SCIENTIFIC INVESTIGATIONS

Using the Berlin Questionnaire to Predict Obstructive Sleep Apnea in the General Population

Adeline Tan, MBBS; Jason D.C. Yin, MSc; Linda W.L. Tan, MPH; Rob M. van Dam, PhD; Yan Yi Cheung, BSc; Chi-Hang Lee, MD

1Department of Respiratory Medicine, Ng Teng Fong General Hospital, Jurong Health Services, Singapore; 2Saw Swee Hock School of Public Health, National University of Singapore and National University Health System; 3Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System; 4Department of Cardiology, National University Heart Centre, Singapore.

Study Objectives: The Berlin questionnaire is a self-administered questionnaire that was developed to identify subjects with obstructive sleep apnea (OSA) in primary care settings. This study evaluated the performance of the questionnaire to predict OSA in the general population.

Methods: A sample of 242 subjects in a population-based cohort completed a home-based sleep study with an Emblettta device (type 3 monitor). Subjects completed the Berlin questionnaire on the evening just prior to the sleep study. The sleep studies were manually scored according to the 2012 American Academy of Sleep Medicine (AASM) criteria.

Results: The prevalence of moderate-to-severe and severe OSA defined as apnea-hypopnea index (AHI) of ≥ 15 and ≥ 30 was 28.1% and 10.7%, respectively. Seventy-nine subjects (32.6%) were classified as high risk according to the Berlin questionnaire. The sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) of the questionnaire to predict an AHI ≥ 15 was 58.8%, 77.6%, 82.9%, and 50.6%, respectively. The area under the receiving operator characteristic (ROC) curve for moderate-to-severe OSA was 0.682. When used to predict an AHI ≥ 30, the sensitivity of the questionnaire increased to 76.9% with a small drop in specificity to 72.7%. The corresponding NPV, PPV, and area under the ROC curve of the questionnaire to predict severe OSA were 96.3%, 25.3%, and 0.748, respectively.

Conclusions: The Berlin questionnaire may have utility in the general population setting as a screening tool for OSA in view of its good sensitivity and high NPV in ruling out severe OSA.

Keywords: Berlin questionnaire, general population, obstructive sleep apnea, screening


INTRODUCTION

Obstructive sleep apnea (OSA) is a condition characterized by repeated upper airway collapse during sleep, leading to desaturations and arousals. It is associated with negative health consequences such as excessive daytime sleepiness; cardiovascular and metabolic morbidities such as hypertension, heart failure, and diabetes; and increased all-cause mortality. Recent population-based studies have shown an alarming increase in the prevalence of OSA, with reported rates of moderate-to-severe OSA (defined as an apnea-hypopnea index (AHI) ≥ 15 events/h) up to 50%. The rise in prevalence is likely due to a combination of factors, such as the obesity epidemic, increased sensitivity of sleep study recording sensors and reduced stringency of scoring criteria as a result of updates. The majority of OSA subjects in these studies were previously undiagnosed, suggesting a large burden of sleep apnea subjects that remains unrecognized in the general population. OSA subjects in the general population are unique in two aspects. First, the prevalence of OSA is lower as compared to referral populations such as preoperative or sleep clinic patients. Second, OSA subjects in the general population are usually asymptomatic or have unrecognized symptoms, meaning they are less likely to present to a healthcare provider for further evaluation. Hence, screening questionnaires could be particularly useful in risk stratifying large numbers of subjects in an unselected community-based population for further diagnostic testing.

The Berlin questionnaire is a survey that has been used to predict OSA have been validated mainly in referral populations where the prevalence of OSA is high. The Berlin questionnaire was initially developed as a tool to identify patients who are likely to have OSA in the primary care setting. Our findings show that the Berlin questionnaire can be used as a screening tool in the general population in view of its good sensitivity and high negative predictive value in ruling out severe OSA. However, the positive predictive value is low. Subjects from a population-based cohort who are classified as high risk on the questionnaire can be prioritized to undergo further testing to confirm the presence of OSA.
primary care setting. The questionnaire is self-administered and consists of 10 questions in three categories related to the presence and severity of snoring, frequency of daytime sleepiness, and the presence of obesity or hypertension. A high-risk classification on the Berlin questionnaire in the primary care setting was associated with a sensitivity of 86% and specificity of 77% for a respiratory disturbance index (RDI) greater than 5 and a sensitivity of 54% and specificity of 97% for a RDI greater than 15, respectively. To date, there have been limited studies on the performance of the Berlin questionnaire in the general population. This study aimed to examine the diagnostic properties of the Berlin questionnaire to predict OSA in a general population setting using the latest 2012 American Academy of Sleep Medicine (AASM) scoring criteria.

METHODS

For the current study, subjects were selected from the Singapore Health Study 2012 (SH2012), a population-based cohort that aimed to study the prevalence of non-communicable diseases such as hypertension and diabetes in Singapore, a multi-ethnic Asian nation. The sampling strategy has been previously described. Briefly, inclusion criteria included Singaporeans aged 18–79 y as of July 2012. Exclusion criteria were mental retardation, pregnancy, or stroke resulting in loss of speech or bedridden status. A total of 1,984 subjects completed the Pittsburgh Sleep Quality Index (PSQI). There were 336 snorers and 1,648 nonsnorers based on the PSQI. All snorers and 500 of the 500 nonsnorers were randomly selected and contacted via phone to participate in a home-based sleep test to study the prevalence of OSA in the community. Due to logistic constraints, only 332 of the 500 nonsnorers were contacted. One hundred thirty-seven of 336 snorers (40.7%) and 119 of 332 nonsnorers (35.8%) agreed to a sleep study. Fourteen sleep studies were rejected due to insufficient recording time of at least 4 h, leaving a total of 242 successful sleep studies in the final analysis. The signals recorded included airflow through a nasal pressure sensor, oxyhemoglobin saturation, pulse rate from pulse oximetry, and thoracic and abdominal movements through respiratory inductance plethysmography belts. The sleep recordings were manually scored by a certified polysomnologist and subsequently reviewed by a sleep physician. Both the polysomnologist and sleep physician were blinded to the subjects’ demographic characteristics and questionnaire results. Respiratory events were scored according to the latest 2012 AASM scoring criteria. Apneas were defined as a ≥ 90% reduction in airflow for at least 10 sec, and hypopneas were defined as a ≥ 30% reduction in airflow for at least 10 sec with an oxyhemoglobin decrease of ≥ 3%. The AHI was then calculated as the number of apneas and hypopneas per hour of recording time. We defined OSA severity using thresholds of greater than or equal to 15 and 30 to represent moderate to severe and severe OSA, respectively.

Statistical Analysis

Demographics, anthropometric, sleep test, and questionnaire data for the study sample were summarized as means and standard deviation for continuous variables and percentages for categorical variables. The sensitivities, specificities, positive predictive value (PPV) and negative predictive value (NPV), positive and negative likelihood ratios, and area under the receiving operator characteristic (ROC) curve were calculated for a high-risk Berlin questionnaire classification to predict OSA with different AHI cutoffs of greater than or equal to 15 and 30, respectively, using AASM 2012 scoring criteria. The statistical program used was R version 3.2.1.

RESULTS

All 242 subjects completed a Berlin questionnaire in a language of their choice. Two hundred eleven questionnaires (87.2%) were completed in English, whereas only 20 subjects (8.3%) completed the Chinese version and 11 subjects (4.5%) completed the Malay version. The Tamil version was not used. The demographics and sleep study results of the study sample are...
shown in Table 1. Using Asian cutoff points to define the overweight and obese categories, 99 subjects (40.9%) were overweight (BMI 23–27.5 kg/m²) and 81 subjects (33.4%) were obese (BMI > 27.5 kg/m²). The prevalence of moderate to severe OSA and severe OSA was 28.1% and 10.7%, respectively. Table 2 shows the breakdown of positive responses in the individual categories of the Berlin questionnaire. Overall, 79 subjects (32.6%) were classified as high risk by the questionnaire. A total of 110 subjects (45.5%) scored positively in category one, which evaluates snoring. This is because we oversampled snorers and 137 subjects (56.6%) were already identified as snorers based on the PSQI. Sixty subjects (24.8%) scored positively in category two, which determines the presence of excessive daytime sleepiness, whereas 79 subjects (32.6%) scored positively in category three, which assess presence of hypertension or obesity (defined as BMI > 30 kg/m²).

Twenty-five subjects (10.3%) reported a history of drowsy driving. Among the subjects who reported drowsy driving, 15 (60.0%), 17 (68.0%), and 11 subjects (44.0%) had a high risk score in categories 1, 2 and 3, respectively. A high risk score in category 2 was associated with a positive report of drowsy driving (odds ratio 8.60, 95% confidence interval 3.49–21.24, p < 0.001). However, there was no association between a high risk score in category 1 or category 3 and a positive report of drowsy driving.

Table 3 depicts the diagnostic properties of the Berlin questionnaire to predict moderate to severe OSA and severe OSA. A high-risk classification on the Berlin questionnaire predicted an AHI ≥ 15 with a sensitivity of 58.8%, a specificity of 77.6%, a NPV of 82.9% and a PPV of 50.6%. The area under the ROC curve was 0.682. When used to predict an AHI ≥ 30, the sensitivity of a high risk classification rose to 76.9%, specificity remained moderately high at 72.7% and the corresponding NPV increased to 96.3% whereas PPV dropped to 25.3%. The area under the ROC curve also improved to 0.748.

DISCUSSION

To our knowledge, this is the first population-based study to look at the performance of Berlin questionnaire to predict OSA using the latest AASM 2012 criteria. We found that the questionnaire may have utility as a screening tool in the general population.
population setting in view of the good sensitivity (76.9%) and high NPV (96.3%) in ruling out severe OSA. When used to predict an AHI ≥ 15, we found that the sensitivity was moderate at 58.8%, with a specificity of 77.6%, a moderately high NPV of 82.9%, and a low PPV of 50.6%. The screening properties of the Berlin questionnaire improved significantly when used to predict an AHI ≥ 30, with sensitivity, specificity, NPV, and PPV of 76.9%, 72.7%, 96.3%, and 25.3% respectively. A recent Switzerland study found that the prevalence of moderate to severe sleep-disordered breathing in the general population was 23.4% in females and 49.7% in males. In addition, the authors also found that only the upper quartile for AHI (>20.6 events/h) was associated with comorbidities such as hypertension and diabetes. Another population-based study performed in Iceland found that approximately 20% of subjects had moderate to severe OSA and were largely asymptomatic. Together, these studies suggest that the cutoffs for treatment of OSA may need to be adjusted further upward from the current AHI cutoff of 15. In that context, ruling out severe OSA in a general population setting may be more clinically relevant. However, it is important to take note that despite the high NPVs we found, the PPVs are low, ranging from 20.3% to 50.6%, depending on the cutoff of AHI used. Hence, the questionnaire can only be used primarily as a screening tool to rule out OSA. Subjects who are classified as high risk will require further diagnostic testing.

Our findings are similar to the study by Kang et al., who found that the sensitivity and specificity of the Berlin questionnaire to predict an AHI ≥ 15 was 89.0% and 63.0%, respectively. No NPV or PPV were reported for that study but 26 out of 101 subjects (25.7%) who underwent sleep studies had an AHI ≥ 15, which is close to the prevalence of 28.1% that we found. Kang et al. concluded that the questionnaire could be a useful tool to prioritize subjects at high risk of OSA in the general population. In contrast, Hrubos-Strom et al. found that screening properties of the Berlin questionnaire in the general population was suboptimal, with a sensitivity of 37.2% and a specificity of 84.0% to predict an AHI ≥ 15. One of the possible reasons could be that the prevalence of OSA in this Norwegian sample was very low, with an estimated prevalence of moderate to severe OSA of only 8.0%. Populations with a higher prevalence of OSA are likely to have more subjects who are symptomatic or have associated comorbidities, giving rise to higher sensitivities. Apart from differences in the types of sleep studies performed and the scoring criteria used, we also postulate that the variation in the diagnostic properties of the Berlin questionnaire in these population-based studies performed to date could be because the questionnaire is largely composed of subjective questions. A person’s assessment of his or her snoring severity and daytime fatigue is personal and prone to bias. In comparison, the STOP-Bang questionnaire, which is another tool designed to predict OSA, consists of four simple yes/no questions related to snoring, tiredness, observed apneas, and high blood pressure and four demographic queries that include BMI, age, neck circumference, and sex. A recent meta-analysis looking at the STOP-Bang questionnaire in various populations found that the diagnostic properties of STOP-Bang were generally consistent with high sensitivities and low specificities, which is in contrast to the large variation reported in different studies looking at the predictive accuracy of the Berlin questionnaire.

Netzer et al. administered the Berlin questionnaire to 6,223 subjects in a broad range of primary care settings in the United States and Europe. Similar to our study, they reported that 32.3% of subjects were at high risk of OSA. The proportion of subjects that qualified for high risk in category 1 (snoring) and category 2 (daytime sleepiness) in their study were 43.4% and 24.7%, respectively, which again parallels the values that we found. However, for category 3 (hypertension or obesity), they reported a high-risk proportion of 41.9% whereas we found a lower value of 32.6%. This likely reflects the lower obesity rates in Asia. In particular, we would like to draw attention to the fact that 10.3% of our subjects reported drowsy driving, a percentage that is only slightly lower than the 13.8% found by Netzer et al. This is of clinical significance because there are large numbers of drivers in Singapore, with close to one million motor vehicles registered in a country with a population of five million as of July 2016. We would also like to highlight that we did not find any association between a high risk score in category 1 (snoring) or 3 (hypertension or obesity) and a report of drowsy driving. Hence, we suggest that a complete assessment of OSA risk must also include questions on snoring, hypertension, and obesity in addition to a presence of daytime sleepiness. Overall, our data suggest that there is a high prevalence of symptoms suggestive of OSA even in a cohort that is generally considered to be asymptomatic. This strengthens the case for the use of screening questionnaires to optimize diagnosis of OSA in the general population.

One of the main limitations of our study design was that we oversampled snorers. Hence, slightly more than half of our study sample (56.6%) consisted of snorers. This could have led to spectrum bias and artificially inflated the sensitivities that we reported. High NPVs of a test are usually driven by high sensitivity of the test itself and a low prevalence of a disease. However, even with a moderately high prevalence of severe

### Table 3—Predictive parameters for Berlin Questionnaire to diagnose obstructive sleep apnea using AASM 2012 scoring criteria.

<table>
<thead>
<tr>
<th>AHI ≥ 15</th>
<th>AHI ≥ 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence of OSA, %</td>
<td>28.1</td>
</tr>
<tr>
<td>Sensitivity, %</td>
<td>58.8 (46.2–70.6)</td>
</tr>
<tr>
<td>Specificity, %</td>
<td>77.6 (70.7–83.5)</td>
</tr>
<tr>
<td>PPV, %</td>
<td>50.6 (39.1–62.1)</td>
</tr>
<tr>
<td>NPV, %</td>
<td>82.9 (76.1–88.3)</td>
</tr>
<tr>
<td>LR+</td>
<td>2.62</td>
</tr>
<tr>
<td>LR-</td>
<td>0.53</td>
</tr>
<tr>
<td>Area under ROC curve</td>
<td>0.682 (0.615–0.749)</td>
</tr>
</tbody>
</table>

Data are presented as average (95% confidence interval). AASM = American Academy of Sleep Medicine, AHI = apnea-hypopnea index, LR+ = positive likelihood ratio, LR- = negative likelihood ratio, NPV = negative predictive value, OSA = obstructive sleep apnea, PPV = positive predictive value, ROC = receiver operating characteristic.
OSA (10.7%) in our study sample, we found that the NPV of a low-risk Berlin questionnaire classification remained very high at 96.3%. We believe that when used in other general population settings where the sensitivity of the test may be reduced due to a lower proportion of snorers, the NPV would still remain high, especially if the prevalence of severe OSA is lower than the 10.7% reported in our sample. Another limitation is that we used portable home-based sleep studies instead of in-laboratory polysomnogram, which is the gold standard to diagnose OSA. However, there are limited slots for polysomnogram in Singapore because of a long waiting list. In addition, many participants are unwilling to travel to the hospital for a polysomnogram for a research study due to the inconvenience. Type 3 sleep studies tend to underestimate AHIs because they do not have electroencephalogram leads to stage sleep times accurately. However, a recent population-based study that re-scored complete home-based polysomnogram with only parameters available on type 3 portable monitor devices found good concordance. In fact, rescoring with only portable monitor parameters using a 3% desaturation criteria, which was similar to what we used for our primary analysis, had the best diagnostic accuracy.26

In conclusion, we found that the Berlin questionnaire may be used as a screening tool in the general population in view of its good sensitivity and high negative predictive value in ruling out severe OSA.

ABBREVIATIONS

AASM, American Academy of Sleep Medicine
AHI, apnea-hypopnea index
BMI, body mass index
NPV, negative predictive value
OSA, obstructive sleep apnea
PPV, positive predictive value
PSQI, Pittsburg Sleep Quality Index
RDI, respiratory disturbance index
ROC, receiver operating characteristic
SH2012, Singapore Health Study 2012

REFERENCES

ACKNOWLEDGMENTS
The authors thank the staff of Easmed Private Limited for assistance with the home-based sleep studies; Venesa Loh and Tan Ching Yee for help with patient recruitment; Glenn Rolden (RGPST) and Hafiz Firdaus Bin Abdul Haddy for the analysis of the sleep tracings.

SUBMISSION & CORRESPONDENCE INFORMATION
Submitted for publication July 18, 2016
Submitted in final revised form October 4, 2016
Accepted for publication October 18, 2016
Address correspondence to: Adeline Tan, MBBS, Department of Respiratory Medicine, Ng Teng Fong General Hospital, Jurong Health Services, 1 Jurong East Street 21, Singapore 609606; Tel: +65 96268485; Fax: +65 67165500; Email: ayytan@gmail.com

DISCLOSURE STATEMENT
The work for this study was performed at Department of Respiratory Medicine, Ng Teng Fong General Hospital, Jurong Health Services, 1 Jurong East Street 21, Singapore 609606. This study was funded by the FY2014 Health Services Research and Quality Improvement Grant of Ng Teng Fong General Hospital, Jurong Health Services. The authors have indicated no financial conflicts of interest.