NEW RESEARCH

JCSM
Journal of Clinical Sleep Medicine

DOI: 10.5664/JCSM.1306

The STOP-Bang Equivalent Model and Prediction of Severity of Obstructive Sleep Apnea: Relation to Polysomnographic Measurements of the Apnea/Hypopnea Index

Robert J. Farney, M.D.[1]; Brandon S. Walker[1]; Robert M. Farney[1]; Gregory L. Snow, Ph.D.[2]; James M. Walker, Ph.D.[1]

[1]*Intermountain Sleep Disorders Center, LDS Hospital, Salt Lake City, UT;* [2]*Statistical Data Center, LDS Hospital, Salt Lake City, UT*

SCIENTIFIC INVESTIGATIONS

Background: Various models and questionnaires have been developed for screening specific populations for obstructive sleep apnea (OSA) as defined by the apnea/hypopnea index (AHI); however, almost every method is based upon dichotomizing a population, and none function ideally. We evaluated the possibility of using the STOP-Bang model (SBM) to classify severity of OSA into 4 categories ranging from none to severe.

Methods: Anthropomorphic data and the presence of snoring, tiredness/sleepiness, observed apneas, and hypertension were collected from 1426 patients who underwent diagnostic polysomnography. Questionnaire data for each patient was converted to the STOP-Bang equivalent with an ordinal rating of 0 to 8. Proportional odds logistic regression analysis was conducted to predict severity of sleep apnea based upon the AHI: none (AHI < 5/h), mild (AHI ≥ 5 to < 15/h), moderate (≥ 15 to < 30/h), and severe (AHI ≥ 30/h).

Results: Linear, curvilinear, and weighted models (R^2 = 0.245, 0.251, and 0.269, respectively) were developed that predicted AHI severity. The linear model showed a progressive increase in the probability of severe (4.4% to 81.9%) and progressive decrease in the probability of none (52.5% to 1.1%). The probability of mild or moderate OSA initially increased from 32.9% and 10.3% respectively (SBM score 0) to 39.3% (SBM score 2) and 31.8% (SBM score 4), after which there was a progressive decrease in probabilities as more patients fell into the severe category.

Conclusions: The STOP-Bang model may be useful to categorize OSA severity, triage patients for diagnostic evaluation or exclude from harm.

Keywords: STOP-Bang model, Berlin Questionnaire, screening questionnaire, obstructive sleep apnea syndrome, proportional odds logistic regression, polysomnography

Citation: Farney RJ; Walker BS; Farney RM; Snow GL; Walker JM. The STOP-Bang equivalent model and prediction of severity of obstructive sleep apnea: relation to polysomnographic measurements of the apnea/hypopnea index. *J Clin Sleep Med* 2011;7(5):459-465.

Obstructive sleep apnea (OSA) is a highly prevalent condition underlying numerous co-morbidities routinely encountered in all fields of medicine. OSA has been linked with cardiovascular and neurovascular diseases, metabolic disorders, and impaired neurocognitive function.[1,2] Patients with untreated sleep apnea are at increased risk for motor vehicle accidents,[3] and they have more frequent cardiorespiratory complications in the perioperative environment compared to patients diagnosed and treated prior to surgery.[4,5] Therefore, a simple and reliable method to screen certain populations is needed to identify patients with OSA. The choice of optimal screening method depends upon the major objective: to include patients with sleep apnea so that they can either be referred for appropriate sleep testing and therapy or to exclude those from potential harm to themselves or others pending appropriate evaluation. With increasing numbers of patients suspected of having OSA being referred to sleep centers, a screening method is also needed in order to prioritize those patients for urgent evaluation and therapy.

Numerous clinical prediction models have been developed based upon self-reported symptoms, demographics, anthropometric variables, and comorbidities.[6-20] Limited screening measurements have also been used independently or in combination with clinical prediction models in order to provide greater accuracy.[21-26] All of these screening methods incorporate similar variables and are compared with a single descriptive parameter, apnea/hypopnea index (AHI) derived from polysomnography (PSG). However, as a single metric of the obstructive apnea syndrome, the AHI is clearly not the gold standard. It may not necessarily correlate with symp-

BRIEF SUMMARY

Current Knowledge/Study Rationale: Numerous clinical prediction models have been developed for screening certain populations for obstructive sleep apnea syndrome (OSA); however, all use a dichotomous analysis which precludes definition of multiple severity categories. This study was based on the premise that the severity of OSA is correlated with the number of risk factors present and that the probabilities of none, mild, moderate and severe sleep apnea could be predicted by means of a proportional logistic regression analysis.

Study Impact: This study shows that a simple assessment of the number of known risk factors for OSA present in a patient can be used to estimate the severity of sleep apnea/hypopnea, thus enabling clinicians to prioritize patients for definitive testing and therapy as well as to exclude patients from possible harm due to unrecognized sleep apnea. The development of prediction models that utilize a proportional logistic regression analysis rather than a dichotomous analysis may provide more useful information.

toms such as hypersomnolence, fatigue, and cardiovascular outcomes, nor is the severity of hypoxemia reflected by this parameter. Patients with other risk factors, comorbidities, or sensitive occupations may still require therapy with an AHI < 15/h. Despite the many technical considerations and the inconsistent correlation of the AHI with specific symptoms and clinical outcome measures,[27] identification of patients with obstructive sleep apnea syndrome by every previous screening method has been ultimately based upon a binary decision (absence/presence) and the chosen threshold for AHI. Unless the distribution is highly non-normal with the majority of cases occurring at either extreme of the scale, dichotomizing a continuous variable (AHI) will misclassify subjects no matter which method is used.[28]

A recently reported meta-analysis showed that none of the screening tools function ideally and none of the questionnaires or the majority of the models will correctly identify patients with OSA.[29] For example, the STOP-Bang model (SBM) was validated as an easy and effective screening method to exclude pre-surgical patients with undiagnosed severe sleep apnea from going to the operating room.[11] Consisting of 8 yes or no questions based upon snoring, tiredness/sleepiness, observed apneas, hypertension, BMI > 35 kg/m^2, age > 50 years, neck circumference > 40 cm, and male gender, a score of any 3 affirmative responses was used to dichotomize the population. The sensitivities measured 83.6%, 92.9%, and 100%, respectively, for patients with at least mild (AHI > 5/h), moderate (AHI > 15/h), and severe (AHI > 30/h) sleep apnea. The predictive values negative measured 60.8%, 90.2%, and 100%, respectively. Thus, one could confidently exclude the possibility of at least moderate or severe OSA with a score of 0-2 affirmative responses; however, the test was less reliable in excluding mild sleep apnea. It is important to note that even those patients with AHI ≥ 5 to < 15/h may have clinical symptoms and can also become much more severe postoperatively.[4,5]

On the other hand, using a score of 3 affirmative responses, the predictive values positive measured only 42% for moderate sleep apnea and 31% for severe sleep apnea because of the high false positive rates. Because 3 positive responses are considered the same as 8, a STOP-Bang score of 3 indicates that the patient could have moderate or severe sleep apnea, but not necessarily. Therefore, if the questionnaire were to be used to select high-risk patients for urgent testing based upon standard criteria (SBM score ≥ 3) rather than to exclude them, then this model could potentially result in a higher number of patients undergoing unnecessary tests and reducing the cost-effectiveness.

We assumed that increasing the number of risk factors present not only increases the probability of having at least mild OSA but also increases the likelihood of having more severe disease. The STOP-Bang model was chosen for this study because of its ease of use and simplicity. We hypothesized that the summary score of the STOP-Bang model would correlate with the severity of the AHI and that an enhanced analysis with the proportional odds logistic regression could then be used to categorize ordinal data. Giving probabilities of disease severity on 4 levels (e.g., no disease, mild, moderate, and severe) should provide more information than a binary outcome, so that clinicians can easily consider additional factors in making decisions.

PATIENTS AND METHODS

Subjects

All patients referred to the Intermountain Sleep Disorders Center at LDS Hospital, Salt Lake City, Utah (elevation 1,500 m) who underwent diagnostic polysomnography for any reason from January 2006 to December 2008 were eligible for inclusion. If split-night studies with continuous positive airway pressure were performed, only data from the diagnostic portion of the study were used. Patients were excluded for any of the following reasons: tested while breathing supplemental oxygen, previously diagnosed by PSG or treated for sleep apnea, incomplete or absent questionnaire, incomplete demographic data, and insufficient sleep or technically inadequate polysomnography. The Intermountain Healthcare Institutional Review Board approved the study protocol and waived the patient-consent requirement (IRB#0003815).

Questionnaires

The STOP-Bang questionnaire requires yes/no responses to 8 questions (**Appendix 1**, appendices are available online at www.aasmnet.org/jcsm). For this study, responses to the STOP questionnaire were derived from equivalent questions included in a comprehensive sleep questionnaire obtained on all patients referred for PSG (**Appendix 2**). Demographic data and anthropometric measurements (Bang) were obtained at the same time polysomnography was performed. The use of equivalent questions was validated by prospectively administering the STOP questionnaire and the standard sleep lab questionnaire simultaneously to a second group (200 subjects).

Polysomnography

Standard attended 17-channel polysomnography (Cadwell Laboratories, Inc., Kennewick WA.) was performed, consisting of frontal and/or central and occipital electroencephalogram, right and left electrooculogram, and submentalis electromyogram. Airflow was detected by air pressure transducers (PTAF II, Pro-Tech Services, Inc. Mukilteo, WA), and respiratory effort was determined by measurement of chest and abdominal motion with piezoelectric or respiratory inductive plethysmography transducers. The arterial oxygen saturation (SpO_2) was measured by the Cadwell oximeter with a 4-beat averaging mode. Sleep was manually scored using standard criteria.[30,31] Apneas were scored on the basis of absence of airflow ≥ 10 sec. Obstructive apneas were defined by the presence of respiratory effort; central apneas by the absence of respiratory effort. For patients studied prior to April 2007, scoring of hypopneas was based on earlier criteria.[32] After this date, hypopneas were based upon current standards.[31] Hypopneas were not differentiated as obstructive or central.[32] The apnea/hypopnea index (AHI) was computed as the total of apneas and hypopneas divided by the total sleep time in hours.

Statistical Analysis

The sleep laboratory STOP score was compared to the STOP questionnaire by a Wilcoxon matched pairs test (2-tailed, $p < 0.05$). Individual components of the STOP questionnaire were compared with sleep laboratory equivalent questions by using a nonparametric sign test.

Table 1—Comparison of means and standard deviations of the individual and composite or sum of responses for STOP[1] questions and sleep laboratory STOP equivalent questions for 200 patients

	Composite Score	S	T	O	P
STOP Questionnaire	2.29 (1.05)	0.61 (0.49)	0.84 (0.36)	0.47 (0.50)	0.36 (0.48)
STOP Equivalent Questionnaire	2.25 (1.00)	0.57 (0.50)	0.85 (0.35)	0.44 (0.50)	0.38 (0.49)
p-value[2]	0.460	0.281	0.850	0.326	0.372

[1]STOP refers to a positive response to snoring, tired/sleepiness, observed apneas, and hypertension. The composite score is the sum of all 4 questions (range 0-4), whereas the individual questions for each question is either a yes (1) or no (0) response. [2]Probabilities of comparisons between the STOP and STOP Equivalent questionnaire. A Wilcoxon matched pairs test was used to compare the composite score on the STOP questionnaire whereas a signed test was used to test individual questions.

Table 2—Summary statistics of the patient population

Parameter	All	Males	Females
N	1426	815	611
Snoring[1] (%)	62	70	51
Tired/Sleepy[1] (%)	89	88	90
Observed Apnea[1] (%)	48	61	32
Pressure[1] (%)	42	42	42
BMI[2] (kg/m^2)	33.8 ± 8.1 (16-71)	33.0 ± 7.4 (17-71)	34.9 ± 8.9 (16-68)
Age[2] (y)	49.7 ± 15.2 (15-91)	48.8 ± 15.2 (15-91)	51.0 ± 15.2 (16-88)
Neck size[2] (cm)	40.7 ± 4.8 (26-61)	42.8 ± 4.2 (28-61)	38.0 ± 4.1 (26-53)
Gender (%)		57.2	42.8
STOP-Bang Score[2]	4.3 ± 1.7 (0-8)	5.1 ± 1.5 (1-8)	3.3 ± 1.4 (0-7)
AHI[3]	32.9 ± 30.1 (0-147)	37.7 ± 30.3 (0-142)	26.4 ± 28.6 (0-162)

[1]Percent positive responses to the first 4 questions or STOP questionnaire (snoring, tired/sleepy, observed apnea, and pressure). [2]Values for BMI, age, neck size, and composite score for the STOP-Bang model are reported as Mean ± SD (Range). [3]AHI or apnea/hypopnea index is the number of apneas plus hypopneas divided by hours of total sleep time while breathing room air reported as Mean ± SD (Range).

The primary outcome measure was the severity of sleep apnea based upon the AHI as categorized into 4 groups: none (AHI < 5/h), mild (AHI ≥ 5 to < 15/h), moderate (AHI ≥ 15 to < 30/h), and severe (AHI ≥ 30/h). A proportional odds logistic regression analysis[33] was conducted using the total score of the STOP-Bang equivalent model as the independent variable with equal weight given to each response (linear model). Curvilinear and weighted models were analyzed similarly. The curvilinear model used the total score with cubic splines to allow for a possible curved relationship. The weighted model used each of the 8 questions as independent variables, allowing for each to have a different effect on the model. Coefficient of determination or the proportion of variance was calculated (R^2) and the area under the receiver operator curve was expressed as a C-statistic. In order to assess whether changes in definitions and instruments during the data collection period had any effect on the model, an additional proportional odds model was fit to the first 500 patients using earlier criteria for hypopneas[32] and the last 500 patients using the most recent criteria.[31] This included all 8 of the STOP-Bang terms as well as an indicator variable for time grouping and all interactions between this grouping variable and the STOP-Bang questions. This model was then compared to a model on the same data without the grouping variable and interactions. The 2 models were then compared using a full vs. reduced likelihood ratio test.

RESULTS

Comparison of the sum of STOP questions to the sleep lab equivalent questions indicated no difference between the 2 forms of questions. Similarly, there were no differences between the individual questions of the STOP and the sleep lab equivalent questions (**Table 1**).

Descriptive summary statistics for the patient population reflected by the STOP-Bang model and the AHI are displayed in **Table 2**. The mean ± SD scores for the STOP-Bang model for males measured 5.1 ± 1.5 compared to 3.3 ± 1.4 for females. Because male gender is one of the 8 variables, the maximum score for males is 8, versus 7 for females. Snoring and observed apneas were reported more frequently by males but the responses to tired/sleepy and history of hypertension were similar. BMI and age were similar; however, males had greater neck size (42.8 ± 4.2 cm versus 38.0 ± 4.1 cm), and their AHI was more severe (37.8 ± 30.3/h versus 26.4 ± 28.6/h). The distributions of the STOP-Bang summary scores for the entire group, males, and females are shown in **Figure 1**.

Three models were constructed using proportional odds logistic regression (**Table 3**). Models based upon only males and females showed no differences compared to the models based upon all subjects. Therefore, only group data are presented. Model one was based upon equal weight of each of the 8 re-

Figure 1—Distribution of affirmative STOP-Bang responses for all (1,426), males (815), and females (611)

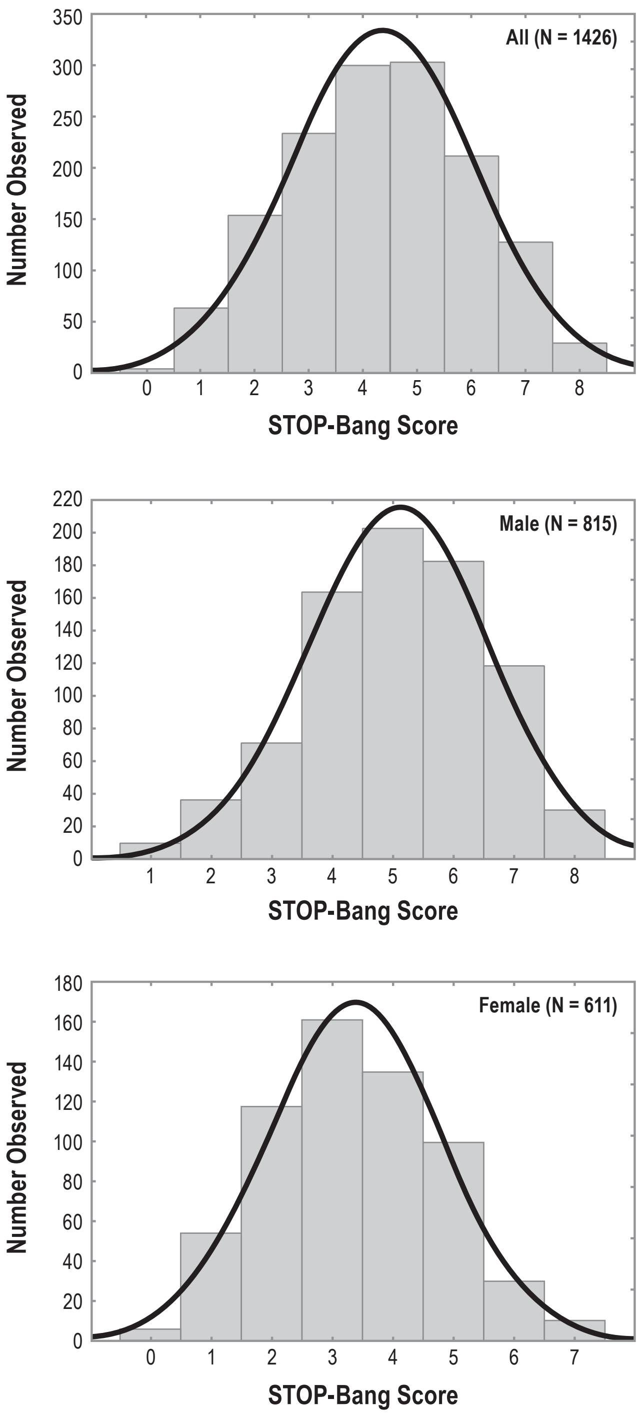


Table 3—The coefficient of determination (R^2) and C-Statistic for STOP-Bang models

Model	R^2	C Statistic
Linear	0.245	0.712
Curvilinear	0.251[1]	0.712
Weighted	0.269[2]	0.723

[1]p < 0.05 improvement over linear model. [2]p < 0.001 improvement over linear model.

Table 4A—Number of patients with sleep apnea vs. positive STOP-Bang Model responses

	Apnea/Hypopnea Index			
STOP-Bang Score[1]	< 5/hr	≥ 5 to < 15/hr	≥ 15 to < 30/hr	≥ 30/hr
0	3	0	1	0
1	32	17	9	7
2	52	54	29	18
3	32	67	70	64
4	15	82	96	106
5	10	62	92	138
6	4	22	56	129
7	2	9	23	93
8	0	2	3	25

[1]Numerical sum of the positive responses of the STOP-Bang model questionnaire.

Table 4B—Predicted probability of sleep apnea severity based upon a linear model of the total population studied vs. positive STOP-Bang Model responses

	Apnea/Hypopnea Index			
STOP-Bang Score[1]	< 5/hr	≥ 5 to < 15/hr	≥ 15 to < 30/hr	≥ 30/hr
0	52.5%	32.9%	10.3%	4.4%
1	41.1%	37.6%	14.6%	6.7%
2	30.6%	39.3%	19.8%	10.3%
3	14.9%	33.2%	29.6%	22.3%
4	10.0%	26.9%	31.8%	31.3%
5	6.5%	20.4%	31.1%	41.9%
6	4.2%	14.7%	27.8%	53.3%
7	1.7%	6.8%	17.4%	74.1%
8	1.1%	4.4%	12.5%	81.9%

[1]Numerical sum of the positive responses of the STOP-Bang model questionnaire.

sponses on the SBM. The proportion of variance accounted for by this model was 0.245 with the C-statistic area under the operator receiver curve of 0.712. The second model was significantly improved ($p = 0.016$), but there were only marginal changes in R^2 (0.251) and the C-statistic (0.712). The third model developed on weighting each response indicated a significant improvement ($p < 0.001$) as compared to the first model (R^2 0.269, C-statistic 0.723). The test on the first 500 and last 500 patients did not show any effect of the group variable ($p = 0.125$), indicating no significant changes in the criteria used for scoring hypopnea and the overall effect of the STOP-Bang model.

In order to maintain simplicity and because there was no major advantage of using the more complex weighted model (C Statistic 0.712), the linear model (C Statistic 0.723) was derived from the entire patient population to estimate the probability of having no, mild, moderate, or severe sleep apnea as defined above. The number of patients in each category of sleep apnea severity (AHI) versus the STOP-Bang model scores are shown in **Table 4A**. The predicted probabilities of sleep apnea severity based upon a linear model of the total population studied versus the STOP-Bang model score are shown in **Table 4B**. With each incremental increase in the score from 0 to 3, the probability of having no sleep apnea diminished, while the probability of having mild, moderate, or severe sleep apnea increased. With any score > 3, the probability continuously increased for having severe sleep apnea, while the probability for anything else decreased.

DISCUSSION

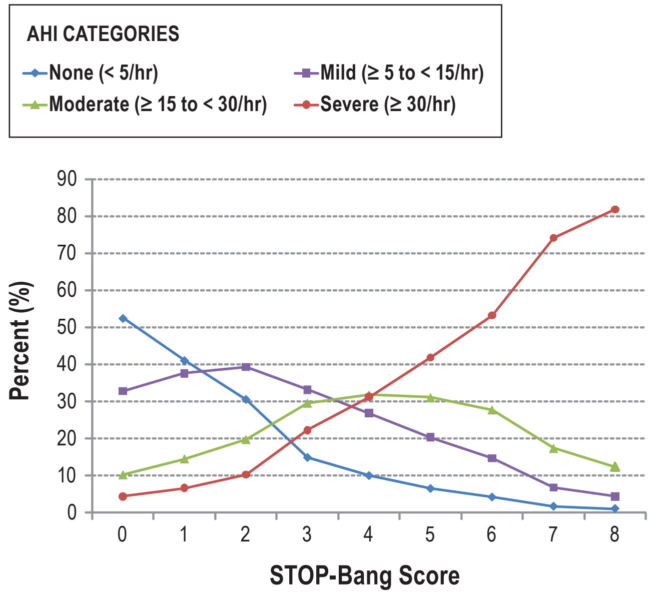
Clinical Importance

This study confirms that in a population referred for evaluation of various sleep disorders, the STOP-Bang[11] or equivalent model can be used to estimate the probabilities of no, mild, moderate, and severe obstructive sleep apnea as defined by the AHI. As opposed to dichotomizing a continuous variable in which there is no distinction between a STOP-Bang score of 3 and 8, enhanced analysis with the proportional odds logistic regression provides a set of probabilities for various levels of severity. These can then be used as a guide within a given clinical context to make more reasoned diagnostic decisions. Although intuitive, this study indicates that the greater a single cumulative score of known risk factors as reflected by the STOP-Bang model, the greater the probability of more severe sleep apnea (**Figure 2**).

Of the various models to screen for OSA, the Berlin Questionnaire[8] has possibly been the most frequently used. It consists of 11 questions with multiple choices and organized into three categories (snoring/apnea, fatigue/sleepiness, and obesity/hypertension). Certain important predictive demographic parameters were not utilized (gender, age, and neck circumference). Of 1008 adults surveyed, 100 underwent *unattended* cardiorespiratory sleep testing and were then dichotomized into a low- or a high-risk group at various cut points. The sensitivity for having an AHI > 5/h measured 86%. The greatest limitations are that the scoring can be confusing or cumbersome and that in the end, patients are still categorized into only low-risk or high-risk groups.

The STOP-Bang model has the appeal of being very straightforward and requires no more than a few minutes to complete and score. Of 2,467 subjects screened in pre-surgical clinics, it was validated with attended polysomnography in 177 patients. Consistent with Chung et al.,[11] who reported a sensitivity of 83.6% for AHI > 5/hr, our data shows that there is an 85.1% probability of having an AHI ≥ 5/hr if the score ranges from 3 to 8. Therefore, if the only intent is to exclude patients who may have any degree of OSA from undergoing surgery, then the SBM appears to be a valid initial screening device. However, it does not provide sufficient direction for individual clinical decision-making regarding the need or urgency for polysomnography and therapy. For AHI > 15/hr and > 30/hr, the specificity measured 43% and 37%, and the predictive value positive measured 51.6% and 31%, respectively.[11] Analyzing the data with various cut-points is useful but does not avoid the dilemma of grouping patients on either side of the boundary with the extremes at the ends of the spectrum, regardless of which cut point is selected. With a binary approach, a patient with a 51% chance of having moderate or severe sleep apnea could be treated exactly the same as one with a 90% chance—while being treated differently than someone with a 49% chance—with no option for the physician to include other known information about the patient.

Figure 2—Predicted probability of sleep apnea severity based upon a linear model of the total population studied vs. affirmative STOP-Bang Model responses



When the composite score is zero, there is a 52.5% probability of having no sleep apnea and little probability of having either moderate or severe sleep apnea (10.3% and 4.4%, respectively). There is 32.9% probability of having mild OSA with a zero score, which reflects the false negative rate that could be relevant depending upon the purpose of the screening method. With each incremental increase in the score from 0 to 3, the probability of having no sleep apnea diminishes, while the probability of having moderate or severe sleep apnea increases. With a composite score of 3, there is a 14.9% probability of having no sleep apnea and a 33.2%, 29.6%, and 22.3% probability of mild, moderate, and severe sleep apnea, respectively (overall 85% probability for any degree of OSA). With any score > 4, the probability increases continuously for having severe sleep apnea, while the probability for anything else decreases. With a score of 8, the probability of severe sleep apnea is 81.9%.

To the extent that the AHI is a valid parameter for defining this complex syndrome, the composite STOP-Bang model score provides a method for stratifying patients into categories of sleep apnea. For example, if the SBM score is 6-8, then the corresponding probability of having severe OSA is quite high (53.3%, 74.1%, and 81.9%, respectively), and further evaluation is relatively urgent. The physician can weigh a 53% chance of severe sleep apnea with other factors to decide on the appropriate clinical course of action, but would likely balance this differently for someone with a 90% or 10% probability.

If the score is 3-5, then the probabilities of mild, moderate, and severe OSA are more evenly balanced. Greater discrimination between cases of no, mild, and moderate sleep apnea may not be possible due to the intrinsic deficiencies of the STOP-Bang model and limitations imposed by defining severity of OSA by only the AHI. In these cases, secondary screening, for example with oximetry or limited cardiopulmonary sleep testing, and consideration of additional clinical details would be appropriate. The presence of clinically significant sleep disordered breathing found in patients with a composite score of 0-2 suggests that the number of false negatives may be substantial and that there are other risk factors not identified by this questionnaire. Note, however, that there were relatively few subjects in these low-risk categories, probably because specific population tested consisted predominantly of patients referred for OSA.

Limitations and Strengths of the Present Study

A potential limitation of this retrospective study is that we used a standard sleep questionnaire from our laboratory to obtain the equivalent responses for the STOP questionnaire (Appendix 2); however, the categorization of ordinal responses of the sleep laboratory questionnaire were not different than the STOP responses obtained on a subset of 200 patients.

An assumption of the proportional odds logistic regression analysis is that the regression functions for different response categories are parallel on the logit scale and the sum of the probabilities must equal 100%. Of interest, each component of the STOP-Bang model is itself a dichotomy. Three of the eight variables are continuous (BMI, age, and neck circumference), while three can be quantified on an ordinal scale (snoring, tiredness, and observed apneas); however, reducing these variables to a binary response is necessary for the simplicity of the SBM. In our own study, we showed that there is a positive correlation between the ordinal scale of the subjective symptoms (STOP) and between increasing BMI, age, and neck circumference that make up the demographic data (Bang) with the AHI.

A degree of underlying uncertainty regarding the probabilities distribution may be created because each point of the STOP-Bang score was given equal weight. In fact, our data indicate that a weighted model for each response improves prediction over the linear model, but only slightly. More research needs to be done and part of that would be to revisit the curvilinear and weighted models with additional data from a more general population. With additional data, it will be easier to detect differences in the effects of the STOP-Bang predictors so that models with better information may be derived. But the incorporation of these weights detracts from the simplicity and ease of use and with the current data, the difference in the predictions from the weighted and curvilinear models does not justify the additional complication.

The SBM also does not include other risk factors, such as craniofacial morphology, and comorbidities such as diabetes, coronary artery disease, and stroke. A critical limitation of all screening studies for sleep apnea, is that central apneas (i.e., idiopathic central sleep apnea, Cheyne-Stokes breathing, and Biot's respiration) are ignored in the final analysis, even if obstructive and central events are differentiated in the methods. The extent to which this methodology practice contributes to the false negative rates is not known but could be substantial, particularly in the older patients with cardiac or neurologic disease or in those receiving chronic opioids.[34,35]

Lastly, these findings may be criticized because we studied a population referred for evaluation of sleep disorders. Suspected sleep apnea syndrome was the predominant reason for referral to our sleep center; however, we included all patients being evaluated with polysomnography regardless of the reason for referral. Every study used to screen for OSA may be criticized because the population tested is biased in one way or another. In the majority, subjects have been referred for sleep testing because of suspected sleep apnea. Other studies are also likely biased because conditions such as hypertension, depression, obesity, diabetes, use of opioids, and other risk factors are intrinsic in any clinic population (general medicine or surgery).

For patients referred to a sleep center, we believe that this study has validity because of the large number of patients examined with attended polysomnography (1426), the broad patient mix (which includes patients not referred for OSA), and the ability to stratify patients into none, mild, moderate, and severe risk categories. The statistical method used (proportional odds logistic regression) gives greater information than a yes/no risk determination. Whether or not the probabilities observed in this study apply exactly to other clinic or general populations is not known, but the concept of increasing risk factors being associated with greater likelihood for the presence of more severe sleep apnea is intuitive and seems reasonable. It is implausible that a SBM score of 7-8 in any population would be less predictive of severe OSA; however, the accuracy of discriminating no, mild, and moderate disease in a less selective population with a lower prevalence of OSA is unknown and the validity and application of these results in a more general or different clinical population must be established.

The STOP-Bang questionnaire has already been successfully applied in preoperative screening.[4,36] Despite the various limitations noted, this study builds on that experience and provides a simple, practical basis for initial screening and risk stratification in a population referred for the evaluation of sleep disorders. This may be particularly valuable for sleep centers in order to identify those patients most in need of urgent evaluation

CONCLUSION

Ultimately, the decision to test or treat any individual is reached on the basis of the clinical context and a determination of the severity of sleep disordered breathing. Although the AHI is an imperfect parameter, it remains widely used in practice. Being able to estimate the severity of sleep apnea with a range of probabilities may enable one to make more appropriate decisions for an individual regarding the need for testing with polysomnography. In clinical practice, these decisions are based on consideration of numerous factors and the probabilities of various conditions that may be contributing to symptoms. Although there are limitations, the STOP-Bang model provides a simple method for screening, estimating the AHI severity, and triaging patients for testing. Using the SBM or some other equivalent questionnaire to provide sets of probabilities may provide greater guidance to clinical decision making than a dichotomous analysis. This concept needs to be verified in a larger prospective clinical trial. No matter what tool is employed, further intervention depends upon clinical judgment before embarking upon expensive tests with limited resources.

ABBREVIATIONS

AHI, Apnea plus hypopnea index
BMI, Body mass index
OSA, Obstructive sleep apnea
PSG, Polysomnography
SBM, STOP-Bang model

REFERENCES

1. Jennum P, Riha RL. Epidemiology of sleep apnoea/hypopnoea syndrome and sleep-disordered breathing. *Eur Respir J* 2009;33:907-14.
2. McNicholas WT, Bonsignore MR and the Management Committee of EU COST ACTION B26. Sleep apnoea as an independent risk factor for cardiovascular disease: current evidence, basic mechanisms and research priorities. *Eur Respir J* 2007;29:156-78.
3. Hartenbaum N, Collop N, Rosen IM, et al. Sleep apnea and commercial motor vehicle operators. Statement from the Joint Task Force of the American College of Chest Physicians, the American College of Occupational and Environmental Medicine, and the National Sleep Foundation. *Chest* 2006;130:902-5.
4. Chung SA, Yuan H, Chung F. A systematic review of obstructive sleep apnea and its implications for anesthesiologists. *Anesth Analg* 2008;107:1543-63.
5. Gupta RM, Parvizi J, Hanssen AD, et al. Postoperative complications in patients with obstructive sleep apnea syndrome undergoing hip or knee replacement: A case-control study. *Mayo Clin Proc* 2001;76:897-905.
6. Rowley JA, Aboussouan LS, Badr MS. The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. *Sleep* 2000;23:929-38.
7. Ross SD, Sheinhait IA, Harrison KJ, et al. Systematic review and meta-analysis of the literature regarding the diagnosis of sleep apnea. *Sleep* 2000;23:1-14.
8. Netzer NC, Stoohs RA, Netzer CM, et al. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. *Ann Intern Med* 1999;131:485-91.
9. Sharma SK, Vasudev C, Sinha S, et al. Validation of the modified Berlin Questionnaire to identify patients at risk for the obstructive sleep apnoea syndrome. *Indian J Med Res* 2006;124:281-90.
10. Chung F, Yegneswaran B, Liao P, et al. Validation of the Berlin Questionnaire and American Society of Anesthesiologists Checklist as screening tools for obstructive sleep apnea in surgical patients. *Anesthesiology* 2008;108:822-30.
11. Chung F, Yegneswaran B, Liao P, et al. STOP Questionnaire. A tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008;108:812-21.
12. Flemons WW, Whitelaw WA, Brant R, et al. Likelihood ratios for a sleep apnea clinical prediction rule. *Am J Respir Crit Care Med* 1994;150:1279-85.
13. Maislin G, Pack AI, Kribbs NB, et al. A survey screen for prediction of apnea. *Sleep* 1995;18:158-66.
14. Rodsutti J, Hensley M, Thakkinstian A, et al. A clinical decision rule to prioritize polysomnography in patients with suspected sleep apnea. *Sleep* 2004;27:694-9.
15. Hoffstein V, Szalai JP. Predictive value of clinical features in diagnosing obstructive sleep apnea. *Sleep* 1993;16:118-22.
16. Dixon JB, Schachter LM, O'Brien PE. Predicting sleep apnea and excessive day sleepiness in the severely obese. Indicators for polysomnography. *Chest* 2003;123:1134-41.
17. Viner S, Szalai JP, Hoffstein V. Are history and physical examination a good screening test for sleep apnea? *Ann Intern Med* 1991;115:356-9.
18. Schellenberg JB, Maislin G, Schwab RJ. Physical findings and the risk for obstructive sleep apnea. The importance of oropharyngeal structures. *Am J Respir Crit Care Med* 2000;162:740-8.
19. Tsai WH, Remmers JE, Brant R, et al. A decision rule for diagnostic testing in obstructive sleep apnea. *Am J Respir Crit Care Med* 2003;167:1427-32.
20. Kirby SD, Eng P, Danter W, et al. Neural network prediction of obstructive sleep apnea from clinical criteria. *Chest* 1999;116:409-15.
21. Deegan PC, McNicholas WT. Predictive value of clinical features for the obstructive sleep apnoea syndrome. *Eur Respir J* 1996;9:117-24.
22. Gurubhagavatula I, Maislin G, Nkwuo JE, et al. Occupational screening for obstructive sleep apnea in commercial drivers. *Am J Respir Crit Care Med* 2004;170:371-6.
23. Farney RJ, Walker LE, Jensen RL, et al. Ear oximetry to detect apnea and differentiate rapid eye movement (REM) and non-REM (NREM) sleep. Screening for the sleep apnea syndrome. *Chest* 1986;89:533-9
24. Williams AJ, Yu G, Santiago S, Stein M. Screening for sleep apnea using pulse oximetry and a clinical score. *Chest* 1991;100:631-5.
25. Netzer N, Eliasson AH, Netzer C, et al. Overnight pulse oximetry for sleep-disordered breathing in adults. A review. *Chest* 2001;120:625-33.
26. Magalang UJ, Dmochowski J, Veeramachaneni S, et al. Prediction of the apnea-hypopnea index from overnight pulse oximetry. *Chest* 2003;124:1694-701.
27. McNicholas WT. Diagnostic criteria for the sleep apnoea syndrome: time for consensus? *Eur Respir J* 1996;9:634-5.
28. Altman DG, Royston P. The cost of dichotomizing continuous variables. *BMJ* 2006;332:1080.
29. Ramachandran SK, Josephs LA. A meta-analysis of clinical screening tests for obstructive sleep apnea. *Anesthesiology* 2009;110:928-39.
30. Rechtschaffen A, Kales AD. *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. Los Angeles, CA: UCLA Brain Information Service/Brain Research Institute, 1968.
31. Iber C, Ancol-Isreal S, Chesson AL, et al. *The AASM manual for scoring of sleep and associated events*. Westchester, IL: American Academy of Sleep Medicine, 2007
32. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The report of an American Academy of Sleep Medicine Task force. *Sleep* 1999;22:667-89.
33. Venables WN, Ripley BD. *Modern Applied Statistics with S, 4th ed.* Springer, 2002.
34. Eckert DJ, Jordan AS, Merchia P, et al. Central sleep apnea. Pathophysiology and treatment. *Chest* 2007;131:595-607.
35. Walker JM, Farney RJ, Rhondeau SM, et al. Chronic opioid use is a risk factor for the development of central sleep apnea and ataxic breathing. *J Clin Sleep Med* 2007;3:455-61.
36. Vasu TS, Doghramji K, Cavallazzi R, et al. Obstructive sleep apnea syndrome and postoperative complications: clinical use of the STOP-BANG questionnaire. *Arch Otolaryngol Head Neck Surg* 2010;136:1020-4.
37. Kobayashi M, Kazuyoshi N, Tsuiki S, et al. Clinical characteristics in two subgroups of obstructive sleep apnea syndrome in the elderly. Comparison between cases with elderly and middle-age onset. *Chest* 2010;137:1310-5.

ACKNOWLEDGMENTS

The authors acknowledge R.T. Nuttall R.R.T., R.C.P., and Marabeth Morton, R.N., for their crucial help in data entry/management, and to Jill Rhead, M.A., for her help in the creation of the figures. This work was performed at Intermountain Sleep Disorder Center, Salt Lake City, Utah. Data from this work was presented in abstract form at the 24th Annual Meeting of the Associated Professional Sleep Societies in San Antonio, Texas on 9 Jun 2010.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication December, 2010
Submitted in final revised form March, 2011
Accepted for publication March, 2011
Address correspondence to: Robert J. Farney, M.D., Intermountain Sleep Disorders Center, LDS Hospital, 325 8th Ave & C Street, Salt Lake City, Utah 84143; Tel: (801) 408-3617; Fax: (801) 408-5110; E-mail:rjfmd@msn.com

DISCLOSURE STATEMENT

This was not an industry supported study. The authors have indicated no financial conflicts of interest.

Appendix 1—STOP-Bang model questionnaire[11]

1. ***S*noring**
   Do you snore loudly (louder than talking or loud enough to be heard through closed doors)?
   Yes No
2. ***T*ired**
   Do you often feel tired or sleepy during daytime?
   Yes No
3. ***O*bserved**
   Has anyone observed you stop breathing during your sleep?
   Yes No
4. **Blood *P*ressure**
   Do you have or are you being treated for high blood pressure?
   Yes No
5. ***B*MI**
   BMI more than 35 kg/m^2
   Yes No
6. ***A*ge**
   Age over 50 years old?
   Yes No
7. ***N*eck Circumference**
   Neck Circumference greater than 40 cm
   Yes No
8. ***G*ender**
   Gender male?
   Yes No

High risk of OSA (AHI ≥ 5/hr): answering yes to three or more items
Low risk of OSA (AHI < 5/hr): answering yes to less than three items

Appendix 2

STOP questionnaire	Sleep center questionnaire	Considered equivalent if
(S) Do you snore loudly (louder than talking or loud enough to be heard through closed doors)?	Have you ever had loud snoring? Yes/No. **AND** When you snore, it is usually	Yes **AND** (2) Moderately loud (disturbs my bed partner) **OR** (3) Extremely loud (can be heard by people outside of my bedroom) **OR** (4) Beyond loud (offends my neighbors)
(T) Do you often feel tired, fatigued, or sleepy during daytime?	How often do you have a problem throughout the day due to tiredness and fatigue? **OR** Following your usual night's sleep, do you have difficulty with becoming drowsy when not physically active (for example: while reading, watching television, at movies)?	(3) Frequently **OR** (4) Almost always **OR** (2) Occasionally **OR** (3) Frequently **OR** (4) Almost always
(O) Has anyone observed you stop breathing during your sleep?	As far as either you or your bed partner know, your breathing during sleep is	(3) frequently interrupted by long pauses **OR** (4) continuously interrupted by pauses
(P) Do you have or are you being treated for high blood pressure?	Have you ever been diagnosed or treated for hypertension (high blood pressure)?	Yes

Comparison between the STOP questionnaire and the equivalent responses derived from the Intermountain Sleep Disorders Center questionnaire, which uses an ordinal rating of 0 to 4 for questions regarding snoring, tiredness/sleepiness and observed apneas. The wording of the question regarding diagnosis or therapy of hypertension is essentially identical and both questionnaires require a yes/no response.