ORIGINAL ARTICLE

The Berlin questionnaire for sleep apnea in a sleep clinic population: relationship to polysomnographic measurement of respiratory disturbance

Negar Ahmadi · Sharon A. Chung · Alison Gibbs · Colin M. Shapiro

Abstract The Berlin questionnaire (BQ) has been used to help identify patients at high risk of having sleep apnea in primary care and atrial fibrillation patients. The BQ may be a useful adjunct in sleep medicine and research, but it has never been validated in a sleep clinic population. The aim of the study is to determine the specificity and sensitivity of the BQ compared to the respiratory disturbance index (RDI) values obtained from two nights of polysomnographic recording in a sleep clinic population. This is a retrospective chart review study of 130 sleep clinic patients. Patients' demographics, BQ scores, RDI measurements, and sleep study parameters were extracted from the patients' chart. Of the 130 charts reviewed, the BQ identified 76 (58.5%) as being at high-risk of having sleep apnea, but overnight polysomnography found only 34 of the 130 patients (26.2%) had an RDI>10. The BQ performed with 0.62 sensitivity and 0.43 specificity at the RDI>10 level. Due to the low sensitivity and specificity as well as the large number of false negatives and positives, the Berlin questionnaire is not an appropriate instrument for identifying patients with sleep apnea in a sleep clinic population.

Keywords Sleep Apnea · Berlin questionnaire · Excessive daytime sleepiness

N. Ahmadi · S. A. Chung (✉) · C. M. Shapiro
Department of Psychiatry, Toronto Western Hospital,
399 Bathurst Street, 7M-429,
Toronto, ON M5T 2S8, Canada
e-mail: sachung@uhnres.utoronto.ca

A. Gibbs
Department of Statistics, University of Toronto,
Toronto, ON, Canada

Introduction

Sleep apnea affects approximately 2–5% of males and 2% of females [1]. If left untreated, sleep apnea has been shown to result in numerous medical complications and an elevated mortality rate [2–6]. Overnight polysomnography with recording of respiratory variables such as Respiratory Disturbance Index (RDI) and levels of oxygen saturation is the gold standard for the diagnosis of sleep apnea [7]. However, several studies have shown that there exists a significant rate of mistaken diagnosis for sleep apnea based on RDI recordings of only one night. A review of the existing literature on this issue by Stepnowsky et al. [8] has shown that the misclassification occurs in 6 to 54% of patients based on RDI cut-off of 5 and 12% to 16% based on cut-off of 10. Hence, two or more nights of polysomnographic testing would be the ideal diagnostic procedure for detecting sleep apnea. However, the economic pressure of reducing health care spending and the long wait times in a number of sleep clinics dictate otherwise, and the majority of patients receive only one night of polysomnography.

Given the greater chance of a missed diagnosis based on only one night of recording and the need to triage patients who urgently require sleep studies, the clinical interview has been relied upon to identify patients who may have sleep apnea and to help with decision making when the sleep study results are ambiguous. In addition, in the face of rising health care costs and increased pressure for more stringent use of health care resources, it falls to the sleep physician to recommend sleep studies for only those patients at high risk of having a sleep disorder. Furthermore, most sleep clinics have long waiting lists, and the more ill patients, in particular those with severe sleep breathing disorders, warrant quicker diagnosis and treatment. Most patients are sent to sleep physicians with

complaints of poor sleep, daytime sleepiness, and fatigue, but the differential diagnosis could include insomnia, depression [9], and hypothyroidism [10], to name a few. It therefore falls to the physician to determine if their patient is at high risk of having a sleep disorder based solely on the clinical interview. This is particularly the case for sleep apnea. In addition, the sleep clinic population tends to be representative of a population with excessive sleepiness. However, excessive daytime sleepiness is a common complaint both in sleep disordered patients and a number of other nonmedical and medical populations. These include individuals who restrict their sleep due to social or family reasons, shift workers [11] and daytime non-shift workers [12], and patient groups such as those, for example, with neurological disorders [13, 14]. Therefore, a cheap and quick screening tool for sleep apnea is useful for both research purposes and clinical triage.

The purpose of this study is to explore if a screening questionnaire can be used to help identify patients at greater risk of having sleep apnea. The Berlin questionnaire (BQ) has commonly been used as a measure in screening for sleep apnea especially over the past few years. The BQ has been shown to perform with high sensitivity and moderate specificity (0.97 and 0.54, respectively) in a primary care population [15]. However, the sensitivity and specificity values obtained by Netzer and colleagues have been criticized [16] as being erroneous and the suggested sensitivity and specificity for Netzer's study are 0.95 and 0.48, respectively. Furthermore, the BQ has never been validated in a sleep clinic population where the presence of other sleep disorders, in particular excessive daytime sleepiness, may alter its sensitivity and specificity. The primary aim of this study is to determine the specificity and sensitivity of the BQ in patients referred to a sleep clinic. Furthermore, the sensitivity and specificity of individual items of the BQ will be closely examined to see if it is feasible to develop an abbreviated version of the BQ. This study will also explore if there is a difference in the sensitivity and specificity of the BQ in male and female sleep clinic referrals.

Patients with sleep apnea commonly endorse symptoms of excessive daytime sleepiness [17], fatigue [18], low mood [19] and poor or unrefreshing sleep [20]. Hence, this study will also determine if there is a relationship between BQ scores or RDI measurements and scores on other questionnaires that assess symptoms of sleepiness, fatigue, low mood, and insomnia.

Materials and methods

Study population

A total of 130 patients who were referred to see one of the sleep physicians at the Sleep and Alertness Clinic at the Toronto Western Hospital between June 2004 and June 2005 were included in this retrospective chart review study. For the chart selection, the person collecting the data chose charts in order from the alphabetically filed charts in the clinic. Those charts that met the criteria were selected, and the process was continued until 130 qualified charts, as determined by the inclusion and exclusion criteria, and the power analysis below, were collected.

The Sleep and Alertness Clinic is a large, nine-bed tertiary-center affiliated sleep clinic in downtown Toronto associated with a tertiary care center, the Toronto Western Hospital of the University Health Network. Unlike many respiratory-based sleep clinics, the patient population in this clinic is very diverse and patients present with various sleep disorders including those secondary to psychiatric or neurological disorders. On average, about one in every four patient is referred to this clinic for investigation of possible sleep apnea. The physician whose charts were used in this study is a sleep specialist who exclusively sees patients with sleep disorders. The 130 patients were a subset of the referred patients in that they endorsed symptoms of excessive daytime sleepiness and/or poor, unrefreshing sleep and were asked to undergo polysomnographic (PSG) recording to investigate these symptoms. At the time the study was being conducted, it was standard practice for sleep clinic patients, especially those with symptoms of excessive daytime sleepiness, to receive two overnight sleep studies. About 90% of the patient charts available for this study, before application of the inclusion and exclusion criteria, contained two nights of diagnostic PSG recordings. Very recently and subsequent to this study, the use of two overnight diagnostic PSGs has been discontinued due to cutbacks in health care funding.

Every patient was also asked to complete a questionnaire battery that includes the BQ along with other scales, and these questionnaire results are used in conjunction with the sleep study findings when clinical diagnoses are made. The inclusion criteria for this study were patients referred for sleepiness and/or poor sleep, recordings available from two overnight PSGs and a completed questionnaire battery including the BQ. Charts of patients with incomplete or missing data for the respiratory variables were excluded. The study protocol was approved by the ethics board of the University Health Network.

Berlin questionnaire

The BQ includes five items on snoring (category 1, items 1–5), three items on daytime somnolence (category 2, items 6–8), and one item on the history of hypertension (category 3, item 9). The questionnaire also includes information about age, gender, height, weight, and neck size. The BQ was scored as previously reported by Netzer and colleagues

[15]. The overall score is based on the patients' responses to each of the three categories of the BQ. The snoring or daytime somnolence categories are positive if responses indicate persistent symptoms (>three to four times a week) on the questionnaire items. A positive score on the third category requires a history of hypertension or a Body Mass Index (BMI) of greater than 30 kg/m^2. Study patients were classified as being at a high risk of having sleep apnea if scores were positive on two or more of the three categories. Those patients who scored positively on less than two categories were identified as being at a low risk of having sleep apnea.

For the abbreviated BQ, category 2 containing the three items on daytime somnolence (how often they felt tired or fatigued after sleep; how many times they felt tired or fatigued during their waking time; if they had ever fallen asleep while driving a vehicle) and question 4 (if their snoring had ever bothered anyone else) were removed leaving questions 1, 2, 3, 5, and 10. The sleepiness items were removed to see if other causes of sleepiness in sleep clinic patients were responsible for the low sensitivity and specificity. The question about their snoring bothering anyone else was also removed, as this question was poorly answered in the subject population. When the abbreviated version was used, study subjects were classified as being at high risk of having sleep apnea if they scored positive on at least one of the two categories. Those subjects who did not score positive on any of the two categories were classified as being at low risk of having sleep apnea. The scoring of the two categories in the abbreviated version was the same as the scoring of the categories in the original BQ despite the fact that item 4 was removed from category 1.

Polysomnography

The polysomnographic data used in this study were extracted from the two consecutive overnight sleep studies. Charts of patients who slept less than 240 min on either of the two nights were excluded from the study. The sleep studies were scored according to standardized procedures [21]. Apneas/hypopneas were scored where there was a 50% or greater reduction in the baseline amplitude of respiration or at least 3% reduction in oxygen saturation, either lasting for a minimum of 10 s [22]. The maximum RDI for the two nights was used for the sensitivity and specificity analysis.

Other questionnaires

As excessive daytime sleepiness is a common complaint in sleep apnea, the Epworth Sleepiness Scale [23] (ESS) was included in the study and a cutoff of greater than 10 was considered as pathological sleepiness [24]. Other symptoms characteristic of sleep apnea such as fatigue, poor and unrefreshing sleep, and low mood were measured by the Fatigue Severity Scale [25] (FSS), Athens Insomnia Scale [26] (AIS), and the Center for Epidemiological studies in Depression scale [27] (CES-D), respectively. A score of greater than 3.7 on the FSS, representing two SDs above the normal healthy adult score, was used to indicate excessive fatigue [25]. In addition, a score of greater than 6 on the AIS was considered indicative of insomnia [26], and a score of greater than 16 on the CES-D was indicative of depressed mood [27].

Statistical analysis

Power analysis determined that a minimum sample size of 123 patient charts were required to detect a difference of 3 in the RDI value to a 95% confidence level (95% CI). Data in the text are expressed as means ± SDs. Statistical analyses were performed using SAS (version 9.1). Correlational analysis using the Spearman's rank-order correlational coefficient was used to determine the relationship between the estimate of risk on the Berlin Questionnaire and the patients' RDI and also to assess the correlation between RDI and BMI, RDI and neck size, and RDI and age. Statistical significance was set at $p < 0.05$.

Results

Study population

The study consisted of 130 sleep clinic patients. Males comprised 54% ($n=70$) of the population. The mean age, body mass index (BMI), and neck size for the male patients were: 42.2 years, 27.9 kg/m^2, and 39.9 cm, respectively. The mean age, BMI and neck size for the female patients were: 45.1 years, 28.1 kg/m^2 and 35.4 cm, respectively.

Polysomnography

The distribution of the RDI measurements obtained from the two overnight polysomnographic studies are shown in Table 1. Furthermore, the prevalence of a maximum RDI>5 was 56 (43.1%), maximum RDI>10 was 34 (26.2%), and maximum RDI>15 was 28 (21.5%) in the total study population.

Berlin questionnaire

Of the 130 participants, 76 (58.5%) scored positively in category 1, 97 (74.6%) scored positively in category 2, and 52 (40.0%) scored positively in category 3. The BQ identified 58.5% ($n=76$) study patients as being at a high

Table 1 RDI distribution for night 1 and night 2

RDI Distribution	RDI of night 1: # patients (percentage)	RDI of night 2: # patients (percentage)	Maximum RDI: # patients (percentage)
Females			
<5	45 (75%)	46 (77%)	42 (70%)
5–10	9 (15%)	7 (12%)	8 (13%)
11–20	4 (7%)	4 (7%)	6 (10%)
21–30	1 (2%)	0 (0%)	1 (2%)
>31	1 (2%)	3 (5%)	3 (5%)
Males			
<5	37 (53%)	39 (56%)	32 (46%)
5–10	13 (19%)	13 (19%)	14 (20%)
11–20	8 (11%)	4 (6%)	8 (11%)
22–30	3 (4%)	4 (6%)	3 (4%)
>31	9 (11%)	10 (14%)	13 (19%)

Total: 60 sleep clinic females; 70 sleep clinic males
RDI Respiratory disturbance index, *# patients* Number of patients

Table 3 BQ Classification for each RDI Range

RDI	High BQ risk # patients (percentage)	Low BQ risk # patients (percentage)
<5	38 (50%)	36 (67%)
5–<10	17 (22%)	5 (9%)
10–<15	5 (7%)	1 (2%)
>15	16 (21%)	12 (22%)

BQ Berlin Questionnaire, *RDI* Respiratory Disturbance Index (maximum), *# patients* Number of patients

risk of having sleep apnea and 41.5% ($n=54$) at a low risk of having sleep apnea. BQ sensitivity and specificity were calculated at various RDI cut-offs as shown in Table 2. Overall, the BQ performed with moderate specificity and low sensitivity. The analysis showed a large number of false negatives (patients with RDI>15 designated as being at low risk of having sleep apnea) and false positives (patients with RDI<5 identified as having a high sleep apnea risk) when the BQ was used as a screening tool. Table 3 shows the number of individuals with positive (high risk of sleep apnea) and negative (low risk of sleep apnea) scores on the BQ in each RDI range.

The individual items of the BQ and several abbreviated models of the BQ were further analyzed. It was found that an abbreviated version of the BQ restricted to items 1, 2, 3, 5, and 10 (i.e., removal of category 2 and item 4) performed with a sensitivity and specificity of 0.80 (95% CI: 0.62–0.91) and 0.42 (95% CI: 0.31–0.51), respectively, compared to 0.63 (95% CI: 0.44–0.77) and 0.43 (95% CI: 0.33–0.53) for the full questionnaire at a RDI cut-off of 10. If the abbreviated BQ is rescored such that a score of only 1 or more is required for a positive score in category 1, the

sensitivity and specificity are 0.89 (CI: 0.72–0.96) and 0.28 (CI: 0.19–0.37) respectively. Moreover, when assessed separately, the sensitivity and specificity of the BQ in males and females were very similar (males: sensitivity= 0.63 (95% CI: 0.44–0.77), specificity=0.41 (95% CI: 0.32–0.52); females: sensitivity=0.60 (95% CI: 0.41–0.74), specificity=0.44 (95% CI: 0.34–0.54).

Association between RDI and BQ score

The scores on BQ and the RDI measurements were in agreement for 47.7% of the study population. To qualify as being in agreement, those patients the BQ identified as being at high risk also had an RDI>10 or that those identified by the BQ at being at low risk had an RDI<10. Furthermore, there is weak evidence (Wilcoxon rank sum test, maximum RDI of two nights $p=0.056$) that the RDI scores are higher in the group identified as being at high risk of having sleep apnea by the BQ. The Spearman's correlation analysis showed a moderate correlation between RDI and category 1 scores (maximum RDI of two nights: $r=0.331$, $p=0.0001$). Furthermore, the correlation between RDI and category 1 score for men (maximum RDI of two nights: $r=0.319$, $p=0.007$) was slightly stronger than the correlation between RDI and category 1 for women (maximum RDI of two nights: $r=0.235$, $p=0.07$). However, there is no correlation between RDI and category 2 regardless of gender (maximum RDI of two nights: $r=-0.105$, $p=0.24$). Analysis of how the subjects answered category 3 (i.e., question 10) of the BQ and RDI measurement showed that for those with RDI>10, the response to category 3 was an equal split (50% responded 'yes' and 50% responded 'no'). However, for those subjects with RDI<10, they were more likely to respond 'no' (82.5%).

Association between RDI and BMI, neck size, and age

The analysis showed a relationship between RDI and BMI (maximum RDI of two nights: $r=0.262$, $p=0.003$). Furthermore, neck size was found to be strongly correlated with RDI (maximum RDI of two nights: $r=0.353$, $p<0.0001$). The relationship between RDI and age is also

Table 2 Sensitivity and Specificity of the BQ at varying severities of sleep apnea in the sleep clinic population

RDI	BQ sensitivity (95% CI: lower limit–upper limit)	BQ specificity (95% CI: lower limit–upper limit)
RDI>5	0.68 (0.54–0.79)	0.49 (0.37–0.60)
RDI>10	0.62 (0.44–0.77)	0.43 (0.33–0.53)
RDI>15	0.57 (0.37–0.75)	0.41 (0.32–0.51)

BQ Berlin Questionnaire, *RDI* Respiratory Disturbance Index

shown to be strong (maximum RDI of two nights: $r=0.512$, $p<0.0001$). Therefore, people diagnosed with sleep apnea are statistically significantly older (t test, $p<0.0001$, mean age for those with RDI>10: 56.1 years, mean age for those with RDI<10: 39.1 years). However, there is no difference in age between those who scored at low and high risk on the BQ (t test, $p=0.11$, mean age for low risk 36.1, mean age for high risk 41.6).

Relationship between BQ, RDI, and other questionnaires

The analysis showed that there was an association between RDI measurements and scores on the CES-D where subjects with RDI>10 had lower CES-D scores (Wilcoxon rank sum test, $p=0.046$). The median score of the CES-D was 15 for subjects with RDI>10 compared to a median score of 19 for subjects with RDI<10. There was no difference in ESS scores between subjects with RDI>10 and those with RDI<10 (Wilcoxon rank sum test, $p=0.91$). The median ESS score for subjects with RDI>10 was 9 and for subjects with RDI<10 was 10. In addition, there was no difference in FSS or AIS scores between subjects with RDI>10 and those with RDI<10 (FSS: Wilcoxon rank sum test, $p=0.78$; AIS: Wilcoxon rank sum test, $p=0.24$). The median FSS score for subjects with RDI>10 was 5.1 and for subjects with RDI<10 was 5.3, and the median AIS score was 12 for subjects with RDI>10 and 13 for subjects with RDI<10.

Analysis of the relationship between the BQ and other questionnaires showed that subjects identified as being at high risk on the BQ had higher CES-D scores (Wilcoxon rank sum test, $p=0.03$), higher ESS scores (Wilcoxon rank sum test, $p=0.02$) and higher FSS scores (Wilcoxon rank sum test, $p=0.04$). The median scores of the CES-D, ESS, and FSS for the group identified as being at high risk on the BQ were 21, 11, and 5.8, respectively, whereas the median scores of the CES-D, ESS, and FSS for the group identified as being at low risk on the BQ were 14.5, 7, and 4.7, respectively. However, there was no difference in AIS scores between subjects who were identified as being at high risk versus low risk on the BQ (Wilcoxon rank sum test, $p=0.19$). The median score of the AIS for the group identified as being high risk on the BQ was 13 and for the group identified as being at low risk on the BQ was 12.

Discussion

The present study demonstrates that the BQ is not a good instrument for the estimation of sleep apnea risk in patients referred to the sleep clinic. The low sensitivity and specificity of the BQ and the large number of false positives and false negatives indicate that the Berlin Questionnaire

would not be a useful clinical adjunct for screening for sleep apnea in patients referred to a sleep clinic. This study further shows that in this population, the ability of the BQ to predict sleep apnea is independent of the gender of the patient as the BQ performs with similar low specificity and sensitivity in both males and females.

Rowley and colleagues [28] have previously examined four potential questionnaires that were thought to be predictive of sleep apnea in a sleep clinic population. They similarly concluded that none of the four questionnaires was an accurate predictor of sleep apnea in a sleep clinic population. One of the questionnaires evaluated by Rowley and colleagues was developed by Maislin's group [29] and is almost an exact duplication of the BQ with the exception of the sleepiness component. Rowley and colleagues reported the sensitivity for this questionnaire to be 0.87 at RDI cut-off of 10. This is substantially higher than the sensitivity found in our study for the BQ (0.62, at the same RDI cut-off). The main difference between the two questionnaires is the sleepiness component. Hence, one might hypothesize that this component was at least partially responsible for the reduction in sensitivity. In fact, our abbreviated version of the BQ with the sleepiness component removed (along with item 4) and a slightly modified scoring method for category 1 performed with a much higher sensitivity (0.80). Furthermore, the result showed that 74.6% of the participants in our study population responded positively to category 2, which represents the sleepiness component of the BQ. This proportion is substantially higher than those found in other populations. Mustafa's study [30] used the BQ to find the frequency of those at high risk for sleep apnea in a primary care population and found 43% scoring positive on the sleepiness category. In another study, Principe-Rodriguez et al. [31] used the BQ in a sample of patients with heart failure for the same purpose and found that 48% of the population responded positively to the sleepiness category. Therefore, it can be proposed that a subject population that is comprised of a large proportion of sleepy individuals may reduce the ability of the BQ to detect sleep apnea. However, as one would expect that a higher prevalence of sleepiness would decrease the specificity and not the sensitivity, other factors are likely involved.

In addition, the mean BMI of the population in this study was slightly elevated with many patients being overweight and obese (mean BMI=28). The higher than average BMI of the sleep clinic population likely affected the sensitivity and specificity of the BQ, as the BMI also factors heavily into the scoring of the BQ. In fact, in our study population, the BMI and the neck size were strongly correlated with the RDI measurements and that is in line with previous findings [32] where a larger neck size and higher BMI were identified as being predictive of sleep apnea.

However, the study by Ogretmenoglu and colleagues [32] involved only patients who were referred for suspected sleep apnea, while the present study involved all sleepy patients or those with poor sleep who were referred to a general sleep clinic. Overall, our clinic population is composed of patients with a wide variety of sleep and mood disorders. However, similar to Rowley's group [28], we showed that snoring and the presence of hypertension (category 3 on BQ) was not significantly different in patients with sleep apnea and those without sleep apnea. In contrast to Rowley's group [28], this current study showed that the mean age of patients diagnosed with sleep apnea is significantly higher than that of patients not diagnosed with sleep apnea.

The results of this study differ from the findings in the literature where the BQ was shown to detect sleep apnea with a high sensitivity and specificity [15, 33]. However, Netzer and colleagues [15] validated the BQ in a primary care population and Gami [33] and co-workers conducted their validation in a general cardiology population. The present study evaluates the BQ in a sleep clinic population that is characterized by individuals with high subjective sleepiness and above average BMI. This is supported by the finding that those patients predicted by the BQ as being at a high risk of having sleep apnea had a significantly higher ESS score than those predicted to be at low risk, while in fact, those patients clinically diagnosed with sleep apnea were sleepy, but not sleepier than those who were not diagnosed with sleep apnea.

This study has several limitations that need to be considered. One limitation of this study is that it is retrospective in nature. The BQ data were obtained from the charts of the patients referred to the sleep clinic where the BQ along with a number of other questionnaires had been completed by the patients before their initial consultation with a sleep specialist. The study population included all of those who endorsed symptoms of daytime sleepiness, poor sleep or were suspected of having a sleep disorder based on clinical interview, and those who underwent two nights of polysomnographic study. In general, having patients in our clinic undergo two nights of testing was not a problem for the vast majority, and in fact, only a very few were unable to complete their sleep studies, these being mostly due to anxiety issues. The patient charts used in the current study are representative of a sleep clinic population with a majority of patients complaining of excessive sleepiness. Another limitation is that the chart selection process was not entirely random, as charts were chosen in consecutive order in the filing system and selected based on the inclusion and exclusion criteria.

Another limitation of the current study has to do with the generalizability of the findings. The data for this study were collected from the largest sleep clinic in Downtown Toronto, which is associated with the Department of Psychiatry of a tertiary care center, the Toronto Western Hospital of the University Health Network. Because of its association with this large center, the patient population is very diverse, and diagnoses include various sleep disorders. About one in every four patients referred to the sleep clinic is diagnosed with sleep apnea. Therefore, the composition of the patient population might be different from other sleep clinics especially those run by respirologists. On the other hand, having a population comprised primarily of sleep apnea patients would also have introduced other biases to the study.

In conclusion, the findings of the current study strongly recommend against using the BQ for clinical or research purposes to determine the level of risk of sleep apnea in a sleep clinic population and would suggest caution when using the BQ in a research study composed of primarily sleepy subjects. However, the findings of this study suggest alternative ways through which the sensitivity and specificity of the BQ can be improved. This study also emphasizes the need to validate the BQ in differing study populations before use in the clinical setting or research studies. The BQ is an inexpensive and quick screening tool, a fact that is especially beneficial in research, but further studies are needed in other sleepy populations, for example shiftworkers, to determine if the BQ can reliably identify undiagnosed sleep apnea.

References

1. Young T, Evans L, Finn L, Palta M (1997) Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. Sleep 20(9):705–706 (Sep)
2. Bahammam A, Delaive K, Ronald J, Manfreda J, Roos L, Kryger MH (1999) Health care utilization in males with obstructive sleep apnea syndrome two years after diagnosis and treatment. Sleep 22 (6):740–747 (Sep 15)
3. Bixler EO, Vgontzas AN, Lin HM, et al (2000) Association of hypertension and sleep-disordered breathing. Arch Intern Med 160(15):2289–2295
4. Dyken ME, Somers VK, Yamada T, Ren ZY, Zimmerman MB (1996) Investigating the relationship between stroke and obstructive sleep apnea. Stroke 27(3):401–407 (Mar 1)

5. He J, Kryger MH, Zorick FJ, Conway W, Roth T (1988) Mortality and apnea index in obstructive sleep apnea. Experience in 385 male patients. Chest 94(1):9–14 (Jul)
6. Hung J, Whitford EG, Parsons RW, Hillman DR (1990) Association of Sleep–Apnea with Myocardial-Infarction in Men. Lancet 336(8710):261–4 (Aug 4)
7. Riley RW, Powell NB, Guilleminault C, Clerk A, Troell R (1995) Obstructive sleep apnea. Trends in therapy. West J Med 162 (2):143–148 (Feb)
8. Stepnowsky C, Orr W, Davidson T (2004) Nightly variability of sleep-disordered breathing measured over 3 nights. Otolaryngol Head Neck Surg 131(6):837–843 (Dec)
9. Ford DE, Kamerow DB (1989) Epidemiologic study of sleep disturbances and psychiatric disorders. An opportunity for prevention? JAMA 262(11):1479–1484. (Sep 15)
10. Lu CL, Lee YC, Tsai SJ, Hu PG, Sim CB (1995) Psychiatric disturbances associated with hyperthyroidism: an analysis report of 30 cases. Zhonghua Yi Xue Za Zhi (Taipei) 56(6):393–398 (Dec)
11. Guilleminault C, Czeisler C, Coleman R, Miles L (1982) Circadian rhythm disturbances and sleep disorders in shift workers. Electroencephalogr Clin Neurophysiol Suppl 36:709–714
12. Melamed S, Oksenberg A (2002) Excessive daytime sleepiness and risk of occupational injuries in non-shift daytime workers. Sleep 25(3):315–322 (May 1)
13. Bassetti CL (2005) Sleep and stroke. Semin Neurol 25(1):19–32 (Mar)
14. Thorpy MJ (2004) Sleep disorders in Parkinson's disease. Clin Cornerstone 6(Suppl 1A):S7–S15
15. Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP (1999) Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. Ann Intern Med 131(7):485–491 (Oct 5)
16. Strauss R, Browner W (2000) Letter to the editor: sensitivity, specificity and positive predictive value of Berlin Questionnaire. Ann Intern Med 132(9):758–759
17. Guilleminault C, Dement WC (1977) 235 cases of excessive daytime sleepiness. Diagnosis and tentative classification. J Neurol Sci 31(1):13–27 (Jan)
18. Chervin RD (2000) Sleepiness, fatigue, tiredness, and lack of energy in obstructive sleep apnea. Chest 118(2):372–379 (Aug)
19. Balan S, Spivak B, Mester R, Leibovitz A, Habot B, Weizman A (1998) Psychiatric and polysomnographic evaluation of sleep disturbances. J Affect Disord 49(1):27–30 (Apr)
20. Zwillich CW, Pickett C, Hanson FN, Weil JV (1981) Disturbed sleep and prolonged apnea during nasal obstruction in normal men. Am Rev Respir Dis 124(2):158–60 (Aug)
21. Rechtschaffen A, Kales A. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. 1968. Washington, US Department of Health, Education and Welfare, National Institute of Health [Ref Type: Serial (Book, Monograph)]
22. Sleep-related breathing disorders in adults (1999) recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. Sleep 22(5):667–689 (Aug 1)
23. Johns MW (1991) A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 14(6):540–545 (Dec)
24. Johns M, Hocking B (1997) Daytime sleepiness and sleep habits of Australian workers. Sleep 20(10):844–849 (Oct)
25. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD (1989) The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 46 (10):1121–1123 (Oct)
26. Soldatos CR, Dikeos DG, Paparrigopoulos TJ (2003) The diagnostic validity of the Athens Insomnia Scale. J Psychosom Res 55(3):263–277 (Sep)
27. Radloff L (1977) The CES-D scale: a self-report depression scale for research in the general population. App Psychol Meas 385–401 (Ref Type: Generic)
28. Rowley JA, Aboussouan LS, Badr MS (2000) The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. Sleep 23(7):929–938 (Nov 1)
29. Maislin G, Pack AI, Kribbs NB, et al (1995) A survey screen for prediction of apnea. Sleep 18(3):158–166 (Apr)
30. Mustafa M, Erokwu N, Ebose I, Strohl K (2005) Sleep problems and the risk for sleep disorders in an outpatient veteran population. Sleep Breath 9(2):57–63 (Jun)
31. Principe-Rodriguez K, Strohl KP, Hadziefendic S, Pina IL (2005) Sleep symptoms and clinical markers of illness in patients with heart failure. Sleep Breath 9(3):127–133 (Sep)
32. Ogretmenoglu O, Suslu AE, Yucel OT, Onerci TM, Sahin A (2005) Body fat composition: a predictive factor for obstructive sleep apnea. Laryngoscope 115(8):1493–1498 (Aug)
33. Gami AS, Pressman G, Caples SM, et al (2004) Association of atrial fibrillation and obstructive sleep apnea. Circulation 110 (4):364–367 (Jul 27)