Sleep Breath (2011) 15:447–454
DOI 10.1007/s11325-010-0356-1

ORIGINAL ARTICLE

A two-tier screening model using quality-of-life measures and pulse oximetry to screen adults with sleep-disordered breathing

Ning-Hung Chen · Min-Chi Chen · Hsueh-Yu Li · Chang-Wei Chen · Pa-Chun Wang

Received: 20 January 2010 / Revised: 6 April 2010 / Accepted: 18 April 2010 / Published online: 7 May 2010

Abstract
Purpose Using quality-of-life measures and pulse oximetry, this study developed a two-tiered prediction algorithm with an aim to prioritize sleep-disordered breathing patients for polysomnography.
Methods Data from 355 patients were evaluated to obtain their clinical information, Chinese version of Epworth sleepiness scale, and snore outcomes survey scores against respiratory disturbance index (RDI). In the first-tier screening, receiver-operating characteristics were calculated with an initial strategy of choosing optimal prediction sensitivity. The second-tier strategy investigated the association between pulse oximetry data (desaturation index of 3%) against RDI to optimize prediction specificity.
Results The "SOS score of 55 and ESS score of 9" was the optimal combination that yielded the highest sensitivity (0.603) in the first-tier screening. The strategy can includ 94.93% possible patients (probability=0.6) with positive predictive value of 0.997. The area under the curve (AUC) was 0.88 ($p<0.001$). Desaturation index of 3% would optimized specificity (0.966, probability=0.5) in the second-tier screening to exclude 54% of innocent patients, with negative predictive values of 0.93 and AUC of 0.951 ($p<0.001$). The two-tier screening model jointly excluded 4.8% of innocent subjects and prioritized 40% of severe patients for polysomnography.
Conclusions The prediction model is sufficiently accurate and feasible for large-scale population screening.

Keywords Pulse oximetry · Quality-of-life measure · Screening · Sleep-disordered breathing

N.-H. Chen
Sleep Center, Department of Pulmonary and Critical Care Medicine,
Chang Gung Memorial Hospital,
Taoyuan, Taiwan, China

N.-H. Chen
Department of Respiratory Care, Chang Gung University,
Taoyuan, Taiwan, China

M.-C. Chen
Department of Public Health & Biostatistics Consulting Center,
School of Medicine, Chang Gung University,
Taoyuan, Taiwan, China

H.-Y. Li
Department of Otolaryngology, Chang Gung Memorial Hospital,
Chang Gung University,
Taoyuan, Taiwan, China

C.-W. Chen
Department of Psychology, National Chengchi University,
Taoyuan, Taiwan, China

P.-C. Wang
Department of Otolaryngology, Cathay General Hospital,
280 Sec.4 Jen-Ai Rd.,
Taipei, Taiwan, China

P.-C. Wang
Fu Jen Catholic University School of Medicine,
Taipei, Taiwan, China

P.-C. Wang (✉)
Department of Public Health, China Medical University,
Taichung, Taiwan, China
e-mail: drtony@seed.net.tw

Abbreviations

AUC	area under the curve
AASM	American Academy of Sleep Medicine
BMI	body mass index
CESS	Chinese version of Epworth Sleepiness Scale
CSOS	Chinese version of Sleep Outcomes Survey
DI	desaturation index
NPV	negative predictive value
OSAS	obstructive sleep apnea
PPV	positive predictive value
PSG	polysomnography
RDI	respiratory disturbance index
ROC	receiver-operating curve
SDB	sleep-disordered breathing

Introduction

Sleep-disordered breathing (SDB) is a prevalent disorder among the middle-aged that can seriously compromise a patient's quality-of-life [1, 2]. Patients of SDB may suffer from symptoms ranging from snoring to apnea (obstructive sleep apnea syndrome, OSAS). They have higher risks of developing cardiovascular complications and neurocognitive dysfunctions. The SDB can also raise the risk of accidents in traffic and working places [3, 4].

Due to insufficient capacity and long waiting time for overnight polysomnography (PSG), there have been several attempts to develop screening approaches that will simplify diagnostic procedures and reduce costs. Studies based on clinical features [5–7], quality-of-life measures [7–9], and pulse oximetry have been conducted to predict SDB, with some extent of success [5, 10, 11]. Unfortunately, there is little consensus as to the most reliable clinical features that will discriminate the absence or presence of SDB [5, 6].

A simple but cost-effective screening system can help clinicians to prioritize patients for full overnight PSG, especially for those who need immediate surgical or medical attention. For screening methods widely used by researchers, the questionnaire is generally regarded as simple and sensitive, but less specific, while the oximeter is more sophisticated but specific [5–11]. This study combined the merits of these two methods to design a two-tier screening model, using a sensitive questionnaire in the first-tier to exclude innocent subjects and the more specific oximeter in the second-tier to identify severely diseased subjects for early PSG. We hypothesize that a stepwise approach with proper risk stratification strategies can overcome the limitation of individual screening tools to optimize effectiveness of the whole prediction algorithm.

Methods

Patients

In this retrospective study, 355 consecutive patients (aged 18–80 years) who received PSG test in the sleep clinic were examined to evaluate their sleep status. All had a variety of sleep-related complaints that necessitated consult and all provided informed consent for this study. Their demographic and characteristics data were collected upon entry.

The patients were administered with the Chinese versions of sleep outcomes survey (SOS) and Epworth sleepiness scale (ESS) [12, 13]. All surveys were validated and considered statistically equivalent to their original English versions [12, 13]. Permissions to use these surveys were secured and the ethics committee of Chang Gung Memorial Hospital approved this study.

Sleep study

The patients all received standard overnight in-lab polysomnography (Nicolet, Nicolet Inc. Madison, WI) to obtain at least 6 h of sleep data recording. We used nasal pressure combined with oronasal thermocouples to detect airflow. The sleep respiratory disturbance index (RDI) obtained was used as the gold standard for data analysis. RDI was defined as the sum of total apnea and hypopnea episodes per hour of sleep. Apnea episode was defined as cessation of airflow lasting longer than 10 s, whereas hypopnea was defined as ≥30% reduction of oral and nasal flow lasting longer than 10 s with 4% desaturation. Based on the definition of the American Academy of Sleep Medicine (AASM), patients with RDI >5 episodes/h had OSAS and over 30 episodes/h were severe cases [14]. To improve the clinical relevance of the screening algorithm, RDI of 5 and 30 episodes/h were used as cut-off points to dichotomize variables for further analyses.

Quality-of-life measures

The Chinese version of the SOS and ESS were used for the first-tier screening. Both of them were outcome measures to evaluate the health impact and treatment effectiveness for adults with SDB and had been previously translated and validated by the authors [14, 15].

Chinese version of Snore Outcomes Survey

The SOS is a validated outcome measure that evaluates the health impact and treatment effectiveness of adults with SDB and snoring [15]. It contains eight items that evaluates the duration, severity, frequency, and consequences of problems associated with SDB on a Likert scale, each with

five-to-six response options. The SOS total score is transformed into a scale ranging from 0 (worst) to 100 (best). The Chinese version of SOS was translated and validated by the authors in a previous study, with good correlation to PSG results [12]. Patients with SOS scores of 55 or less are considered to be a loud snorer.

Chinese version of the Epworth Sleepiness Scale

The eight-item ESS is widely used for evaluating adults on the average sleep propensity in daily life [16]. Scores for each item range from 0 to 3 and the total Epworth score ranges from 0 to 24 (lowest to highest sleep propensity). The reliability, unitary structure and validity of the ESS are supported by experimental evidences in distinguishing the excessive daytime sleepiness of SDB from that of normal subjects [16]. Patients with ESS scores higher than 12 are considered to have pathologic sleepiness. Chinese version of ESS was also translated and validated by the authors in previous study, with good correlation to PSG results. [13]

Pulse oximetry

Pulse oximetry is frequently used in the clinical hospital setting to measure the oxygen saturation of patients. It is a small and sophisticated device clipped on the fingertip to record oxygen saturation. Desaturation 2%, 3%, or 4% mean a 2%, 3% or 4% oxygen saturation drop from previous recording. The number of desaturation events of 2%, 3%, and 4% was recorded in selected cases overnight. Desaturation index of 2%, 3%, and 4% was defined as the number of the episodes of 2%, 3%, and 4% desaturation over the hours of sleep recording.

The Pulsox-3i (Minolta Co., Ltd, Osaka, Japan) was chosen as oxygen saturation monitoring in the second-tier screening. Patients had Pulsox-3i monitoring and recording simultaneously with standard polysomnography. The sleep oxygen desaturation events were retrieved and stored using Pulsox-3 DS-3 Data Analysis (Minolta Co., Ltd, Osaka, Japan) software.

Statistical analysis

Association between RDI and patient demographics and survey scores

The Spearman correlation coefficient was used to examine the association between RDI, patient demographics, and survey scores.

First-tier screening modeling

According to the definition of AASM, RDI was dichotomized as "non-obstructive sleep apnea syndromes (non-OSAS)" for RDI <5 vs. "obstructive sleep apnea syndromes (OSAS)" for RDI≧5. Multiple logistic regression was applied to examine the possibility of "having OSAS" using the variables chosen from the demographic characters that were significantly association with RDI, such as gender, age, body mass index (BMI), Chinese version of the Snore Outcomes Survey (CSOS), and Chinese version of the Epworth Sleepiness Scale (CESS). Using these demographic characters against OSAS (RDI≧5), the receiver-operating characteristic (ROC) curve was applied to determine the diagnostic thresholds for CSOS/CESS combinations that were more likely to differentiate "OSAS" from "non-OSAS".

The area under curve (AUC) was calculated. CESS and CSOS were dichotomized simultaneously at various cut points and were entered into the estimated logistic regression model with age, gender, and BMI, and the patient was considered a "OSAS" case when the estimated probability from multiple logistic regression was greater than 0.5. As a result, the sensitivity, specificity, positive and negative predictive values (PPV and NPV) were derived based on different CSOS and CESS combination. The bootstrapping technique was used for cross-validation since it is impossible to collect more new samples to evaluate the validation of our predictive logistic regression, and it was also helpful to identify the cut-off point, the optimal CSOS and CESS combination which would yield relatively higher sensitivity of this model to include as many OSAS patients as possible.

Second-Tier screening modeling

In the second-tier screening, the pulse oximeter was used. Although it was easy to use and clinically available than standard PSG, it still took the whole night to record. It was also more sophisticated to calculate than the questionnaire. For cost-effectiveness reasons and to achieve a power of 80% with a significance level of 5%, we performed power analysis based on a preliminary study which showed 85% of patients from first-tier were correctly identified as cases. In order to demonstrate a difference between our preliminary study (85%) and 75% in other literature [17], at least 98 subjects were required. So, we randomly selected 100 possible OSA patients that were identified of having OSAS (predicted positive for RDI≧5) in first-tier screening for pulse oximeter examination. Binary RDI in the second-tier screening was defined as "severe OSAS" with RDI≧30 against "non-severe OSAS" with RDI <30. The area under the curve (AUC) of ROC for desaturation index (DI, episodes/h) of 2% (DI2), 3% (DI3), and 4% (DI4) desaturation were calculated and DI3 was best fitted to predict the severity of OSAS. Logistic regression was used to evaluate the relationship between "severe OSAS" and DI3.

The sensitivity, specificity, and PPV and NPV of DI3 were also tabulated. The optimal DI3 cut-off point

Table 1 Correlations between RDI and patients' demographics

Variable	Mean±SD	γ (p value)[a]
Age (years old)	44.7±11.3	0.14 (0.008)
BMI (kg/m^2)	27.4±4.1	0.309(<0.001)
CSOS	44.9±15.3	−0.362(<0.001)
CESS	10.9±5.2	0.248(<0.001)

The mean RDI is 23.31±32.19 episodes/h of female and 40.21±29.28 episodes/h of male, the p value of t statistic from two-sample t test is less than 0.0001

CESS Chinese version of Epworth Sleepiness Scale, *CSOS* Chinese version of Snore Outcomes Survey

[a] Spearman's correlation coefficient

yielded relatively higher specificity of the second-tier screening model, without sacrificing sensitivity, to exclude as many "non-severe OSAS" patients as possible. Similarly, the bootstrapping was used in the 2-tier screening for cross-validation.

Data management

All data were stored in Access 7.0 database (Microsoft, Redmond, Seattle) and analyzed using the SAS software package (SAS Institute, Cary, North Carolina). A p value <0.05 was considered statistically significant.

Results

Study population

The initial study group consisted of 355 patients. There were 312 (87.9%) males and 43 (12.1%) females. The mean RDI was 38.3±29.9 episodes/h. The mean RDI was 40.21±29.28 episodes/h for men and it was significantly higher than females (23.31±32.19 episodes/h) with a p value <0.001 using a two-sample t test. The distribution of RDI was: 48 (13.5%) less than 5 episodes/h, 69 (19.4%) between 5 and less than 15 episodes/h, 52 (14.6%) between 15 to less than 30 episodes/h, and 186 (52.4%) equal or greater than 30 episodes/h. The demographic data and the severity of RDI in these patients are shown in Table 1.

First-tier screening prediction

Estimated probability of OSAS

Gender, age, BMI, CESS, and CSOS were used to predict the probability of having OSAS (RDI≥5). Multiple logistic regression was used to predict the probability of having OSAS (RDI≥5) in the first-tier screening and the results are shown in Table 2. Based on this model, the probability of having OSAS was:

$$\widehat{P}(\text{having OSAS}) = \frac{e^{-5.935+1.096X_{sex}+0.064X_{age}+0.264X_{BMI}+0.039X_{ESS}-0.062X_{SOS}}}{1+e^{-5.935+1.096X_{sex}+0.064X_{age}+0.264X_{BMI}+0.039X_{ESS}-0.062X_{SOS}}}$$

For example, a 50-year-old male with BMI of 30, CESS score 12, and CSOS 50 would have a predicted probability of having OSAS of 0.97

Cut-off point and model predictability

The ROC curve of the first-tier screening is shown in Fig. 1. The sensitivity, specificity, and PPV and NPV of different possible CSOS/CESS combinations in predicting OSAS are shown in Table 3. The combination of "CSOS score of 55 and CESS score of 9" was the optimal cut-off point that yielded relatively higher sensitivity (0.603) and specificity in this first-tier screening model.

Second-tier screening prediction

Study population

The second-tier screening study group consisted of 100 randomly selected patients after power analysis from the predicted positive population (RDI≥5, presumably having

Table 2 Multiple logistic regression model to predict the probability of having OSAS (RDI≥5) in the first-tier screening

Variables		Estimated β	Odds ratio (OR)	95% CI for OR	p value[a]
Gender	Male	1.096	2.99	1.05–8.55	0.041
	Female		1		
Age		0.064	1.07	1.03–1.11	0.001
BMI		0.264	1.30	1.15–1.47	<0.001
CESS		0.039	1.04	0.96–1.13	0.34
CSOS		-0.062	0.94	0.92–0.97	<0.001

The intercept was −5.935 in this multiple logistic regression.

[a] Adjusted p value indicates the significance of the parameters by multiple logistic regression.

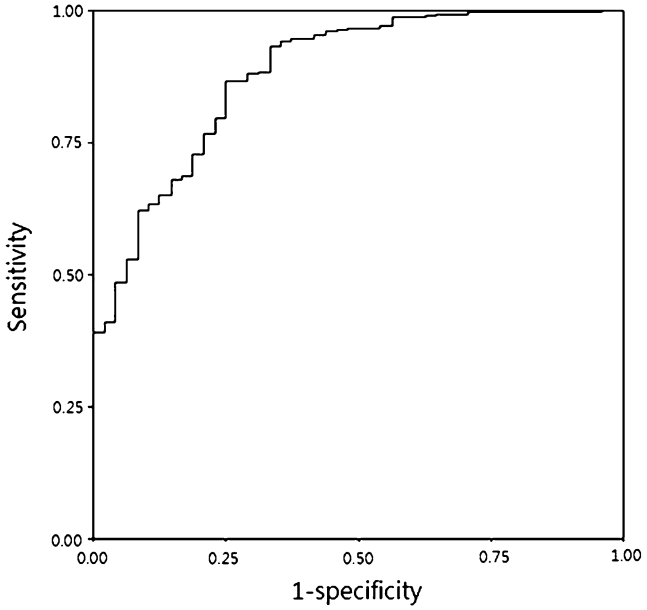

Fig. 1 Receiver-operating characteristic curve using gender, age, BMI, CSOS, and CESS against OSAS (RDI≧5). (area under curve 0.88, standard error 0.026, Z 14.62, $p<0.001$)

OSAS, $n=337$) of the first-tier screening. There were 83 (83%) males and 17 (17%) females, with mean age of 43.3±11.5 years and BMI of 26.5±3.7. The mean RDI was 32.2±28.4 episodes/h. Nineteen (19%) patients did not have OSAS (RDI<5 episodes/h), while 21 (21%) had RDI≥5 but <15 episodes/h, 18 (18%) had RDI≥15 but <30 episodes/h, and 42 (42%) have RDI>30. The mean DI3 of this cohort was 22.3±21.5 episodes/h.

Desaturation index

The ROC curve using DI3 against severe OSAS (RDI>30) showed that the area under the curve (AUC) was 0.951 (standard error=0.024, $Z=18.792$, $p<0.001$). The ROC curves using DI2 and DI4 against severe OSAS (RDI>30) showed that the AUC was 0.942 (standard error=0.027, $Z=16.3763$, $p<0.001$) for DI2, and similarly, the AUC was 0.942 (standard error=0.027, $Z=16.3763$, $p<0.001$) for DI4. The DI3 was therefore chosen as the desaturation index in this study (Fig. 2).

Probability of having severe OSAS

The logistic regression model showed that DI3 positively related to the possibility of having severe OSAS (RDI>30; estimated beta=0.170, $p<0.001$).

The probability of having severe OSAS was:

$$\widehat{P}(\text{having severe OSAS}) = \frac{e^{-3.627+0.170X_{DI3}}}{1+e^{-3.627+0.170X_{DI3}}}$$

Cut-off point and model predictability

The sensitivity, specificity, and PPV and NPV of DI3 in predicting severe OSAS are shown in Table 4.

The DI3 of 30 optimized specificity (0.966) of the second-tier screening model to exclude as many non-severe OSAS patients as possible (Table 4). With NPV of 0.93 (54/58) and calculated probability of 0.5, this second-tier screening model excluded as many patients ($n=54$, 54%) as possible that did not have severe OSAS.

Table 3 Relative discriminatory powers of CESS and CSOS

Surveys' scores	Sensitivity	Specificity	PPV%	NPV%
CESS≧9, CSOS≦40	0.381	0.833	93.60	17.39
CESS≧9, CSOS≦45	0.495	0.792	93.83	19.69
CESS≧9, CSOS≦50	0.541	0.75	93.26	20.34
CESS≧9, CSOS≦55	0.603	0.729	93.43	22.29
CESS≧10, CSOS≦40	0.358	0.917	96.49	18.26
CESS≧10, CSOS≦45	0.453	0.875	95.86	20.00
CESS≧10, CSOS≦50	0.498	0.833	95.00	20.51
CESS≧10, CSOS≦55	0.538	0.813	94.83	21.55
CESS≧11, CSOS≦40	0.326	0.917	96.15	17.53
CESS≧11, CSOS≦45	0.407	0.896	96.15	19.11
CESS≧11, CSOS≦50	0.437	0.854	95.04	19.16
CESS≧11, CSOS≦55	0.472	0.833	94.77	19.80
CESS≧12, CSOS≦40	0.296	0.958	97.85	17.56
CESS≧12, CSOS≦45	0.375	0.938	97.46	18.99
CESS≧12, CSOS≦50	0.401	0.917	96.85	19.30
CESS≧12, CSOS≦55	0.437	0.896	96.40	19.91

CESS Chinese version of Epworth Sleepiness Scale, *CSOS* Chinese version of Snore Outcomes Survey

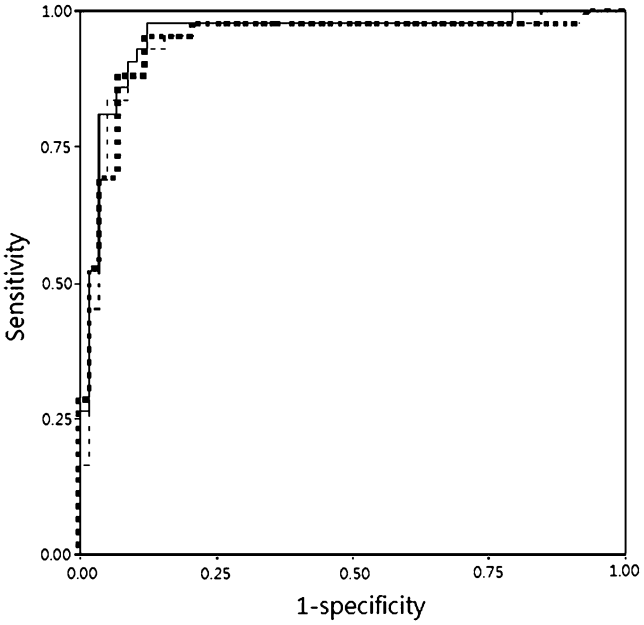

Fig. 2 Receiver-operating characteristic curve using DI3 (*solid line*) against severe OSAS (RDI≧30; area under curve 0.951, standard error=0.024, Z=18.792, p<0.001). For DI2 (*dashed line*) and DI4 (*dashed-dot line*), the AUC are identical (0.942, with standard error=0.027, Z=16.3763, p<0.001)

The Upper panel of Table 5 (model predictability) shows the predicted positive and predicted negative values from the proposed model for the first-tier screening. It was calculated by plugging in the parameters in the multiple logistic regression model to obtain the estimated probability of having OSAS (RDI≥5). If the estimated probability was >0.5, it was considered a case, and vice versa. As a result, the number of true positive was compared with the estimated positive, and the number of true negative with the estimated negative. Similarly, the predicted positive and negative listed in the lower panel of Table 6 were based on the proposed model for the second-tier screening model. A calculated probability of 0.6 included as many patients (n=337, 94.93%) as possible that had PPV 0.997 (306/307) for the diagnosis of OSAS (Table 5).

Table 4 Relative discriminatory powers of DI3 for severe OSAS (RDI≧30)

DI3 (episodes/h)	Sensitivity	Specificity	PPV%	NPV%
5	0.976	0.448	75.93	97.83
10	0.976	0.655	78.43	95.92
20	0.905	0.914	81.63	96.08
30	0.571	0.966	82.98	94.34
40	0.357	0.983	84.78	94.44
50	0.075	0.994	86.67	94.55

DI3 Desaturation index 3, desaturation more than 3%/h

Table 5 Model predictability for first-tier screening

N=355	Predicted positive	Predicted negative
True positive (n=307)	hit 306	miss 1
True negative (n=48)	false alarm 31	hit 17

The accuracy of the presented two-tier model is confirmed by cross-validation using the bootstrapping technique [18]. Given the probability of greater than 0.5, the correct prediction rates are 0.92 (minimum–maximum, 0.88–0.96), 0.91 (minimum–maximum, 0.83–0.96) for first- and second-tier screening models, respectively.

Discussion

SDB is a major quality-of-life issue. Patients with SDB often show increased difficulty in concentrating, learning new tasks, and performing repetitive tasks. Lindberg and others [3, 4, 19] report that OSAS patients have higher risk of occupational and traffic accidents. In order to reduce professional liability, it is important to identify patients with the highest risks of severe SDB as early as possible. This study attempts to develop a cost-effective screening approach in order to prioritize candidates for early PSG.

Combined with clinical information, standard sleep quality-of-life measures are widely used to describe the prevalence of snoring, observed apneas, daytime sleepiness in the general population, and the relationships of sleep disturbances to health [7, 20]. It is generally regarded that questionnaires alone are not sufficient to discriminate patients with SDB, although these may be useful in prioritizing patients for split-night PSG. The reported sensitivity of questionnaires varies from 72% to 96% in predicting OSAS, but the specificity is as low as 13% to 54% [6, 9, 17]. The highest specificity of 0.77 reported from a Berlin questionnaire has been challenged because of underestimation using a four-channel sleep monitor as the validated gold standard [8].

Sleepiness and snoring are two major clinical symptoms in SDB patients. This study combines the widely circulated measurement tools, CESS and CSOS, which cover these two important but distinct dimensions (sleepiness and snoring) in SDB. Compared to other studies that use only indices or symptom scores to evaluate patients [9, 17, 20],

Table 6 Model predictability for second-tier screening

N=100	Predicted positive	Predicted negative
True positive (n=42)	hit 36	miss 6
True negative (n=58)	false alarm 4	hit 54

CSOS and CESS are both well validated by our group and show good associations to SDB severity [12, 13, 21]. The applicability in using CSOS and CESS for community SDB screening was tested previously [22]. With CSOS >55 and CESS >9, a sensitivity of 0.603 and specificity of 0.729 can be attained, which is the optimal cut-off value that provides good positive predicted values and highest negative prediction.

By using the regression model, the probability of having disease can be easily calculated by this formula. For example, a 50-year-old male with BMI of 30, CESS score 12, and CSOS 50 will have a predicted probability 0.97 of having OSAS. Physicians then have to make clinical judgment for the second-tier screening based on this calculation. After the second-tier screening with similar calculations, patients will be prioritized for further examination (PSG) if the risks of having severe disease is high as identified by the algorithm we developed.

The AUC of the ROC curve reaches the level of 0.88, which is compatible with the reported data of 0.55–0.83 from similar studies in literatures [7, 9, 23, 24]. With a calculated probability of 0.6, as many patients (94.93%) as possible can be included that probably have OSAS. Excluded subjects (estimated RDI<5) are "least likely" to have the disease and their chances of having even very mild sleep respiratory disturbance is very low. Using this algorithm, 17 patients will be exempted from PSG because their risks of having OSAS are so low and only one (out of 355 patients) with true OSAS will be missed (Table 5).

Pulse oximetry is another frequently used tool for screening OSAS with great economical benefit [10, 11]. The report from the Technology Assessment Task Force of the Society of Critical Care Medicine in 1993 indicate that pulse oximetry is a non-invasive tool to measure oxygen saturation with a high degree of accuracy over a range of 80–100% saturation [11]. The 1995 British Thoracic Society Report concludes that pulse oximetry criteria is highly specific when positive (specificity 100%), but may miss patients with hypopneic arousal without significant oxygen desaturation (sensitivity 31%) [24]. In the second-tier screening, the strategy is to increase the screening specificity. Even though the differences among DI2, DI3, and DI4 are small, the highest AUC of 0.951 indicates that DI3 is the ideal threshold against RDI≧30.

The desaturation index of 3% used in the second-tier screening yields a sensitivity of 0.57 and a specificity of 0.96, which are comparable to those reported by Golpe et al. (for RDI>40.5, specificity 97%) [25]. With a calculated probability of 0.5, 60% of patients who are not likely to have severe OSAS can be identified, while the excluded patients need not to be prioritized for PSG. Using this algorithm, 36 (out of 100) patients will definitely need early PSG because of high risks of having severe OSAS, while four patients will be recruited for unnecessary sleep study (Table 6).

Since neither quality-of-life measures nor pulse oximeter is individually ideal, some authors advocate the usefulness of pulse oximetry to establish the diagnosis of OSAS and highlight the value of clinical scoring to improve the sensitivity of screening tools [5]. This study sought to optimize the prediction algorithms by developing a step-wise, two-tiered screening model. Using CESS and CSOS, the study can exclude 4.8% (18 out of 355, including one false negative) of patients from PSG testing in the first-tier screening since their risks of having OSAS is low. Using pulse oximetry, 40% (40 out of 100, including four false alarm) of patients can be prioritized for early PSG testing since their risks of having severe OSAS are high. These cost-effective data are equivalent to those reported by Keenan et al. [26] and Gurubhagavatula et al. [23].

However, the cost-effectiveness is highly dependent on the prevalence of OSAS in the study population. When the two-tier model is applied to the general population, rather than to this validation population, we expect more targeted patients will be identified to achieve screening objectives (excluding low-risk patients and prioritizing high-risk patients with greater cost-effectiveness ratio).

Conclusion

In conclusion, the two-tier screening model can jointly exclude 4.8% of innocent subjects from sleep studies and can prioritize up to 40% of severe OSAS patients to receive complete in-laboratory PSG with 0.603 sensitivity for OSAS and 0.966 specificity for severe OSAS. Even though this model may not identify other causes of sleep disorders, the prediction algorithm is sufficiently accurate for patients with sleep complaints. Quality-of-life and pulse oximetry information can help clinicians to identify patients who need early PSG diagnosis. It could also be a cost-effective solution to assist community or occupational SDB screening.

References

1. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S (1993) The occurrence of sleep-disordered breathing among middle-age adults. N Engl J Med 328:1230–1235
2. Lindberg E, Elmasry A, Gislason T, Janson C, Bengtsson H, Hetta J, Nettelbladt M, Boman G (1999) Evolution of sleep apnea syndrome in sleepy snorers: a population-based prospective study. Am J Respir Crit Care Med 159:2024–2027
3. Young T, Blustein J, Finn L, Palta M (1997) Sleep-disordered breathing and motor vehicle accidents in a population-based sample of employed adults. Sleep 20:608–613

4. Powell NB, Schechtman KB, Riley RW, Li K, Guilleminault C (2002) Sleepy driving: accidents and injury. Otolaryngol Head Neck Surg 126:217–227
5. Schäfer H, Ewig S, Hasper E, Lüderitz B (1997) Predictive diagnostic value of clinical assessment and non-laboratory monitoring system recordings in patients with symptoms suggestive of obstructive sleep apnea syndrome. Respiration 64:194–199
6. Tami TA, Duncan HJ, Pfleger M (1998) Identification of obstructive sleep apnea in patients who snore. Laryngoscope 108:508–513
7. Pouliot Z, Peters M, Neufeld H, Kryger MH (1997) Using self-reported questionnaire data to prioritize OSA patients for polysomnography. Sleep 20:232–236
8. 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 131:485–491
9. Rowley JA, Aboussouan LS, Badr MS (2000) The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. Sleep 23:929–938
10. Pradhan PS, Gliklich RE, Winkelman J (1996) Screening for obstructive sleep apnea in patients presenting for snoring surgery. Laryngoscope 106:1393–1397
11. A model for technology assessment applied to pulse oximetry (1993) The technology assessment task force of the society of critical care medicine. Crit Care Med 21:615–624
12. Chen NH, Li HY, Gliklich RE, Chu CC, Liang SC, Wang PC (2002) Validation assessment of the Chinese version of the Snore Outcome Survey (SOS). Qual Life Res 11:601–607
13. Chen NH, Johns MW, Li HY, Chu CC, Liang SC, Shu YH, Chuang ML, Wang PC (2002) Validation assessment of the Chinese version of the Epworth Sleepiness Scale (ESS). Qual Life Res 11:817–821
14. Hirshkowitz M, Kryger MH (2005) Monitoring techniques for evaluating suspected sleep-disordered breathing. In: Kryger MH, Roth T, Dement WC (eds) Principle and practice of sleep medicine, 4th edn. Elsevier, Philadelphia, pp 1378–1379
15. Gliklich RE, Wang PC (2002) Validation of the Snore Outcomes Survey for Patients with sleep-disordered breathing. Arch Otolaryngol Head Neck Surg 128:819–824
16. Johns MW (1993) Daytime sleepiness, snoring and obstructive sleep apnea. The Epworth Sleepiness Scale. Chest 103:30–36
17. Herer B, Roche N, Carton M, Roig C, Poujol V, Huchon G (1999) Value of clinical, functional, and oximetric data for the prediction of obstructive sleep apnea in obese patients. Chest 116:1537–1544
18. Efron B (1994) How biased is the apparent error rate of a prediction rule? J Am Stat Assoc 81:463–470
19. Ulfberg J, Carter N, Edling C (2000) Sleep-disordered breathing and occupational accidents. Scand J Work Environ Health 26:237–242
20. Enright PL, Newman AB, Wahl PW, Manolio TA, Haponik EF, Boyle PJR (1996) Prevalence and correlates of snoring and observed apneas in 5, 201 older adults. Sleep 19:531–538
21. Gurubhagavatula I, Maislin G, Nkwuo JE, Pack AI (2004) Occupational screening for obstructive sleep apnea in commercial drivers. Am J Respir Crit Care Med 170:371–376
22. Chen NH, Shih TS, Wang PC, Lin Y, Chen CW, Chen CJ, Lin JD, Li HY, Gliklich RE (2007) Sleep quality and the prevalence of related breathing problems in Taiwanese factory workers. Fu Jen J Med 5:67–80
23. Gurubhagavatula I, Maislin G, Pack AI (2001) An algorithm to stratify sleep apnea risk in a sleep disorders clinic population. Am J Respir Crit Care Med 164:1904–1909
24. Ryan PJ, Hilton MF, Boldy DA, Evans A, Bradbury S, Sapiano S, Prowse K, Cayton RM (1995) Validation of British Thoracic Society guidelines for the diagnosis of the sleep apnea/hypopnea syndrome: can polysomnography be avoided? Thorax 50:972–975
25. Golpe R, Jimenez A, Carpizo R, Cifrian JM (1999) Utility of home oximetry as a screening test for patients with moderate to severe symptoms of obstructive sleep apnea. Sleep 22:932–937
26. Keenan SP, Anderson B, Wiggs B, Ryan CF, Fleetham JA (1993) The predictive accuracy of home oximetry in patients with suspected obstructive sleep apnea. Sleep 16:S133–S134