



# Evaluation of a Modified STOP-BANG Questionnaire for Sleep Apnea in Adults from the Korean General Population

Chol Shin, MD, PhD<sup>1</sup>, Inkyung Baik, PhD<sup>2</sup>

<sup>1</sup>Division of Pulmonary, Sleep and Critical Care Medicine, Department of Internal Medicine, Korea University Ansan Hospital, Korea University College of Medicine, Ansan, Korea

<sup>2</sup>Department of Foods and Nutrition, College of Science and Technology, Kookmin University, Seoul, Korea

**Background and Objective** There are few studies assessing diagnostic accuracy of the STOP-BANG questionnaire, which is a screening tool for detecting obstructive sleep apnea (OSA), in the general population. Our study aimed to develop a modified version of this screening questionnaire and compare its diagnostic accuracy to the standard version in general Korean adults.

**Methods** A cross-sectional polysomnographic study was conducted among 2065 male and female participants aged 48–79 years, who were members of a population-based cohort study. Half of them were assigned to the exploratory sample and the other half to the validation sample.

**Results** In this study, 958 individuals (487 in an exploratory sample) were identified to have OSA. To develop a model in an exploratory sample, stepwise analysis was used with information from the standard version and additional variables; snoring, witnessed breathing cessation, hypertension, body mass index > 25 kg/m<sup>2</sup>, age groups, male, diabetes, and waist circumference > 85 cm were included in a final modified model. In the evaluation of diagnostic accuracy in a validation sample, the modified model showed higher sensitivity [79.1%, 95% confidence interval (CI): 77.3, 80.9] in detecting severe OSA than that of the standard version model (66.0%, 95% CI: 64.0, 68.0). Overall accuracy indicated as area under the receiver-operating characteristic curve for the modified model was greater ( $p = 0.001$ ) than that of the standard version model.

**Conclusions** These findings suggest that the modified version of the STOP-BANG questionnaire may be useful in detecting severe OSA in general adults. **Sleep Med Res 2021;12(1):28-35**

**Key Words** Obstructive sleep apnea, Mass screening, Polysomnography.

Received: December 11, 2020

Revised: March 16, 2021

Accepted: April 15, 2021

## Correspondence

Chol Shin, MD, PhD  
Division of Pulmonary,  
Sleep and Critical Care Medicine,  
Department of Internal Medicine,  
Korea University Ansan Hospital,  
123 Jeokgeum-ro, Danwon-gu,  
Ansan 15355, Korea

Tel +82-31-412-5603

Fax +82-31-412-5604

E-mail chol-shin@korea.ac.kr

Inkyung Baik, PhD  
Department of Foods and Nutrition,  
College of Science and Technology,  
Kookmin University, 77 Jeongneung-ro,  
Seongbuk-gu, Seoul 02707, Korea

Tel +82-2-910-4774

Fax +82-2-910-5249

E-mail ibaik@kookmin.ac.kr

## ORCID

Chol Shin  
<http://orcid.org/0000-0002-2928-8576>

Inkyung Baik  
<http://orcid.org/0000-0002-9524-4344>

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## INTRODUCTION

The prevalence of obstructive sleep apnea (OSA) has gradually increased, partly due to population aging and a rising prevalence of risk factors such as obesity, and thus it is now considered to be a prevalent disease [1,2]. Moreover, OSA is recognized as an emerging public health problem because it has been reported to be associated with cardiovascular and metabolic diseases [3,4] and a large number of OSA cases remain undiagnosed and untreated in the community [5]. It is recommended that a diagnosis of OSA should be established by laboratory-based or home-based polysomnography (PSG) [6]. PSG provides objective and quantitative data on sleep status and the severity of OSA, but it can be costly and time-consuming, as well as uncomfortable, and disruptive to sleep for patients undergoing the test. In addition, because PSG-related resources are unavailable in most primary care settings, patients at risk of OSA are less likely to be identified and have limited access to proper care [7]. Use of screening questionnaires such as the STOP-BANG and Berlin questionnaires has been suggested as an alternative approach for identifying potential cases of OSA in the primary care setting [8]. The STOP-BANG questionnaire in particular seems to be a useful screening tool for OSA because it is simple and has

been validated [9]. A number of studies evaluating the diagnostic accuracy of the STOP-BANG questionnaire among sleep clinic patients showed the sensitivity greater than 80% for detecting all cases of OSA and greater than 95% for detecting severe cases, although the specificity was found to be less than 50% [10-17]. On the other hand, there are few studies testing the STOP-BANG questionnaire in the general population [18,19]. These studies have observed lower sensitivity and higher specificity (between 60% and 70%) for severe OSA compared to studies performed in sleep clinic patients [18,19]. To improve such a modest sensitivity, Tan et al. [19] modified the questionnaire, using a lower cutoff point for body mass index (BMI). However, they found similar sensitivity for the standard and modified version models. Thus, further studies are warranted to develop a modified version model that can improve the predictive ability of the STOP-BANG questionnaire, making it more applicable and useful for detecting OSA in the general population.

The objectives of this study were to develop predictive models using potential risk factors of OSA in addition to the components of the standard STOP-BANG questionnaire, construct a scoring method for the modified version models, and compare diagnostic accuracy parameters in the standard and modified versions for detecting all and severe cases of OSA in the general adult population.

## METHODS

### Study Design and Population

We conducted a cross-sectional PSG study in a population-based cohort study conducted in Ansan, Republic of Korea, a part of the Korean Genome and Epidemiology Study. This prospective cohort study is ongoing and has been carried out since 2001. Detailed information of this cohort study is available elsewhere [20].

Between August 17, 2009 and February 28, 2013, 2065 participants aged 48–79 years from Ansan underwent PSG and were included in the present study. They were assigned either to the exploratory sample or the validation sample by a gender-stratified random sampling procedure based on a half-split-sample validation method. The exploratory sample was used to construct a predictive model and scoring method in order to develop a modified version of the questionnaire. The validation sample was used for evaluating diagnostic accuracy of the standard and modified versions.

The Human Subjects Review Committee at the Korea University Ansan Hospital approved the study protocol as well as the informed consent form, which each participant signed (IRB number: ED0624).

### Polysomnography and Outcome

Unattended overnight PSG was carried out either at the par-

ticipant's home or at the Korea University Ansan Hospital sleep laboratory using a portable monitoring device (Embletta X-100; Embla Systems, Broomfield, CO, USA). A detailed description of PSG method is available elsewhere [21]. To calculate the apnea-hypopnea index (AHI) (average number of apnea and hypopnea events per sleep hour), apnea was defined as a > 90% reduction in airflow from baseline for at least 10 s and hypopnea was defined as a  $\geq 30\%$  reduction in airflow from baseline accompanied by a  $\geq 4\%$  decrease in oxygen saturation. In this study, OSA was defined as having AHI  $\geq 5$  in order to develop a predictive model. Subgroups such as mild to moderate cases ( $5 < \text{AHI} < 30$ ) and severe cases (AHI  $\geq 30$ ) were also used to calculate diagnostic accuracy.

### Potential Risk Factors

Information about potential risk factors such as age, sex, smoking status, alcohol consumption status, and sleep-related symptoms was collected from the interview-based questionnaire. Anthropometric and clinical data including BMI, neck circumference (NC), waist circumference (WC), and blood pressure were collected through a comprehensive health examination conducted by health professionals according to a standardized protocol. Biomarkers such as high-sensitivity C-reactive protein (hs-CRP), insulin, leptin, and ghrelin concentrations were selected based on previous reports [22,23] and assayed in serum samples collected following minimum 8-hour fast. The presence of hypertension and diabetes mellitus (DM) was determined based on blood pressure, fasting and post-load blood glucose levels, and medication history.

### Statistical Analysis

Descriptive statistics of characteristics were calculated for the exploratory sample, the validation sample, and all participants taken together. To compare characteristics between the exploratory sample and the validation sample, the chi-squared test for categorical variables and the Student's t-test for continuous variables were used.

To analyze associations between OSA and potential risk factors, we conducted multivariate logistic regression analysis and obtained odds ratios (ORs) with 95% confidence interval (CI). First, we conducted multivariate association analysis for the components of the standard STOP-BANG questionnaire including daytime sleepiness, snoring, witnessed breathing cessation, presence of hypertension, BMI  $> 35 \text{ kg/m}^2$ , age  $> 50$  years, NC  $> 40$  cm, and male sex in all participants as well as in the exploratory sample separately. Next, we constructed three multivariate models. For the first model, we considered all components of the standard STOP-BANG questionnaire, but modified the BMI cutoff point. To determine this cutoff point, we conducted univariate logistic regression analysis for the association between BMI and OSA and selected a BMI having a minimum value of the square root of  $[(1 - \text{sensitivity})^2 + (1 - \text{specificity})^2]$ , which in-

indicates the minimum distance from the upper left corner to the point on the receiver-operating characteristic (ROC) curve that is the optimal cutoff point [24]. For the second model, we considered WC, DM, and smoking and alcohol consumption status as well as all variables in the first model. In particular, the WC variable was fitted as a binary variable that was created using the optimal cutoff point, which was determined by univariate logistic regression analysis for the association between WC and OSA. For the third model, we considered biomarkers including hs-CRP, insulin, leptin, and ghrelin in addition to all variables in the second model and then selected significant variables using the forward method in multivariate regression analysis. A final model consisting of all significant variables was established for scoring.

We made a scoring system for each model. In the standard scoring method for the STOP-BANG questionnaire, one or zero is assigned for each risk factor according to its presence or absence and then all point values are summed. In this study, we mainly used this method but further considered a modified scoring method based on the effect size estimates of risk factors. After scoring each model, we evaluated the diagnostic accuracy of the standard version and the modified versions using  $\geq 3$  as the cutoff score and calculated measures of diagnostic accuracy such as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the ROC curve (AUC). Furthermore, we tested the difference in the AUC between the standard version model and each of the modified models.

All testing was based on a 2-sided level of significance ( $p < 0.05$ ) and was conducted using the SAS software (SAS 9.1.3; SAS Institute, Cary, NC, USA).

## RESULTS

Table 1 shows characteristics for all participants ( $n = 2065$ ) in both the exploratory sample ( $n = 1032$ ) and the validation sample ( $n = 1033$ ). Among all participants, 0.24% had a BMI  $> 35$  kg/m<sup>2</sup>, 4.4% had a NC  $> 40$  cm, 32% had daytime sleepiness, 79% snored, 21% had witnessed breathing cessation, and 46% had a diagnosis of OSA based on PSG. In our sample population, a total of 958 participants (487 in an exploratory sample) were determined to have OSA, which is double the number of participants who reported witnessed breathing cessation. The distribution of measured characteristics was similar between the two samples.

The results regarding the associations between the components of the STOP-BANG questionnaire and OSA among all participants as well as in the exploratory sample are presented in Table 2. Daytime sleepiness and BMI  $> 35$  kg/m<sup>2</sup> were not associated with OSA in either group. NC  $> 40$  cm was significantly associated with OSA in the exploratory sample, but not when all participants were evaluated together.

Table 3 shows multivariate ORs (95% CI) of OSA for the components of the STOP-BANG questionnaire and other potential risk factors in the exploratory sample. In the modified model 1, the BMI cutoff point was modified and all other variables were the same as the STOP-BANG questionnaire. The optimal BMI cutoff point, which was obtained from the univariate regression analysis for the association between BMI and OSA, was found to be 25 kg/m<sup>2</sup>. Fig. 1 presents a ROC curve of BMI with points for 25 kg/m<sup>2</sup> and 35 kg/m<sup>2</sup>. In the modified model 1, BMI  $> 25$  kg/m<sup>2</sup> was significantly associated with OSA, but daytime sleepiness and NC  $> 40$  cm were not associated with OSA. For the modified model 2, WC, DM, smoking status, and alcohol consumption status were also included. The optimal WC cutoff point was determined as 85 cm based on the association between WC and OSA. In this model, daytime sleepiness, NC  $> 40$  cm, and smoking and alcohol consumption status were not associated with OSA. Using the modified model 2, the forward procedure identified the following significant variables which were included in the modified model 3: snoring, witnessed breathing cessation, hypertension, BMI  $> 25$  kg/m<sup>2</sup>, age  $> 50$  years, male sex, WC  $> 85$  cm, and DM. All of these variables were found to be significantly associated with OSA.

Next, we scored the modified models using the standard scoring method: one or zero was assigned for each risk factor according to its presence or absence because effect size estimates for all risk factors except age  $\leq 50$  years were observed to be similar as presented in Table 3. When we further analyzed using the three age groups ( $\leq 50$  years, 51–64 years and  $\geq 65$  years), in the modified model 3, multivariate ORs (95% CI) were observed to be 2.70 (1.63, 4.48) and 4.50 (2.54, 7.98) for the 51–64 age group and  $\geq 65$  age group, respectively, compared to the  $\leq 50$  age group. The scoring method was modified to accommodate the differences found between age groups as follows: zero points for  $\leq 50$  years, one point for 51–64 years, and two points for  $\geq 65$  years. One or zero points were assigned for all other risk factors. Total scores were calculated by summing the point values assigned to risk factors for each model. Both the standard and the modified scoring methods were applied to the modified model 3. The distribution of total scores for the standard and the modified models is displayed in Fig. 2. The proportion of individuals with total score  $\geq 3$  was observed to be greater in the modified models than in the standard version model.

Table 4 demonstrates the diagnostic accuracy parameters, which were calculated for each model in the validation sample. For detecting mild to moderate OSA, sensitivity and specificity estimates were found to be between 56% and 64%, PPV and NPV estimates between 57% and 73%, and AUC estimates between 0.66 and 0.72. In terms of diagnostic accuracy for detecting severe OSA, sensitivity was found to be between 66% and 79%, specificity between 51% and 55%, PPV between 5% and 6%, NPV between 98% and 99%, and AUC between 0.67 and 0.78. On the basis of these results, the modified model 3 showed the

**Table 1.** Characteristics of 2065 study participants and comparison between two samples

| Characteristics                       | All<br>(n = 2065) | Exploratory sample<br>(n = 1032) | Validation sample<br>(n = 1033) | p-value* |
|---------------------------------------|-------------------|----------------------------------|---------------------------------|----------|
| Male                                  | 52.2              | 52.1                             | 52.2                            | 0.983    |
| Age, years                            | 59.1 ± 7.6        | 59.0 ± 7.3                       | 59.2 ± 7.9                      | 0.499    |
| BMI, kg/m <sup>2</sup>                | 24.7 ± 2.9        | 24.6 ± 2.9                       | 24.8 ± 3.0                      | 0.098    |
| BMI > 35 kg/m <sup>2</sup>            | 0.24              | 0.19                             | 0.29                            | 0.999    |
| NC, cm                                | 34.8 ± 3.2        | 34.8 ± 3.3                       | 34.9 ± 3.1                      | 0.559    |
| NC > 40 cm                            | 4.41              | 4.55                             | 4.26                            | 0.744    |
| WC, cm                                | 81.7 ± 8.3        | 81.5 ± 8.4                       | 81.9 ± 8.2                      | 0.294    |
| Current smokers                       | 12.5              | 11.6                             | 13.5                            | 0.210    |
| Current alcohol drinkers              | 46.5              | 47.2                             | 45.9                            | 0.553    |
| Presence of hypertension              | 38.1              | 37.9                             | 38.3                            | 0.834    |
| Presence of DM                        | 21.9              | 23.0                             | 20.9                            | 0.259    |
| Self-reported sleep-related variables |                   |                                  |                                 |          |
| Daytime sleepiness                    | 31.8              | 32.9                             | 30.7                            | 0.292    |
| Snoring                               | 78.5              | 76.7                             | 80.3                            | 0.052    |
| Witnessed breathing cessation         | 20.8              | 19.6                             | 22.1                            | 0.162    |
| Polysomnographic recordings           |                   |                                  |                                 |          |
| AHI, events/hour                      | 7.36 ± 8.76       | 7.46 ± 8.65                      | 7.27 ± 8.87                     | 0.609    |
| Severity of OSA                       |                   |                                  |                                 | 0.627    |
| Mild (5 ≤ AHI < 15)                   | 32.7              | 33.0                             | 32.4                            |          |
| Moderate (15 ≤ AHI < 30)              | 10.8              | 11.5                             | 10.1                            |          |
| Severe (AHI ≥ 30)                     | 2.86              | 2.62                             | 3.10                            |          |
| Serum biomarkers (mg/dL)              |                   |                                  |                                 |          |
| Hs-CRP                                | 1.43 ± 3.28       | 1.41 ± 3.32                      | 1.46 ± 3.24                     | 0.622    |
| Insulin                               | 8.73 ± 4.46       | 8.74 ± 4.54                      | 8.73 ± 4.38                     | 0.952    |
| Leptin                                | 6.74 ± 5.60       | 6.73 ± 5.75                      | 6.75 ± 5.44                     | 0.933    |
| Ghrelin                               | 777.8 ± 311.1     | 771.7 ± 301.6                    | 783.9 ± 320.4                   | 0.393    |

Values are expressed as percentage or mean ± standard deviation unless otherwise indicated.

\*Analyses of the variables were performed by t-test.

BMI: body mass index, NC: neck circumference, WC: waist circumference, DM: diabetes mellitus, AHI: apnea-hypopnea index, OSA: obstructive sleep apnea, hs-CRP: high-sensitivity C-reactive protein.

**Table 2.** Association between components of the STOP-BANG questionnaire and obstructive sleep apnea based on polysomnography

| Components                    | OR (95% CI) in all participants |                                | OR (95% CI) in the exploratory sample |                                |
|-------------------------------|---------------------------------|--------------------------------|---------------------------------------|--------------------------------|
|                               | Unadjusted                      | Multivariate model             | Unadjusted                            | Multivariate model             |
| Daytime sleepiness            | 0.91 (0.76, 1.10)               | 0.86 (0.70, 1.05)              | 1.02 (0.79, 1.32)                     | 1.01 (0.76, 1.34)              |
| Snoring                       | 2.24 (1.79, 2.80) <sup>‡</sup>  | 1.82 (1.43, 2.31) <sup>‡</sup> | 2.51 (1.84, 3.41) <sup>‡</sup>        | 1.95 (1.40, 2.73) <sup>‡</sup> |
| Witnessed breathing cessation | 3.20 (2.55, 4.02) <sup>‡</sup>  | 2.69 (2.09, 3.47) <sup>‡</sup> | 3.06 (2.20, 4.24) <sup>‡</sup>        | 2.30 (1.60, 3.32) <sup>‡</sup> |
| Presence of hypertension      | 2.30 (1.92, 2.76) <sup>‡</sup>  | 2.05 (1.69, 2.48) <sup>‡</sup> | 2.41 (1.86, 3.11) <sup>‡</sup>        | 2.06 (1.57, 2.71) <sup>‡</sup> |
| BMI > 35 kg/m <sup>2</sup>    | 1.73 (0.29, 10.37)              | 2.20 (0.33, 14.88)             | 1.12 (0.07, 17.95)                    | 1.71 (0.09, 32.72)             |
| Age > 50 years                | 2.98 (2.20, 4.02) <sup>‡</sup>  | 3.40 (2.45, 4.72) <sup>‡</sup> | 3.15 (2.00, 4.95) <sup>‡</sup>        | 2.98 (1.84, 4.85) <sup>‡</sup> |
| NC > 40 cm                    | 2.86 (1.81, 4.53) <sup>‡</sup>  | 1.41 (0.87, 2.31)              | 4.40 (2.16, 8.95) <sup>‡</sup>        | 2.22 (1.05, 4.66) <sup>*</sup> |
| Male                          | 1.87 (1.57, 2.23) <sup>‡</sup>  | 1.54 (1.27, 1.87) <sup>‡</sup> | 1.94 (1.51, 2.49) <sup>‡</sup>        | 1.57 (1.19, 2.07) <sup>‡</sup> |

Data in the multivariate model are adjusted for risk factors presented in the table.

\*p < 0.05, <sup>‡</sup>p < 0.01, <sup>‡‡</sup>p < 0.001 indicate statistical significance.

OR: odds ratio, CI: confidence interval, BMI: body mass index, NC: neck circumference.

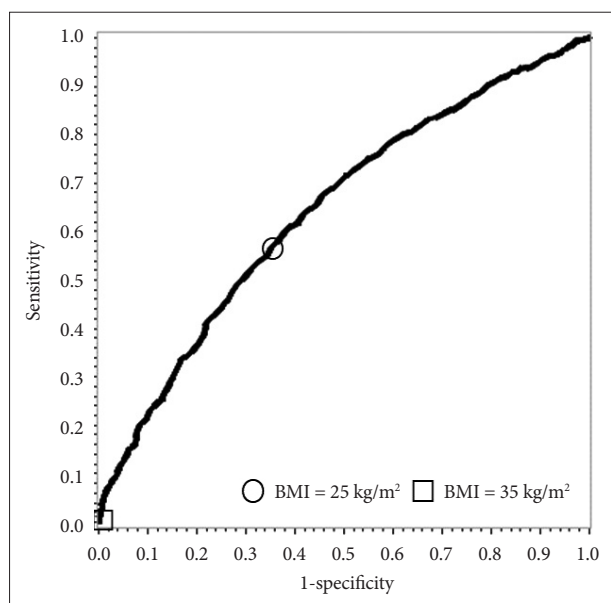
**Table 3.** Association between potential risk factors and obstructive sleep apnea based on polysomnography in the exploratory sample

| Risk factors                  | Variables of the standard version model | Modified model 1               | Modified model 2               | Modified model 3               |
|-------------------------------|---|--------------------------------|--------------------------------|--------------------------------|
|                               |   | Multivariate OR (95% CI)       | Multivariate OR (95% CI)       | Multivariate OR (95% CI)       |
| Daytime sleepiness            | Yes                                     | 0.96 (0.72, 1.29)              | 0.97 (0.72, 1.29)              |                                |
| Snoring                       | Yes                                     | 1.76 (1.25, 2.48) <sup>†</sup> | 1.76 (1.25, 2.48) <sup>†</sup> | 1.76 (1.25, 2.48) <sup>†</sup> |
| Witnessed breathing cessation | Yes                                     | 2.30 (1.58, 3.33) <sup>‡</sup> | 2.23 (1.53, 3.26) <sup>‡</sup> | 2.23 (1.53, 3.24) <sup>‡</sup> |
| Presence of hypertension      | Yes                                     | 1.85 (1.40, 2.44) <sup>‡</sup> | 1.65 (1.24, 2.19) <sup>‡</sup> | 1.69 (1.27, 2.24) <sup>‡</sup> |
| BMI > 25 kg/m <sup>2</sup>    | No                                      | 2.55 (1.92, 3.38) <sup>‡</sup> | 2.03 (1.46, 2.82) <sup>‡</sup> | 2.06 (1.49, 2.86) <sup>‡</sup> |
| Age > 50 years                | Yes                                     | 3.13 (1.90, 5.15) <sup>‡</sup> | 3.00 (1.81, 4.98) <sup>‡</sup> | 2.99 (1.81, 4.94) <sup>‡</sup> |
| NC > 40 cm                    | Yes                                     | 1.27 (0.59, 2.72)              | 1.07 (0.49, 2.32)              |                                |
| Male                          | Yes                                     | 1.67 (1.26, 2.21) <sup>‡</sup> | 1.45 (1.05, 2.00) <sup>*</sup> | 1.52 (1.14, 2.03) <sup>†</sup> |
| WC ≥ 85 cm                    | No                                      |                                | 1.48 (1.04, 2.11) <sup>*</sup> | 1.48 (1.04, 2.10) <sup>*</sup> |
| Presence of DM                | No                                      |                                | 1.73 (1.24, 2.41) <sup>†</sup> | 1.71 (1.23, 2.39) <sup>†</sup> |
| Smoker                        | No                                      |                                | 0.83 (0.53, 1.29)              |                                |
| Alcohol drinker               | No                                      |                                | 1.23 (0.91, 1.66)              |                                |

Data in the models are adjusted for risk factors presented in the table.

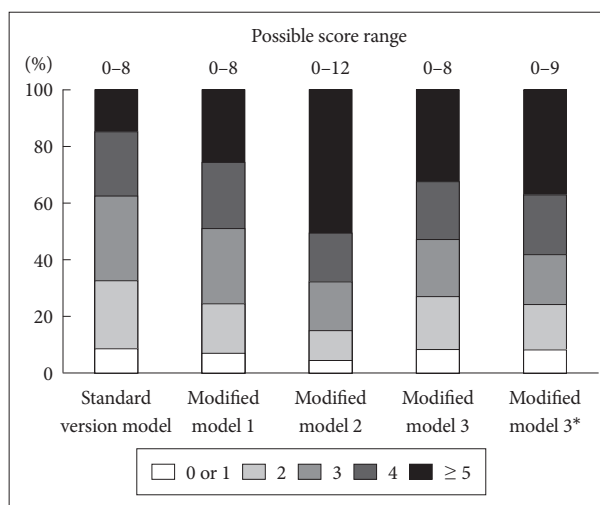
\*p < 0.05, <sup>†</sup>p < 0.01, <sup>‡</sup>p < 0.001 indicate statistical significance.

OR: odds ratio, CI: confidence interval, BMI: body mass index, NC: neck circumference, WC: waist circumference, DM: diabetes mellitus.



**Fig. 1.** The receiver-operating characteristic curve of BMI in the association with obstructive sleep apnea. Points for 25 kg/m<sup>2</sup> and 35 kg/m<sup>2</sup> are depicted on the curve. BMI: body mass index.

greater estimate of AUC compared to the other models for detecting all cases and severe cases of OSA. Regardless of scoring method, the AUC of the modified model 3 was significantly different from that of the standard version model for detecting mild to moderate cases (p < 0.001) and severe cases (p < 0.01) of OSA. Because of the high prevalence of mild to moderate OSA, the measured diagnostic accuracy for detecting all cases of OSA was found to be similar to the accuracy for detecting mild to moderate OSA. Using the five models presented in the Table 4, ROC curves for all OSA cases are presented in Fig. 3.



**Fig. 2.** Distribution of the scores of the standard STOP-BANG questionnaire and the modified models in the exploratory sample. The Y axis indicates the accumulative percentage of score points according to the standard version model and the modified models. \*In the modified scoring method, zero to two points (zero for ≤ 50 years, one for 51–64 years, and two for ≥ 65 years) are assigned for age groups and one or zero points for all other risk factors.

## DISCUSSION

In this cross-sectional study, we carried out PSG in the general adult population of 2065 male and female participants and attempted to develop and validate a predictive model to improve diagnostic accuracy for detecting OSA compared to the standard STOP-BANG questionnaire. Our results show that a predictive model which includes risk factors such as daytime sleepiness, snoring, witnessed breathing cessation, hypertension, BMI > 25 kg/m<sup>2</sup>, middle-aged and older groups (51–64 years and ≥



**Table 4.** Results of diagnostic accuracy for the standard STOP-BANG questionnaire model and modified models using total score  $\geq 3$  as cutoff in the validation sample

| Predictive parameters       | Standard version model<br>+ standard scoring<br>method | Modified model 1<br>+ standard scoring<br>method | Modified model 2<br>+ standard scoring<br>method | Modified model 3<br>+ standard scoring<br>method | Modified model 3<br>+ modified scoring<br>method |
|-----------------------------|--|--|--|--|--|
| <b>All OSA</b>              |  |  |  |  |  |
| Sensitivity, %              | 56.4 (54.5, 58.3)                                      | 60.1 (58.3, 61.9)                                | 60.1 (58.3, 61.9)                                | 59.7 (57.9, 61.5)                                | 62.3 (60.5, 64.2)                                |
| Specificity, %              | 63.0 (61.4, 64.6)                                      | 60.3 (58.7, 61.9)                                | 59.7 (58.0, 61.3)                                | 64.7 (63.1, 66.2)                                | 64.5 (62.9, 66.0)                                |
| PPV, %                      | 59.4 (58.9, 60.0)                                      | 59.4 (58.9, 59.9)                                | 58.9 (58.4, 59.4)                                | 62.9 (62.3, 63.5)                                | 64.0 (63.4, 64.6)                                |
| NPV, %                      | 65.8 (65.4, 66.2)                                      | 67.7 (67.2, 68.1)                                | 68.2 (67.6, 68.7)                                | 69.2 (68.6, 69.8)                                | 71.8 (71.1, 72.5)                                |
| AUC                         | 0.660 (0.628, 0.692)                                   | 0.674 (0.642, 0.705)                             | 0.678 (0.646, 0.710)                             | 0.708 (0.677, 0.738)                             | 0.731 (0.701, 0.761)                             |
| Test statistics* (p values) |  | -2.07 (0.039)                                    | -1.72 (0.086)                                    | -3.97 (< 0.001)                                  | -5.55 (< 0.001)                                  |
| <b>Mild to moderate OSA</b> |  |  |  |  |  |
| Sensitivity, %              | 56.1 (54.2, 58.0)                                      | 59.7 (57.8, 61.5)                                | 59.8 (57.9, 61.6)                                | 59.2 (57.3, 61.0)                                | 62.0 (60.1, 63.9)                                |
| Specificity, %              | 62.8 (61.2, 64.4)                                      | 59.8 (58.1, 61.4)                                | 59.2 (57.5, 60.9)                                | 64.1 (62.5, 65.7)                                | 63.8 (62.2, 65.4)                                |
| PPV, %                      | 57.2 (56.7, 57.8)                                      | 57.0 (56.6, 57.5)                                | 56.7 (56.2, 57.1)                                | 60.4 (59.8, 61.0)                                | 61.6 (61.0, 62.3)                                |
| NPV, %                      | 67.0 (66.5, 67.4)                                      | 68.6 (68.1, 69.0)                                | 69.1 (68.6, 69.6)                                | 70.0 (69.5, 70.5)                                | 72.6 (71.9, 73.3)                                |
| AUC                         | 0.655 (0.622, 0.687)                                   | 0.666 (0.633, 0.698)                             | 0.671 (0.638, 0.704)                             | 0.699 (0.667, 0.730)                             | 0.722 (0.691, 0.752)                             |
| Test statistics* (p values) |  | -1.64 (0.101)                                    | -1.51 (0.131)                                    | -3.56 (< 0.001)                                  | -5.11 (< 0.001)                                  |
| <b>Severe OSA</b>           |  |  |  |  |  |
| Sensitivity, %              | 66.0 (64.0, 68.0)                                      | 73.6 (71.8, 75.5)                                | 71.6 (69.7, 73.5)                                | 76.5 (74.8, 78.3)                                | 79.1 (77.3, 80.9)                                |
| Specificity, %              | 54.8 (53.1, 56.5)                                      | 51.8 (50.1, 53.5)                                | 51.4 (49.7, 53.1)                                | 54.5 (52.9, 56.2)                                | 53.3 (51.6, 54.9)                                |
| PPV, %                      | 5.17 (4.91, 5.43)                                      | 5.45 (5.26, 5.63)                                | 5.18 (5.04, 5.32)                                | 6.20 (6.02, 6.38)                                | 6.03 (6.89, 6.17)                                |
| NPV, %                      | 98.4 (98.3, 98.4)                                      | 98.8 (98.7, 98.8)                                | 98.8 (98.7, 98.8)                                | 99.0 (99.0, 99.1)                                | 99.2 (99.1, 99.2)                                |
| AUC                         | 0.668 (0.583, 0.753)                                   | 0.716 (0.638, 0.793)                             | 0.705 (0.630, 0.779)                             | 0.760 (0.691, 0.830)                             | 0.779 (0.718, 0.839)                             |
| Test statistics* (p values) |  | -2.86 (0.004)                                    | -1.30 (0.195)                                    | -2.81 (0.005)                                    | -3.25 (0.001)                                    |

Data are presented as values (95% confidence interval) unless otherwise indicated.

\*Difference in two areas under the ROC curve between the standard version model (+ standard scoring method) and each of the modified models.

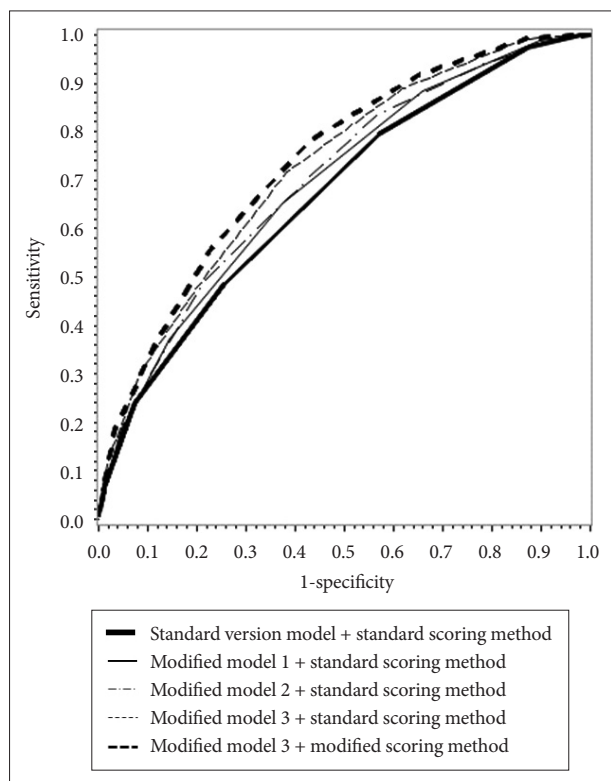
PPV: positive predictive value, NPV: negative predictive value, ROC: receiver-operating characteristic, OSA: obstructive sleep apnea, AUC: area under the ROC curve.

65 years), male sex, WC > 85 cm, and DM had significantly greater AUC with higher sensitivity, specificity, PPV, and NPV when compared to the diagnostic accuracy parameters of the standard version model. In particular, because this predictive model was found to have a higher sensitivity (79%) than the standard version model (66%) for detecting severe OSA, it may be a useful tool for screening for high-risk cases of OSA in the general adult population.

A meta-analysis study reported that the prevalence of OSA ranges from 9% to 38% in the general population [1]. Such a large variation in the estimates may be due to diverse characteristics of the study population, such as age range, gender distribution, and the prevalence of obesity, as well as discrepancies in sampling methods. Nonetheless, OSA prevalence is likely increasing because of aging population trends and increases in obesity prevalence.

Data regarding the causal associations of OSA with cardiovascular and metabolic diseases, occupational accidents, and other

adverse consequences, such as excessive daytime sleepiness or decreased work performance, are accumulating [3,4,25,26]. Improved screening questionnaires to identify potential cases of OSA are needed to help prevent these adverse outcomes, as most cases remain undiagnosed and patients may be unaware of difficulty breathing and repeated awakenings during sleep [7]. Several studies have highlighted the use of the STOP-BANG questionnaire, which is short and includes simple questions, in patients who visited sleep clinics [10-17] as well as in the general population [18,19]. This questionnaire has been found to show comparable diagnostic accuracy, particularly with regards to sensitivity, for detecting OSA in a sleep clinic population when compared to other longer questionnaires [10-17]. However, there are limited studies evaluating this questionnaire in the general population. Two population-based studies observed nearly identical sensitivity for detecting severe OSA: 70% sensitivity was reported in a study of 4770 American adults [18] and 69% sensitivity in a study of 242 Singaporean adults [19]. However, the sensi-



**Fig. 3.** The receiver-operating characteristic curves of the standard and modified model scores for obstructive sleep apnea.

tivity and specificity results from these two studies for detecting moderate to severe cases were inconsistent [18,19].

We also observed similar sensitivity for detecting severe OSA in the standard version model compared with previous population-based studies [18,19]. When using the modified version (modified model 3 + modified scoring method), however, we observed higher sensitivity than what has been found in previous studies. In particular, we found that the modified version can detect all severe cases of OSA. Explaining in detail, we observed that only 21% of the study population reported witnessed breathing cessation, and thus 46% of severe cases of OSA (or 69% of all cases) were likely to be unaware of apnea. Among severe cases of OSA who reported no witnessed breathing cessation, 85% were detected with the STOP-BANG questionnaire. Meanwhile, all severe cases were detected with the final modified version, regardless of witnessed breathing cessation. The final modified version includes variables such as WC and DM in addition to the components of the standard version, and does not include daytime sleepiness and NC. Both WC and NC have been reported to be associated with OSA, although which anthropometric measure is a better predictor for OSA risk remains controversial [27,28]. In our study, WC was associated with OSA at a high significance level in the multivariate regression analysis, but NC was not (data available on request). Previous reports suggested that OSA is a risk factor for type 2 diabetes, but also vice versa, although more data are needed [29,30]. In our study, a weak as-

sociation between daytime sleepiness and OSA was observed and this might be partly due to snorers, because daytime sleepiness was strongly associated with snoring (data available on request). Previous studies found an association between daytime sleepiness and snoring independent of OSA [31,32]. With regards to the modification of the cutoff for BMI, we obtained an optimal cutoff point based on the ROC curve of the association between BMI and OSA. In a previous study, either 30 kg/m<sup>2</sup> or 27.5 kg/m<sup>2</sup> was used, but diagnostic accuracy was not improved by this modification in the general population [19]. A recent study including Korean patients, which used 30 kg/m<sup>2</sup> for the BMI cutoff and 36.3 cm for the NC cutoff, observed modest improvement in sensitivity and reduced specificity leading to similar AUC values compared with the standard version [33]. In our findings, modification of the BMI cutoff improved the AUC significantly (in modified model 1) for detecting severe OSA. In addition, we considered other variables such as smoking status, alcohol consumption, and some biomarkers in order to improve diagnostic accuracy, but did not observe significant associations with OSA.

The present study has some strengths: we used a large sample size from the general population, a half-split-sample validation approach, and a wide range of variables. One study limitation was that the final modified version could not improve specificity. The standard STOP-BANG questionnaire itself shows low specificity and PPV because of high false positive results [10-19]. Nonetheless, this questionnaire and the modified version may be useful to primary care physicians for educating patients about the risk factors of OSA, which overlap with those for cardiovascular disease. Another limitation of the study is limited generalizability of our findings. Further investigation is warranted to develop an advanced screening tool to improve specificity and PPV.

In conclusion, our findings suggest that the modified version of the STOP-BANG questionnaire which the modified scoring method is applied to may be useful in detecting severe OSA in general Korean adults.

#### Acknowledgments

This study was supported by a fund (2009-E71002-00, 2010-E71001-00, 2011-E71004-00, 2012-E71005-00) by research of Korea Centers for Disease Control and Prevention. This study was also supported by National Research Foundation of Korea Grant funded by the Korean Government (NRF-2019R1A2C2084000).

#### Conflicts of Interest

The authors have no financial conflicts of interest.

#### Authors' Contribution

Conceptualization: Baik I. Data curation: Shin C. Formal analysis: Baik I. Funding acquisition: Shin C, Baik I. Methodology: Baik I. Project administration: Shin C. Resources: Shin C. Software: Baik I. Supervision: Shin C. Validation: Shin C. Visualization: Baik I. Writing—original draft: Baik I. Writing—review & editing: Shin C, Baik I.

#### REFERENCES

1. Senaratna CV, Perret JL, Lodge CJ, Lowe AJ, Campbell BE, Matheson

- MC, et al. Prevalence of obstructive sleep apnea in the general population: a systematic review. *Sleep Med Rev* 2017;34:70-81.
2. Franklin KA, Lindberg E. Obstructive sleep apnea is a common disorder in the population—a review on the epidemiology of sleep apnea. *J Thorac Dis* 2015;7:1311-22.
  3. Wang X, Ouyang Y, Wang Z, Zhao G, Liu L, Bi Y. Obstructive sleep apnea and risk of cardiovascular disease and all-cause mortality: a meta-analysis of prospective cohort studies. *Int J Cardiol* 2013;169:207-14.
  4. Kong DL, Qin Z, Wang W, Pan Y, Kang J, Pang J. Association between obstructive sleep apnea and metabolic syndrome: a meta-analysis. *Clin Invest Med* 2016;39:E161-72.
  5. Simpson L, Hillman DR, Cooper MN, Ward KL, Hunter M, Cullen S, et al. High prevalence of undiagnosed obstructive sleep apnoea in the general population and methods for screening for representative controls. *Sleep Breath* 2013;17:967-73.
  6. Qaseem A, Dallas P, Owens DK, Starkey M, Holty JE, Shekelle P; Clinical Guidelines Committee of the American College of Physicians. Diagnosis of obstructive sleep apnea in adults: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2014;161:210-20.
  7. Mold JW, Quattlebaum C, Schinnerer E, Boeckman L, Orr W, Hollabaugh K. Identification by primary care clinicians of patients with obstructive sleep apnea: a practice-based research network (PBRN) study. *J Am Board Fam Med* 2011;24:138-45.
  8. Senthilvel E, Auckley D, Dasarathy J. Evaluation of sleep disorders in the primary care setting: history taking compared to questionnaires. *J Clin Sleep Med* 2011;7:41-8.
  9. Chung F, Abdullah HR, Liao P. STOP-Bang questionnaire: a practical approach to screen for obstructive sleep apnea. *Chest* 2016;149:631-8.
  10. Ong TH, Raudha S, Fook-Chong S, Lew N, Hsu AA. Simplifying STOP-BANG: use of a simple questionnaire to screen for OSA in an Asian population. *Sleep Breath* 2010;14:371-6.
  11. El-Sayed IH. Comparison of four sleep questionnaires for screening obstructive sleep apnea. *Egypt J Chest Dis Tuberc* 2012;61:433-41.
  12. Boynton G, Vahabzadeh A, Hammoud S, Ruzicka DL, Chervin RD. Validation of the STOP-BANG questionnaire among patients referred for suspected obstructive sleep apnea. *J Sleep Disord Treat Care* 2013;2:10.4172/2325-9639.1000121.
  13. Pereira EJ, Driver HS, Stewart SC, Fitzpatrick MF. Comparing a combination of validated questionnaires and level III portable monitor with polysomnography to diagnose and exclude sleep apnea. *J Clin Sleep Med* 2013;9:1259-66.
  14. Pataka A, Daskalopoulou E, Kalamaras G, Fekete Passa K, Argyropoulou P. Evaluation of five different questionnaires for assessing sleep apnea syndrome in a sleep clinic. *Sleep Med* 2014;15:776-81.
  15. Alhouqani S, Al Manhali M, Al Essa A, Al-Houqani M. Evaluation of the Arabic version of STOP-Bang questionnaire as a screening tool for obstructive sleep apnea. *Sleep Breath* 2015;19:1235-40.
  16. Sadeghniaat-Haghighi K, Montazeri A, Khajeh-Mehrzi A, Ghajarzadeh M, Alemohammad ZB, Aminian O, et al. The STOP-BANG questionnaire: reliability and validity of the Persian version in sleep clinic population. *Qual Life Res* 2015;24:2025-30.
  17. Prasad KT, Sehgal IS, Agarwal R, Nath Aggarwal A, Behera D, Dhooira S. Assessing the likelihood of obstructive sleep apnea: a comparison of nine screening questionnaires. *Sleep Breath* 2017;21:909-17.
  18. Silva GE, Vana KD, Goodwin JL, Sherrill DL, Quan SF. Identification of patients with sleep disordered breathing: comparing the four-variable screening tool, STOP, STOP-Bang, and Epworth Sleepiness Scales. *J Clin Sleep Med* 2011;7:467-72.
  19. Tan A, Yin JD, Tan LW, van Dam RM, Cheung YY, Lee CH. Predicting obstructive sleep apnea using the STOP-Bang questionnaire in the general population. *Sleep Med* 2016;27-28:66-71.
  20. Kim Y, Han BG; KoGES Group. Cohort profile: the Korean Genome and Epidemiology Study (KoGES) consortium. *Int J Epidemiol* 2017;46:e20.
  21. Baik I, Seo HS, Yoon D, Kim SH, Shin C. Associations of sleep apnea, NRG1 polymorphisms, alcohol consumption, and cerebral white matter hyperintensities: analysis with genome-wide association data. *Sleep* 2015;38:1137-43.
  22. Harsch IA, Konturek PC, Koebnick C, Kuehnlein PP, Fuchs FS, Pour Schahin S, et al. Leptin and ghrelin levels in patients with obstructive sleep apnoea: effect of CPAP treatment. *Eur Respir J* 2003;22:251-7.
  23. Kritikou I, Basta M, Vgontzas AN, Pejovic S, Liao D, Tsaoussoglou M, et al. Sleep apnoea, sleepiness, inflammation and insulin resistance in middle-aged males and females. *Eur Respir J* 2014;43:145-55.
  24. Perkins NJ, Schisterman EF. The inconsistency of “optimal” cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol* 2006;163:670-5.
  25. Garbarino S, Guglielmi O, Sanna A, Mancardi GL, Magnavita N. Risk of occupational accidents in workers with obstructive sleep apnea: systematic review and meta-analysis. *Sleep* 2016;39:1211-8.
  26. Mulgrew AT, Ryan CF, Fleetham JA, Cheema R, Fox N, Koehoorn M, et al. The impact of obstructive sleep apnea and daytime sleepiness on work limitation. *Sleep Med* 2007;9:42-53.
  27. Davidson TM, Patel MR. Waist circumference and sleep disordered breathing. *Laryngoscope* 2008;118:339-47.
  28. Onat A, Hergenç G, Yüksel H, Can G, Ayhan E, Kaya Z, et al. Neck circumference as a measure of central obesity: associations with metabolic syndrome and obstructive sleep apnea syndrome beyond waist circumference. *Clin Nutr* 2009;28:46-51.
  29. Wang X, Bi Y, Zhang Q, Pan F. Obstructive sleep apnoea and the risk of type 2 diabetes: a meta-analysis of prospective cohort studies. *Respirology* 2013;18:140-6.
  30. Sanders MH, Givelber R. Sleep disordered breathing may not be an independent risk factor for diabetes, but diabetes may contribute to the occurrence of periodic breathing in sleep. *Sleep Med* 2003;4:349-50.
  31. Gottlieb DJ, Yao Q, Redline S, Ali T, Mahowald MW. Does snoring predict sleepiness independently of apnea and hypopnea frequency? *Am J Respir Crit Care Med* 2000;162(4 Pt 1):1512-7.
  32. Svensson M, Franklin KA, Theorell-Haglöw J, Lindberg E. Daytime sleepiness relates to snoring independent of the apnea-hypopnea index in women from the general population. *Chest* 2008;134:919-24.
  33. Byun JI, Kim DH, Kim JS, Shin WC. Usefulness of using alternative Body-Mass Index and neck circumference criteria for STOP-Bang questionnaire in screening South Korean obstructive sleep apnea patients. *Sleep Med Res* 2020;11:38-43.