., Hayashino Y., Chin K., Sokejima S., Kadotani H., Akashiba T., Kimura H., Ohi M., Fukuhara S., 2009. Simple four-variable screening tool for identification of patients with sleep-disordered breathing. *Sleep*. 2009 July, 32(7):939-48.
- Tonetti L., Mingozi R., Natale V., 2016. Comparison between paper and electronic sleep diary. *Biological Rhythm Research*, 2016 June, 47(5):743-753.
- Weaver T.E., Laizner A.M., Evans L.K., Maislin G., Chugh D.K., Lyon K., Smith P.L., Schwartz A.R., Redline S., Pack A.I., Dinges D.F., 1997. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*. 1997 Oct, 20(10):835-43.
- Weaver T.E., Maislin G., Dinges D.F., Younger J., Cantor C., McCloskey S., Pack A.I., 2003. Self-efficacy in sleep apnea: instrument development and patient perceptions of obstructive sleep apnea risk, treatment benefit, and volition to use continuous positive airway pressure. *Sleep*, 2003 September, 26(6):727-32.
- Winter C., 2014. Personal Sleep Monitors: Do They Work? *The huffington post*, 2014 April 28.