ELSEVIER

Sleep Medicine 3 (2002) 133–138

SLEEP MEDICINE

www.elsevier.com/locate/sleep

Original article

Pulseoximetry: sufficient to diagnose severe sleep apnea☆

Britt Oeverland*, Olav Skatvedt, Kari J. Kværner, Harriet Akre

SRBD Unit, Department of Otorhinolaryngology, Ullevaal University Hospital, N-0403 Oslo, Norway

Received 28 December 2000; received in revised form 14 April 2001; accepted 17 April 2001

Abstract

Objectives: To assess the clinical value of pulseoximetry in the diagnosis of sleep apnea when satisfactory agreement with polysomnography is obtained.

Methods: This was a prospective clinical study, set in the Department of Otorhinolaryngology, Ullevaal University Hospital, Oslo, Norway. One hundred consecutive patients were investigated for sleep related breathing disorders. The main outcome measurements were: measurement success rate, oxygen desaturation thresholds, sensitivity and specificity at apnea–hypopnea-index (AHI) thresholds of 5 and 15.

Results: Pulseoximetry was successfully performed in 93%. When different oxygen desaturation thresholds were calculated, optimal agreement with polysomnography was found at a 3% oxygen desaturation level. The sensitivity and specificity of diagnosing moderate/severe sleep apnea (AHI above 15) were 0.86 and 0.88, respectively. The corresponding figures for milder sleep apnea (AHI above 5) were 0.91 and 0.67. Good agreement was found between the AHI and the oxygen desaturation index (ODI) at the 3% level, with a mean AHI–ODI difference of 2.6 (SD, 7.3), a Pearson correlation of 0.95 and a weighted kappa of 0.86. The best agreement was found for AHI values below 15, where the estimated AHI–ODI difference was only −0.4 (SD, 3.3).

Conclusions: Pulseoximetry is a simple, non-invasive procedure, which is easy to perform and well suited for outpatient registration. When adjusted to polysomnography with high sensitivity of hypopnea registrations, an ODI at the 3% level is optimal to diagnose sleep apnea. In patients with moderate/severe sleep apnea with AHI values above 15, it is sufficient to establish the diagnosis and subsequent treatment. A negative pulseoximetry does not rule out sleep disorders; the patients should complete a full examination. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Polysomnography; Hypopnea; Sleep related breathing disorders; Screening

1. Introduction

Due to the fact that obstructive sleep apnea is claimed to be an important cause of premature death and disability, there is increasing pressure to provide sleep services for the treatment of these patients [1–6]. Although the golden standard diagnostic test is multichannel polysomnography, which enables the detection of obstructive apneas, hypopneas and arousals, the method is expensive and time-consuming.

Efforts have been made to identify screening tools for sleep apnea. Pulseoximetry is fast, easy to perform, and well suited for outpatient registration. Due to the reported sensitivity and specificity variation of the method [7–12], however, its ability to replace polysomnography has been questioned [13,14]. To our knowledge, clinical reports have not used optimized pulseoximetry oxygen desaturation thresholds to compare polysomnography in the diagnosis of sleep apnea syndrome [7,15]; we therefore questioned the clinical significance of these thresholds.

High sensitivity polysomnography with esophagus pressure measurements is routinely applied at the Sleep Related Breathing Disorders Unit of Ullevaal University Hospital, allowing pulseoximetry desaturations to be validated against high sensitivity polysomnography. To assess the clinical value of pulseoximetry when satisfactory agreement with polysomnography was obtained, optimal pulseoximetry oxygen desaturation levels were estimated and subsequently studied to determine whether pulseoximetry is sufficient to establish a diagnosis in patients with moderate/severe sleep apnea with apnea–hypopnea-index (AHI) values above 15 at oxygen desaturation thresholds of 3 and 4%.

☆ Financial disclosure: the authors do not have any financial involvement in any organization with direct financial interest in the subject.

* Corresponding author. Tel.: +47-23015631; fax: +47-22118555.
E-mail address: britt.oeverland@ulleval.no (B. Oeverland).

1389-9457/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S1389-9457(01)00122-8

2. Materials and methods

2.1. Study population

A prospective clinical study was performed in 100 consecutive patients undergoing polysomnography for diagnosis of sleep related breathing disorders (SRBD). The source population included all patients referred to Ullevaal University Hospital for polysomnography in the period from December 1998 to August 1999. All patients who had polysomnography and pulseoximetry performed were included in the study. Patients with incomplete oximetry recordings due to probe disconnection were excluded from further analyses. Since both pulseoximetry and polysomnography are part of the investigative procedures for diagnosing sleep disorders at the Sleep Related Breathing Disorders Unit of Ullevaal University Hospital, informed consent was not obtained. In contrast to the normal procedure, the senior executive officer assigned a priori every third patient on the referral list to both polysomnography and pulseoximetry.

2.2. Measurements

2.2.1. Polysomnography

The polysomnography (Alice3, Healthdyne International) included a four-channel electroencephalogram (EEG), a two-channel electrooculogram (EOG) and a chin electromyogram (EMG) for sleep staging. Oxygen saturation (SaO_2) was measured with a Criticare 504 oximeter and a finger probe. The pulseoximeter averaging time was set to 3 s. Thoracic and abdominal movements were measured by strain gauge thoracic and abdominal belts. A chest sensor for tracking body position was used, as well as an electrocardiogram (ECG). The oro-nasal airflow was measured by an external thermistor and by internal thermistors in an esophageal catheter located in the pharynx. The catheter also measured the pressure in the pharynx and in the esophagus [16–18]. An apnea/hypopnea was scored according to the following criteria: a decrease (>50%) in amplitude from the baseline of pharyngeal air flow measured by internal thermistors, or a clear decrease (<50%) in amplitude followed by either an oxygen desaturation (≥3%) or an arousal, each event lasting 10 s or longer [19]. The diagnosis of SRBD was confirmed when the number of apneas and hypopneas/h of sleep (AHI) was ≥5.

2.2.2. Pulseoximetry

We used an additional pulseoximeter (NONIN model 8500) which samples the SaO_2 every 4 s. Stored data transferred to the computer program Profox version NONIN 06/97 (Profox Associates, Inc., CA) included the oximetry tracing from the first to the last sleep page according to analysis performed with the Alice program. The Profox program calculated the number of SaO_2 reductions ≥2, 3, 4 and 5%, a desaturation event beginning with a saturation decrease of the appropriate percentage or more within a 2 min interval and terminating when the saturation rose by the appropriate percentage above the lowest saturation recorded during that event. Events lasting for 3 min or more were not counted, nor were events containing more than 8 s of deleted or zero data. SRBD diagnosis was established when the number of oxygen desaturations/h (oxygen desaturation index (ODI)) was ≥5.

2.3. Outcome measurements

AHI and ODI were the outcomes of interest. Patients were divided into groups according to the severity of the syndrome: mild (AHI 5 ≤ 15), moderate (15 ≤ 30) and severe (AHI >30) [19]. In order to study the ability of pulseoximetry to differentiate between mild and moderate/severe disease, AHI and ODI were dichotomized with a limit of more than 15 obstructive events/h to define moderate/severe. Sociodemographic factors included body mass index (BMI), age and gender. ODI was divided into the same groups as AHI. The distribution is presented in Table 1.

2.4. Statistical analysis

The sensitivity and specificity for the different thresholds for SaO_2-desaturations were calculated. Paired, two-tailed *t*-tests were used for intraindividual comparisons of AHI–ODI differences.

Measurement of the agreement between pulseoximetry and polysomnography was performed according to a

Table 1
Characteristics of the study population by severity of the sleep apnea syndrome[a]

	Severity of sleep apnea		Cohort ($N = 93$)
	Non-severe sleep apnea: AHI ≤ 15 ($n = 49$)	Moderate/severe sleep apnea: AHI >15 ($n = 44$)	
Gender			
Female	15	5	20
Male	34	39	73
Age			
≤45	32	22	54
45–54	10	9	19
≥54	7	13	20
ODI			
≤5	19	0	19
5–15	24	6	30
15–30	6	15	21
≥30	0	23	23
BMI[b]			
≤25	21	9	30
25–30	16	18	34
≥30	7	15	22

[a] $N = 93$.
[b] Seven missing.

Table 2
Sensitivity and specificity of pulseoximetry estimated at various oxygen desaturation levels by definition of sleep apnea[a] and moderate/severe disease[b]

Pulseoximetry oxygen desaturation reduction	Polysomnography cut off level: AHI ≥ 5	
	Sensitivity	Specificity
2%	1.00	0.00
3%	0.91	0.67
4%	0.73	1.00
5%	0.57	1.00
	Polysomnography cut off level: AHI ≥ 15	
	Sensitivity	Specificity
2%	1.00	0.27
3%	0.86	0.88
4%	0.64	1.00
5%	0.52	1.00

[a] AHI ≥ 5.
[b] AHI ≥ 15.

method described by Bland and Altman [20]. The weighted kappa statistics were calculated.

3. Results

Of the 100 patients, 93 had satisfactory overnight pulseoximetry and polysomnography recordings. Seven patients had incomplete recordings due to technical errors, all due to pulseoximetry probe disconnection. These patients did not differ from the rest of the population in the severity of illness. The study population consisted of 21% ($n = 20$) women and 79% ($n = 73$) men, and the mean age was 44 years (range, 27–76). The mean AHI was 23 (range, 0–78) and the mean BMI was 28 (range, 21–56). The characteristics of the study population by severity of the disease are given in Table 1.

3.1. ODI sensitivity and specificity

The sensitivity and specificity of pulseoximetry with different thresholds of ODI were calculated (Table 2). The sensitivity declined and the specificity increased by the proportion of oxygen reduction required to record a significant desaturation increase. At a threshold of 2% for recording a desaturation, all of the patients with sleep apnea hypopnea syndrome (SAHS) were recognized with pulseoximetry, but the false positive rate was high. At a 4% oxygen desaturation threshold, no patients were incorrectly diagnosed with SAHS, but the corresponding rate of false negatives was high. Optimal agreement with polysomnography was found at the 3% oxygen desaturation level, with the sensitivity and specificity for diagnosing sleep apnea (AHI above 5) being 0.91 and 0.67, respectively (Table 2). Acceptable sensitivities and specificities were only obtained at 3 and 4% desaturation levels, and further analysis makes use of only these two desaturation levels.

3.2. ODI and AHI agreement at 3 and 4% desaturation values

When ODI was plotted against AHI (Fig. 1), high agreement was found for lower ODI and AHI values. Higher values revealed a tendency for higher AHI values than ODI values, particularly at the 4% oxygen desaturation level. The estimated Pearson correlation coefficients were 0.95 and 0.94 for 3 and 4% desaturation levels, respectively.

The difference between AHI and ODI plotted in Fig. 2 shows increasing AHI–ODI as the AHI and ODI values increase. Except for patients with AHI–ODI levels below 15 at the 3% desaturation level, there was a tendency for pulseoximetry to underestimate the severity of the disease (Table 3). The estimated mean AHI–ODI differences were 2.6 (SD, 7.3) and 8.7 (SD, 8.8) at 3 and 4% oxygen desaturation levels, respectively. Intrapair AHI–ODI differences by mild and moderate/severe disease are shown in Table 3.

3.3. Significance of ODI in the diagnosis of moderate/severe apnea

The sensitivity and specificity of diagnosing moderate/severe sleep apnea, defined as an AHI above 15, were 0.86 and 0.88 at the 3% oxygen desaturation level. Corresponding values at 4% desaturation were 0.64 and 1.0, respectively (Table 2).

To study the ability of pulseoximetry to differentiate between mild and more severe disease, the patients were divided into four groups by ODI and AHI values (Table 4). At the 3% oxygen desaturation level, the majority of the patients fitted identical ODI and AHI groups. Pulseoximetry correctly diagnosed most of the patients with moderate/severe sleep apnea, while 6% with moderate sleep apnea were mistakenly diagnosed as mild (Table 4). At the 4% desaturation threshold, most of the patients with moderate/

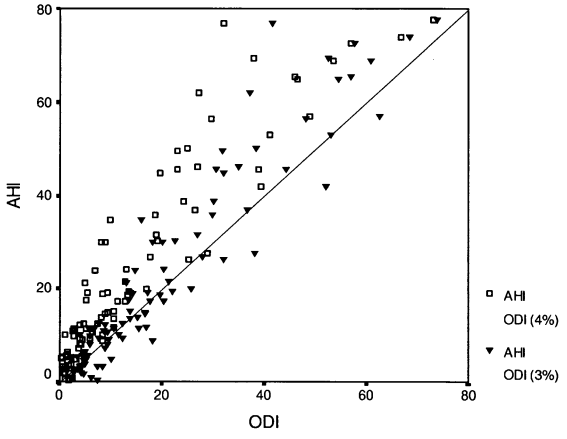

Fig. 1. ODI with a threshold limit of: (▼), 3%; and (□), 4%, and plotted against AHI by polysomnography with the line of equality.

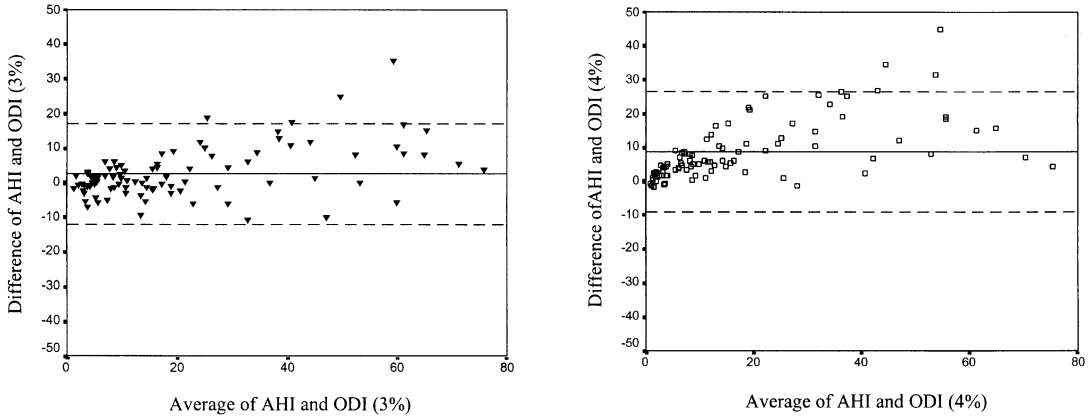

Fig. 2. Difference against mean for ODI (3 and 4%) from pulseoximetry and AHI measured with polysomnography (PSG). The middle line is the mean difference between the two methods. The upper and lower lines are the mean ± 2 SD. The mean differences were 2.6 and 8.7, and the standard deviations were 7.3 and 8.8 for ODI 3% and ODI 4%, respectively.

severe disease are correctly diagnosed, but 14% of the patients with moderate/severe disease were mistakenly diagnosed as mild. The estimated agreement between the two methods was very good ($\kappa = 0.86$) and good ($\kappa = 0.78$) at levels of 3 and 4% oxygen desaturation, respectively (Table 4).

4. Discussion

In order to reduce long-term sequelae and the costs of treatment of severe sleep apnea, efforts have been made to identify screening tools. Although it has been claimed that, due to the prevalence of false positives, pulseoximetry does not reduce the investigative costs of the disease [15], our study showed that, when adjusted to highly sensitive polysomnography, this method may be an easy and valid implement for the diagnosis of severe sleep apnea and for the selection of patients for continuous positive airway pressure treatment.

Due to the fact that there has been some disagreement about the optimal oxygen desaturation limit recommended to score hypopneas, the present study compared the AHI derived from polysomnography with the automatic ODI score with different thresholds for reduction in oxygen saturation. In contrast to former studies, a desaturation level of 3% was highly sensitive in the diagnosis of severe disease, and superior to all recordings at the 4% level. We routinely use a catheter with internal flow and pressure measurements during polysomnography with high reliability in diagnosing hypopneas [17,18]. Since the physiological definition of hypopneas is not standardized and the hypopnea recording techniques vary substantially [21], the sensitivity of the polysomnography used will influence the sensitivity and specificity of the pulseoximetry estimated.

Although our study showed good agreement between AHI and ODI, particularly at low levels, we found that pulseoximetry is most useful for diagnosing moderate/severe disease with a clinical AHI threshold of 15. Other clinical studies have targeted severe disease [7,15], probably because such patients are candidates for continuous positive airway pressure treatment. The titration of optimal positive airway pressure is itself diagnostic for severe disease, and thus extensive examination prior to treatment is unnecessary when thorough otolaryngological examination precedes pulseoximetry. For milder cases, however, treatment choices vary with degree of apneic events, and pulseoximetry cannot replace polysomnography. Furthermore, primary lung disease has to be ruled out in these patients [7]. Negative pulseoximetry screening and clinical investigation will routinely be followed by polysom-

Table 3
The mean difference between AHI and ODI at 3 and 4% pulseoximetry desaturation levels by severe and non-severe disease

	Number of pairs	Mean intrapair AHI–ODI difference	SD	*P* (two-tailed)
3% pulseoximetry desaturation level				
AHI ≤ 15	49	−0.4	3.3	0.369
AHI > 15	44	6.0	8.9	0.000
4% pulseoximetry desaturation level				
AHI ≤ 15	49	3.6	2.8	0.000
AHI > 15	44	14.5	9.7	0.000

Table 4
Distribution of pulseoximetry oxygen desaturations and polysomnography apnea–hypopneas grouped by severity of disease at 3 and 4% oxygen desaturation levels

	3% oxygen desaturation fall				4% oxygen desaturation fall			
	Oxygen desaturations/h (ODI 3%)				Oxygen desaturations/h (ODI 4%)			
	≤5 (normal)	5 ≤ 15 (mild)	15 ≤ 30 (moderate)	≥30 (severe)	≤5 (normal)	5 ≤ 15 (mild)	15 ≤ 30 (moderate)	≥30 (severe)
AHI								
≤5 (normal)	12	6	0	0	18	0	0	0
5 ≤ 15 (mild)	7	18	6	0	19	12	0	0
15 ≤ 30 (moderate)	0	6	11	2	1	14	4	0
≥30 (severe)	0	0	4	21	0	1	12	12
Total	19	30	21	23	38	27	16	12
Weighted kappa (95% CI)	0.86 (0.81–0.91)				0.78 (0.71–0.84)			

nography, and sleep disorders (e.g. upper airway resistance syndrome) will not be overlooked.

Earlier studies, although finding the method satisfactory for identifying the most severe cases, have concluded that pulseoximetry is inadequate as a screening tool due to the large number of false negatives [14,22] and disagreement of manual desaturation scores [14]. This problem can be avoided by using an automatic score for desaturations. Pulseoximetry has been proposed as a tool for ruling out the sleep apnea–hypopnea syndrome [10], but this is contradictory to our results.

Although the costs of introducing pulseoximetry as a screening device have not been estimated, there is a need for screening devices that can successfully reduce the Norwegian health care system's long referral lists for the diagnosis and treatment of sleep disorders. If pulseoximetry could identify severely ill patients needing automatic continuous airway pressure titration, the efficacy of sleep disorder diagnosis in our department would be significantly increased. Our study indicates that pulseoximetry is an inexpensive screening tool, highly sensitive and specific in diagnosing severe sleep apnea. The procedure is easy to perform and well suited for outpatient recording when optimal agreement to polysomnography is obtained.

5. Validity of results

The accuracy of pulseoximeters measuring transient oxygen saturation (SaO_2) changes may be affected by the pulseoximeter time response [23]. The averaging time may vary according to the type of oximeter. Since the averaging time determines the dynamic response of the pulseoximeter, high averaging values cause considerable underestimation of oxygen desaturation [23]. Sleep study response time recommendations are less than 10 s [24]. Dynamic pulseoximeter responses also vary with the type of probe [25,26], and both parameters must be considered when deciding which pulseoximeter to use.

In contrast to earlier reports [11], we found that pulseoximetry tended to underestimate the number of apnea–hypopnea events, except for AHI values below 15. Several limitations associated with pulseoximetry may account for this. A higher AHI value is recorded by increasing hypopnea sensitivity with simultaneous flow and pressure recordings. The method only records obstructive events with simultaneous oxygen desaturations, and its ability to record sleep time is limited. Total pulseoximetry time is recorded, but there is no instrument to register whether the patient has been awake or sleeping and the ODI is likely to be underestimated. The prospective study design and intraindividual comparisons of pulseoximetry and polysomnography results rule out selection bias as an influence on the study results, but since the range of apnea–hypopnea episodes is above those in the general population, they are limited in application to patients suspected of having sleep disorders.

References

[1] He J, Kryger MH, Zorick FJ, Conway W, et al. Mortality and apnea index in obstructive sleep apnea. Experience in 385 male patients. Chest 1988;94(1):9–14.
[2] Lavie P, Herer P, Hoffstein V. Obstructive sleep apnoea syndrome as a risk factor for hypertension: population study. Br Med J 2000;320:479–482.
[3] Partinen M, Palomaki H. Snoring and cerebral infarction. Lancet 1985;2(8468):1325–1326.
[4] Shepard JWJ. Hypertension, cardiac arrhythmias, myocardial infarction, and stroke in relation to obstructive sleep apnea. Clin Chest Med 1992;13(3):437–458.
[5] Sullivan CE, Issa FG, Berthon-Jones M, Eves L. Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. Lancet 1981;1(8225):862–865.
[6] Wright J, Johns R, Watt I, Melville A, et al. Health effects of obstructive sleep apnoea and the effectiveness of continuous positive airways pressure: a systematic review of the research evidence. Br Med J 1997;314(7084):851–860.
[7] Vazquez JC, Tsai WH, Flemons WW, Masuda A, et al. Automated

analysis of digital oximetry in the diagnosis of obstructive sleep apnoea. Thorax 2000;55(4):302–307.
[8] Chiner E, Signes-Costa J, Arriero JM, Marco J, et al. Nocturnal oximetry for the diagnosis of the sleep apnoea hypopnoea syndrome: a method to reduce the number of polysomnographies? Thorax 1999;54(11):968–971.
[9] Gyulay S, Olson LG, Hensley MJ, King MT, et al. A comparison of clinical assessment and home oximetry in the diagnosis of obstructive sleep apnea. Am Rev Respir Dis 1993;147(1):50–53.
[10] Series F, Marc I, Cormier Y, La Forge J. Utility of nocturnal home oximetry for case finding in patients with suspected sleep apnea hypopnea syndrome. Ann Intern Med 1993;119(6):449–453.
[11] Williams AJ, Yu G, Santiago S, Stein M. Screening for sleep apnea using pulse oximetry and a clinical score. Chest 1991;100(3):631–635.
[12] Ryan PJ, Hilton MF, Boldy DA, Evans A, et al. Validation of British Thoracic Society guidelines for the diagnosis of the sleep apnoea/hypopnoea syndrome: can polysomnography be avoided? Thorax 1995;50(9):972–975.
[13] Svanborg E, Larsson H, Carlsson-Nordlander B, Pirskanen R. A limited diagnostic investigation for obstructive sleep apnea syndrome. Oximetry and static charge sensitive bed. Chest 1990;98(6):1341–1345.
[14] Yamashiro Y, Kryger MH. Nocturnal oximetry: is it a screening tool for sleep disorders? Sleep 1995;18(3):167–171.
[15] Golpe R, Jimenez A, Carpizo R, Cifrian JM. Utility of home oximetry as a screening test for patients with moderate to severe symptoms of obstructive sleep apnea. Sleep 1999;22(7):932–937.
[16] Akre H, Skatvedt O, Oeverland B, Borgersen AK. Internal thermistors in differentiating between oral and nasal breathing during sleep. Acta Otolaryngol 1999;119(8):934–938.
[17] Akre H, Borgersen AK, Skatvedt O. Tracing air flow and diagnosing hypopneas in normal subjects. Physiol Meas 2000;21:221–227.
[18] Akre H, Skatvedt O. Advantages of measuring air flow in pharynx with internal thermistors. Eur Arch Otorhinolaryngol 2000;257(5):251–255.
[19] Anonymous. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. Sleep 1999;22(5):667–689.
[20] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–310.
[21] Redline S, Sanders M. Hypopnea, a floating metric: implications for prevalence, morbidity estimates, and case finding. Sleep 1997;20(12):1209–1217.
[22] Douglas NJ, Thomas S, Jan MA. Clinical value of polysomnography. Lancet 1992;339(8789):347–350.
[23] Farre R, Montserrat JM, Ballester E, Hernandez L, et al. Importance of the pulse oximeter averaging time when measuring oxygen desaturation in sleep apnea. Sleep 1998;21(4):386–390.
[24] Clark JS, Votteri B, Ariagno R, Cheung P, et al. Noninvasive assessment of blood gases. Am Rev Respir Dis 1992;145:220–232.
[25] Warley AR, Mitchell JH, Stradling JR. Evaluation of the Ohmeda 3700 pulse oximeter. Thorax 1987;42:892–896.
[26] West P, George CF, Kryger MH. Dynamic in vivo response characteristics of three oximeters: Hewlett–Packard 47201A, Biox III, and Nellcor N-100. Sleep 1987;10(3):263–271.