

Alternative Scoring Models of STOP-Bang Questionnaire Improve Specificity To Detect Undiagnosed Obstructive Sleep Apnea

Frances Chung, M.B.B.S.; Yiliang Yang, M.D.; Russell Brown, M.D.; Pu Liao, M.D.

Department of Anesthesia, Toronto Western Hospital, Women's College Hospital, University of Toronto, Toronto, Ontario, Canada

Background: Obstructive sleep apnea (OSA) is common among surgical patients. The STOP-Bang questionnaire is a validated screening tool with a high sensitivity. However, its moderate specificity may yield fairly high false positive rate. We hypothesized that the specific combinations of predicting factors in the STOP-Bang questionnaire would improve its specificity.

Methods: After research ethics approval, consented patients were asked to complete the STOP-Bang questionnaire and then underwent sleep studies. The predictive performance of the STOP-Bang alternative scoring models was evaluated. Five hundred sixteen patients with complete data on the STOP-Bang questionnaire and polysomnography were reported.

Results: When the STOP-Bang score was ≥ 3 (any 3 positive items), the sensitivity and specificity for identifying moderate-severe OSA was 87% and 31%, respectively. The specificity for any 2 positive items from the 4 STOP questions plus BMI > 35 kg/m², male gender, or neck circumference > 40 cm for identifying moderate-severe OSA was 85%, 77%, and 79%, respectively. Compared with STOP-Bang score

≥ 3 , the predicted probability for severe OSA of the specific combinations of STOP score ≥ 2 + male and STOP score ≥ 2 + BMI increased by 36% and 42%, respectively. For severe OSA, the specific combination of STOP score ≥ 2 + BMI + male demonstrated a specificity of 97% and 86% increase in predicted probability versus any 4 positive items of STOP-Bang questionnaire.

Conclusions: The specific constellations of predictive factors improved the specificity of STOP-Bang questionnaire. For patients with STOP score ≥ 2 , male gender, and BMI > 35 kg/m² were more predictive than age ≥ 50 and neck circumference > 40 cm.

Keywords: obstructive sleep apnea, polysomnography, perioperative care, preoperative screening, sleep apnea questionnaire

Citation: Chung F, Yang Y, Brown R, Liao P. Alternative scoring models of STOP-Bang questionnaire improve specificity to detect undiagnosed obstructive sleep apnea. *J Clin Sleep Med* 2014;10(9):951-958.

The current estimated prevalence of moderate-severe obstructive sleep apnea (OSA, defined by an apnea-hypopnea index [AHI] > 15 events/h) is 17% among 50 to 70-year-old men and 9% among 50 to 70-year-old women.¹ It appears to be prevalent in surgical patients as well as the general population.^{2,3} Up to 80% of men and 93% of women with moderate-severe sleep apnea are undiagnosed,⁴ and a large proportion of surgical patients with OSA remain unrecognized.⁵ Identifying OSA preoperatively might help reduce postoperative complications.^{2,6} The proper identification and management of patients with continuous positive airway pressure (CPAP) appears to nullify the risk of OSA, prevent postoperative exacerbation of OSA, and may even improve patients' health well beyond the perioperative period.^{7,8}

The STOP-Bang questionnaire (**Appendix**) has been developed and validated as a screening tool for OSA in surgical patients. It has been proved to be valuable for screening in the preoperative clinic,² as well as non-surgical populations.⁹ The standard STOP-Bang scoring classifies patients with any 3 positive items from 8 questions as having risk of OSA. The sensitivity of STOP-Bang score 3 or greater is 84%, 93%, and 100% to predict all OSA (AHI > 5), moderate-severe OSA (AHI > 15) and severe OSA (AHI > 30), respectively.² Due to its high sensitivity, the STOP-Bang questionnaire is effective to screen patients with OSA in the preoperative clinical setting and the sleep clinic.⁹

BRIEF SUMMARY

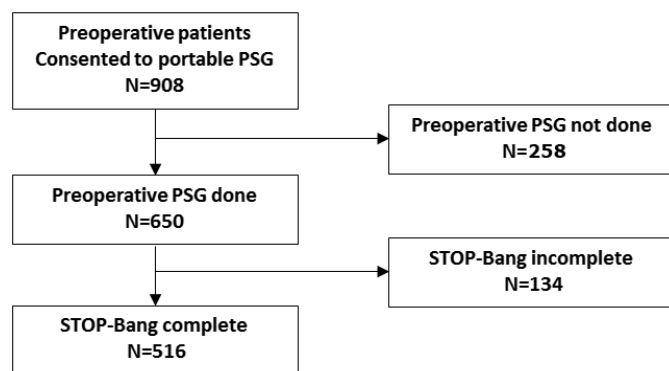
Current Knowledge/Study Rationale: The STOP-Bang questionnaire is a validated screening tool with a high sensitivity. Its moderate specificity may yield a high false-positive rate. The specific combinations of predicting factors in the STOP-Bang questionnaire may improve its specificity.

Study Impact: This study shows that the specific constellations of predictive factors improved the specificity of STOP-Bang questionnaire. For patients with a STOP score ≥ 2 , male gender, and BMI > 35 kg/m² were more predictive of obstructive sleep apnea than age ≥ 50 and neck circumference > 40 cm.

Due to its moderate specificity (56%, 43%, and 37% for predicting mild to severe OSA, moderate-severe OSA, and severe OSA, respectively), the STOP-Bang questionnaire may result in false positives leading to unnecessary referrals for sleep studies.

The predictive analysis at different cutoff values shows that a STOP-Bang score 0-2 essentially excludes patients with moderate-severe OSA, whereas a STOP-Bang score 5 or greater was highly specific for sleep apnea ($\sim 80\%$).¹⁰ The specificity of STOP-Bang questionnaire can also be improved by combining a STOP-Bang score with an elevated serum bicarbonate.¹¹

The predictive values of each individual item in the STOP-Bang questionnaire appear different.^{2,12} The objective of this study was to explore the predictive performance of the different

Figure 1—Flow chart of patient recruitment.

combinations of items from “Bang” with the STOP component, with a main focus on specificity. We hypothesized that the specificity of the score can be improved by the specific combinations of items in the STOP-Bang questionnaire.

METHODS

Patient Population

The study was conducted in the preoperative clinics of Toronto Western Hospital and Mount Sinai Hospital, Toronto, Ontario, Canada. Institutional Review Board approvals were obtained from both institutions (MSH: 07-0183-E; UHN: 07-0515-AE). Surgical patients aged 18 years or older who had an American Society of Anesthesiologists (ASA) physical status I-IV were approached to obtain consent for preoperative polysomnography. Patients with abnormal electroencephalographic findings (e.g., brain tumor, epilepsy surgery) were excluded.

Preoperative Screening with the STOP-Bang Questionnaire and Polysomnography

Consented patients were asked to complete the STOP questionnaire. Information regarding body mass index (BMI), age, neck circumference, and gender (Bang) were collected by a research assistant. The consented patients were invited to undergo a portable polysomnography study at home. The results of the PSG were used to evaluate the various scores of the STOP-Bang questionnaire.

The portable polysomnography was performed with a level 2 portable sleep device (Embletta X100), which is shown to be a reliable alternative for standard polysomnography in surgical patients.¹³ The polysomnographic recordings were performed at the patient’s home, where the device was set up by a trained polysomnography technician. The overnight recording was unattended. The recording montage consisted of 2 EEG channels (C3 and C4), electrooculogram (left or right), and chin muscle EMGs. Thoracic and abdominal respiratory effort bands, body position sensors, and pulse oximeter were also used. Under the supervision of a sleep physician, PSG recordings were scored by a certified polysomnography technologist blinded to the results of the STOP-Bang questionnaire. According to the manual

published by American Academy of Sleep Medicine in 2007 (the American Academy of Sleep Medicine Task Force recommendations),¹⁴ apnea was defined as $\geq 90\%$ drop in air flow from baseline lasting ≥ 10 seconds. Hypopnea was defined as $\geq 50\%$ reduction in air flow lasting ≥ 10 sec and is associated with $\geq 3\%$ decrease in arterial oxyhemoglobin saturation or associated with arousal. Apneas were classified as obstructive if respiratory efforts was present and central if respiratory effort was absent during the event. The diagnosis of OSA was based on PSG results. The patients with AHI > 5 events/h were considered having OSA. Previously reported data, which was involved in evaluating the predictive performance of a high STOP-Bang score,¹⁰ were used to assess the particular combinations of STOP-Bang scores to predict OSA.

Statistical Methods

Statistical analyses were performed using SAS version 9.2. The demographic data was presented with descriptive statistics. Median and interquartile range were used for non-normally distributed continuous data, and comparisons were done using Mann-Whitney U test. Categorical data were presented as frequency and statistical significance was checked by χ^2 or Fisher exact test. Continuous data that were normally distributed were presented as mean \pm SD, and the student *t*-test was used to calculate the p value. Comparisons were considered statistically significant if the p-value was ≤ 0.05 .

The relationship between AHI and age, BMI, and neck circumference, respectively, was assessed by Pearson correlation analysis and simple linear regression. To assess the predictive performance of the STOP-Bang questionnaire, multiple 2×2 contingency tables were used to calculate sensitivity, specificity, positive predictive values (PPV), and negative predictive values (NPV). The response was dichotomized using AHI > 5 , AHI > 15 , and AHI > 30 as the cutoff value for defining all OSA, moderate-severe OSA, and severe OSA. Predictive probability (post-test probability) was calculated by logistic regression.

RESULTS

A total of 908 patients gave consent to this study. Portable polysomnography was completed by 650 patients; 516 patients with complete data from the STOP-Bang questionnaire were included in the analysis (**Figure 1**). The mean age, BMI, and neck circumference were 60 ± 12 years, 31 ± 7 kg/m², and 39 ± 4 cm, respectively (**Table 1**). Forty-six percent were men. Seventy percent of patients underwent orthopedic and general surgery. Of 516 patients, 346 (67.1%) had AHI > 5 , 197 (38.2%) AHI > 15 , and 86 (16.7%) AHI > 30 . As such, 170 (32.9%) patients were classified as non-OSA, 149 (38.9%) as mild OSA, 111 (16.7%) as moderate OSA, and 86 (16.7%) as severe OSA (**Table 2**).

The predictive performance of the standard STOP-Bang scoring and the different combinations of 3 items of STOP-Bang are summarized in **Table 3**. The standard STOP-Bang scoring was to combine any 3 positive items from 8 questions, where baseline sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) was 87.3% (95% CI: 81.8-91.6), 30.7% (95% CI: 25.7-36.1), 43.8% (38.8-48.8), and 79.7% (71.5-86.4), respectively to detect moderate-severe

Table 1—Demographic data of patients.

	Total
N	516
Gender F/M, N (%)	281/235 (54/46)
Age, y	59.5 ± 12
BMI, kg/m ²	30.6 ± 7
Neck circumference, cm	39.0 ± 4
ASA physical status, N (%)	
I	12 (2.3)
II	243 (47.1)
III	247 (47.9)
IV	14 (2.7)
Existing condition, N (%)	
Hypertension	257 (49.8)
Smoker	109 (21.1)
Diabetes Mellitus	96 (18.6)
Asthma	64 (12.4)
CAD	38 (7.5)
COPD	20 (3.9)
Type of Surgery, N (%)	
Orthopedic	264 (51.2)
General	107 (20.7)
Spine	44 (8.5)
Gynecology	29 (5.6)
Urology	28 (5.4)
ENT	15 (2.9)
Ophthalmology	9 (1.7)
Other	20 (3.9)

Data is depicted as mean ± SD or number (percentage). CAD, coronary arterial disease; COPD, chronic obstructive pulmonary disease; ENT, ear nose and throat.

OSA. By using a combination of any 2 positive items from the 4 STOP questions + BMI, specificity was significantly improved and reached 85% (95% CI: 80.6-88.7); PPV was improved to 46.1% (35.4-57.0); however, sensitivity and NPV were decreased. Similarly, specificity was significantly increased by combining any 2 positive items from STOP + neck circumference > 40 cm (79.0% [95% CI: 74.1-83.3]) or male gender (76.8% [95% CI: 71.8-81.3]), respectively at the expense of sensitivity and negative predictive value (NPV). However, the combination of any 2 positive items from the 4 STOP questions + age > 50 only mildly improved the specificity to detect moderate-severe OSA from 30.7% (95% CI: 25.7-36.1) to 56.1% (95% CI: 50.5-61.6). A similar trend of improved specificity to detect mild-to-severe OSA (AHI ≥ 5) and severe OSA (AHI ≥ 30) was also observed for these combinations (**Table 3**).

The predictive performance of a STOP-Bang score 4 as the cutoff and the special combinations of 4 items of STOP-Bang score are summarized in **Table 4**. A STOP-Bang score ≥ 4 had increased baseline specificity of 55.2% (95% CI: 49.5-60.7) and baseline PPV of 48.4% (42.4-54.4), decreased sensitivity of 68.0% (61.0-74.5), and NPV of 73.6% (67.6-79.1) to detect moderate-severe OSA. The specificity of the combination of any 4 items of STOP-Bang was significantly improved by using a specific combination of any 2 positive items from the 4

Table 2—Apnea-hypopnea index and responses to STOP-Bang Questionnaire.

	Total
AHI	10.4 (3.5-23.0)
AHI > 5	346 (67.1%)
AHI > 15	197 (38.2%)
AHI > 30	86 (16.7%)
Severity of OSA, N(%)	
No OSA	170 (32.9)
Mild OSA	149 (28.9)
Moderate OSA	111 (21.5)
Severe OSA	86 (16.7)
Positive response of STOP-Bang, N (%)	
Q1 (Snoring)	272 (52.7)
Q2 (Tired or sleepy)	319 (61.8)
Q3 (Observed apnea)	254 (49.2)
Q4 (Pressure)	142 (27.5)
Score BMI (B)	118 (22.9)
Score Age (A)	395 (76.6)
Score Neck (N)	181 (35.1)
Score Gender (G)	235 (45.5)

Data is depicted as median (interquartile range) or number (percentage). AHI, apnea-hypopnea index.

STOP questions + 2 specific positive items of Bang respectively. This resulted in 6 different combinations which significantly improved specificity 82.8% to 96.9% and PPV 47.8% to 54.5% at the expense of decreased sensitivity. A similar trend of specificity and PPV to detect all OSA (AHI ≥ 5) and severe OSA (AHI ≥ 30) was also observed for these distinct combinations (**Table 4**).

The post-test probabilities of moderate-severe OSA (AHI > 15) and severe OSA (AHI > 30) in patients with a STOP-Bang score ≥ 3 and the particular combinations of STOP-Bang with the same scores ≥ 3 are shown in **Figure 2**. Compared to the patients with any positive 3 items from the STOP-Bang questionnaire, the patients with a combination of STOP score ≥ 2 + male had 18% increase in the predicted probability of moderate-severe OSA (0.52 [CI:0.43-0.60] vs. 0.44 [CI:0.39-0.49]; **Figure 2A**). The patients with a combination of STOP score ≥ 2 + BMI > 35 kg/m² had 42% increase in the predicted probability of severe OSA (0.29 [0.20-0.40] vs. 0.21 [0.17-0.25]; **Figure 2B**).

Similarly, compared to the patients with any 4 positive items of STOP-Bang questionnaire, the combination of STOP score ≥ 2 + male + BMI > 35 kg/m² had 15% increase in the predicted probability of moderate-severe OSA (0.55 [0.32-0.76] vs. 0.48 [0.42-0.54]; **Figure 3A**). Meanwhile, the patients with the same combination had 86% increase in predicted probability of severe OSA (0.46 [0.24-0.68] vs. 0.25[0.20-0.30]; **Figure 3B**).

DISCUSSION

The results from this study further confirmed that a STOP-Bang score ≥ 3 offers high sensitivity to detect sleep apnea and is an excellent cutoff value for screening of OSA patients. The specificity of the STOP-Bang score could be improved by

Table 3—Predictive performance of combination of two items from STOP and one item from BANG for identifying patients with moderate-severe OSA.

All OSA, AHI > 5 as cutoff					
Cutoffs	N (%)	Sensitivity, % (95%CI)	Specificity, % (95%CI)	PPV, % (95%CI)	NPV, % (95%CI)
STOP-Bang \geq 3	290 (56.2)	83.8 (79.5-87.5)	39.4 (32.0-47.2)	73.8 (69.2-78.1)	54.5 (45.3-63.5)
STOP \geq 2 + Bang \geq 1	233 (45.2)	67.3 (62.1-72.3)	52.9 (45.4-60.4)	74.4 (69.2-79.2)	44.3 (37.4-51.5)
STOP \geq 2 + BMI > 35 kg/m ²	69 (13.4)	19.9 (15.9-24.6)	88.2 (82.4-92.7)	77.5 (67.5-85.7)	35.1 (30.6-39.9)
STOP \geq 2 + Neck > 40 cm	101 (19.6)	29.2 (24.5-34.3)	80.0 (73.2-85.7)	74.8 (66.6-81.9)	35.7 (30.9-40.7)
STOP \geq 2 + Male	117 (22.7)	33.8 (28.8-39.1)	78.8 (72.7-85.0)	76.5 (68.9-82.9)	36.9 (31.9-42.1)
STOP \geq 2 + Age > 50 y	193 (37.4)	55.8 (50.4-61.1)	62.4 (54.6-69.7)	75.1 (69.3-82.3)	40.9 (34.9-47.2)
Moderate-severe OSA, AHI > 15 as cutoff					
Cutoffs	N (%)	Sensitivity, % (95%CI)	Specificity, % (95%CI)	PPV, % (95%CI)	NPV, % (95%CI)
STOP-Bang \geq 3	172 (33.3)	87.3 (81.8-91.6)	30.7 (25.7-36.1)	43.8 (38.8-48.8)	79.7 (71.5-86.4)
STOP \geq 2 + Bang \geq 1	141 (27.3)	71.6 (64.7-77.8)	46.1 (40.5-51.7)	45.0 (39.5-50.7)	72.4 (65.7-78.4)
STOP \geq 2 + BMI > 35 kg/m ²	41 (7.9)	20.8 (15.4-27.2)	85.0 (80.6-88.7)	46.1 (35.4-57.0)	63.5 (58.7-68.0)
STOP \geq 2 + Neck > 40 cm	66 (12.8)	33.5 (27.0-40.6)	79.0 (74.1-83.3)	49.6 (40.8-58.4)	65.8 (60.8-70.5)
STOP \geq 2 + Male	79 (15.3)	40.1 (33.2-47.3)	76.8 (71.8-81.3)	51.6 (43.4-59.8)	67.5 (62.4-72.3)
STOP \geq 2 + Age > 50 y	117 (22.7)	59.4 (52.2-66.3)	56.1 (50.5-61.6)	45.5 (39.3-51.8)	69.1 (63.1-74.7)
Severe OSA, AHI > 30 as cutoff					
Cutoffs	N (%)	Sensitivity, % (95%CI)	Specificity, % (95%CI)	PPV, % (95%CI)	NPV, % (95%CI)
STOP-Bang \geq 3	81 (15.7)	94.2 (87.0-98.1)	27.4 (23.3-31.9)	20.6 (16.7-25.0)	95.9 (90.8-98.7)
STOP \geq 2 + Bang \geq 1	68 (13.2)	79.1 (69.0-87.1)	43.0 (38.3-47.9)	21.7 (17.3-26.7)	91.1 (86.4-94.7)
STOP \geq 2 + BMI > 35 kg/m ²	26 (5.0)	30.2 (20.8-41.1)	85.3 (81.7-88.6)	29.2 (20.1-39.8)	85.9 (82.3-89.1)
STOP \geq 2 + Neck > 40 cm	36 (7.0)	41.9 (31.3-53.0)	77.4 (73.2-81.3)	27.1 (19.7-35.5)	86.9 (83.2-90.2)
STOP \geq 2 + Male	43 (8.3)	50.0 (39.0-61.0)	74.4 (70.0-78.5)	28.1 (21.1-35.9)	88.2 (84.4-91.3)
STOP \geq 2 + Age > 50 y	58 (11.2)	67.4 (56.5-77.2)	53.7 (48.9-58.5)	22.6 (17.6-28.2)	89.2 (84.8-92.7)

Data are presented as average (95% confidence interval). AHI, apnea-hypopnea index; NPV, negative predictive value; PPV, positive predictive value.

examining the specific combinations of factors derived from the individual STOP-Bang item. The specificity was 85%, 79%, or 77% for moderate-severe OSA when patients have \geq 2 positive items of STOP + BMI > 35 kg/m², neck circumference > 40 cm, or male gender, respectively. Compared with the combination of any 3 items of STOP-Bang questionnaire, the combination of a STOP score \geq 2 + BMI > 35 kg/m² had a 42% increase in PPV (predicted probability) for identifying patients with severe OSA. Similarly, the definite combination of a STOP score \geq 2 + male + BMI > 35 kg/m² had an 86% increase in PPV (predicted probability) for identifying patients with severe OSA versus any 4 positive items of STOP-Bang questionnaire.

In order to keep the questionnaire ease of use, all items in the STOP-Bang questionnaire are treated equally. However, not all items have an equal predictive weight for OSA.^{2,12} Based on the data of patients' referral to the sleep clinic, Farney et al. evaluated the STOP-Bang model and indicated that a weighted model for each response can slightly improve prediction over the linear model.⁹ Overweight and obesity remain the most important modifiable causes of OSA in adult population.¹ Weight gain and loss have been consistently associated with increasing and decreasing OSA severity, respectively, in both observational and intervention studies.^{1,15,16}

Gender also influences numerous parameters associated with OSA.¹⁷⁻¹⁹ OSA has a male predominance. The prevalence of OSA is more than two times higher in men versus premenopausal women despite similar age and lower weight.¹⁹ Our data demonstrate that the specific combinations of two STOP items

+ male and/or BMI > 35 kg/m² are more effective in improving the specificity and predictive probability of STOP-Bang questionnaire. The finding that BMI > 35 kg/m² was more predictive than neck circumference seems counterintuitive, but it has been shown that the correlation between BMI and AHI was stronger than the correlation between neck circumference and AHI.²⁰

An ideal diagnostic tool should have high sensitivity and specificity at the same cutoff value. However, this is a very rare situation. For most diagnostic tools, there is a tradeoff between sensitivity and specificity. As the cutoff value shifts to increase specificity, sensitivity decreases. The STOP-Bang questionnaire follows a similar trend. When the cutoff value of STOP-Bang score was increased from 3 to 7, the specificity increased from 27.6% to 95.8%, and the sensitivity to detect severe OSA decreased from 94.8% to 11.9%.¹⁰ Essentially, a STOP-Bang score of less than 3 with high sensitivity is useful to "rule out" OSA if the patient scores negative. A STOP-Bang score of 5-8 with high specificity is useful to "rule in" OSA.

The results from this study can further help to differentiate the patients with a STOP-Bang score of 3 or 4 by examining the specific combinations of predictive factors. For the best way to use the STOP-Bang questionnaire, we proposed a two-step strategy.¹⁰ To better decipher the two-step application strategy, we employed a flow chart to show the patient number meeting the different criteria and the probability of moderate-severe OSA (AHI > 15 events/h) and severe OSA (AHI > 30 events/h) in specific patient groups (**Figure 4**). The patients with a STOP-Bang score 0-2 are at low risk of severe OSA (AHI > 30 events/h,

Table 4—Predictive performance of combination of two items from STOP and two items from BANG for identifying patients with moderate-severe OSA.

All OSA, AHI > 5 as cutoff					
Cutoffs	N (%)	Sensitivity, % (95%CI)	Specificity, % (95%CI)	PPV, % (95%CI)	NPV, % (95%CI)
STOP-Bang \geq 4	208 (40.3)	60.1 (54.7-65.3)	58.8 (51.0-66.3)	74.8 (69.3-79.8)	42.0 (35.7-48.6)
STOP \geq 2 + Bang \geq 2	163 (31.6)	47.1 (41.8-52.5)	72.4 (65.0-78.9)	77.6 (71.4-83.1)	40.2 (34.7-45.9)
STOP \geq 2 + BMI > 35 kg/m ² + Male	17 (3.3)	4.9 (2.9-7.8)	97.1 (93.3-99.0)	77.3 (54.6-92.2)	33.4 (29.3-37.8)
STOP \geq 2 + BMI > 35 kg/m ² + Neck > 40 cm	34 (9.8)	9.8 (6.9-13.5)	92.4 (87.3-95.9)	72.3 (57.4-84.4)	33.5 (29.2-38.0)
STOP \geq 2 + BMI > 35 kg/m ² + Age > 50 y	53 (15.3)	15.3 (11.7-19.6)	91.8 (86.6-95.4)	79.1 (67.4-88.1)	34.7 (30.3-39.4)
STOP \geq 2 + Neck > 40 cm + Male	74 (14.3)	21.4 (17.2-26.1)	86.5 (80.4-91.2)	76.3 (66.6-84.3)	35.1 (30.5-39.9)
STOP \geq 2 + Age > 50 + Neck > 40 cm	73 (14.1)	21.1 (16.9-25.8)	85.1 (79.7-90.7)	75.3 (65.5-83.5)	34.8 (30.3-39.6)
STOP \geq 2 + Age > 50 + Male	92 (17.8)	26.6 (22.0-31.6)	85.3 (79.1-92.3)	78.6 (70.1-85.7)	36.3 (31.6-41.3)
Moderate-severe OSA, AHI > 15 as cutoff					
Cutoffs	N (%)	Sensitivity, % (95%CI)	Specificity, % (95%CI)	PPV, % (95%CI)	NPV, % (95%CI)
STOP-Bang \geq 4	134 (26.0)	68.0 (61.0-74.5)	55.2 (49.5-60.7)	48.4 (42.4-54.4)	73.6 (67.6-79.1)
STOP \geq 2 + Bang \geq 2	106 (20.5)	53.8 (46.6-60.9)	67.7 (62.3-72.8)	50.7 (43.3-57.7)	70.4 (64.9-75.5)
STOP \geq 2 + BMI > 35 kg/m ² + Male	12 (2.3)	6.1 (3.2-10.4)	96.9 (94.3-98.5)	54.5 (32.2-75.6)	62.6 (58.1-66.1)
STOP \geq 2 + BMI > 35 kg/m ² + Neck > 40 cm	24 (4.7)	12.2 (8.0-17.6)	92.8 (89.4-95.4)	51.1 (36.1-65.9)	63.1 (58.6-67.5)
STOP \geq 2 + BMI > 35 kg/m ² + Age > 50 y	32 (6.2)	16.2 (11.4-22.2)	89.0 (85.1-92.2)	47.8 (35.4-60.3)	63.3 (58.6-67.7)
STOP \geq 2 + Neck > 40 cm + Male	49 (9.5)	24.9 (19.0-31.5)	85.6 (81.2-89.3)	51.6 (41.1-62.0)	64.8 (60.1-69.4)
STOP \geq 2 + Age > 50 + Neck > 40 cm	47 (9.1)	23.9 (18.1-30.4)	84.6 (80.2-88.4)	49.0 (38.6-59.4)	64.3 (59.5-68.5)
STOP \geq 2 + Age > 50 + Male	62 (12.0)	31.5 (25.0-38.0)	82.8 (78.2-86.7)	53.0 (43.6-62.3)	66.2 (61.3-70.8)
Severe OSA, AHI > 30 as cutoff					
Cutoffs	N (%)	Sensitivity, % (95%CI)	Specificity, % (95%CI)	PPV, % (95%CI)	NPV, % (95%CI)
STOP-Bang \geq 4	68 (13.2)	79.1 (69.0-87.1)	51.4 (46.6-56.2)	24.5 (19.6-30.1)	92.5 (88.4-95.5)
STOP \geq 2 + Bang \geq 2	59 (11.4)	68.6 (57.7-78.2)	65.1 (60.4-69.6)	28.2 (22.2-34.9)	91.2 (87.5-94.1)
STOP \geq 2 + BMI > 35 kg/m ² + Male	10 (1.9)	11.6 (5.7-20.4)	97.2 (95.2-98.6)	45.5 (24.4-67.8)	84.6 (81.1-87.7)
STOP \geq 2 + BMI > 35 kg/m ² + Neck > 40 cm	17 (3.3)	19.8 (12.0-29.8)	93.0 (90.2-95.2)	36.2 (22.7-51.5)	85.3 (81.8-88.4)
STOP \geq 2 + BMI > 35 kg/m ² + Age > 50 y	23 (4.5)	26.7 (17.8-37.4)	89.8 (86.5-92.5)	34.3 (23.2-46.9)	86.0 (82.4-89.1)
STOP \geq 2 + Neck > 40 cm + Male	26 (5.0)	30.2 (20.8-41.1)	84.0 (80.1-87.3)	27.4 (18.7-37.5)	85.7 (82.0-89.0)
STOP \geq 2 + Age > 50 + Neck > 40 cm	27 (5.2)	31.4 (21.8-42.3)	84.0 (80.1-87.3)	28.1 (19.4-38.2)	86.0 (82.3-89.1)
STOP \geq 2 + Age > 50 + Male	35 (6.8)	40.7 (30.2-51.8)	80.9 (76.9-84.5)	29.9 (21.8-39.1)	87.2 (83.5-90.3)

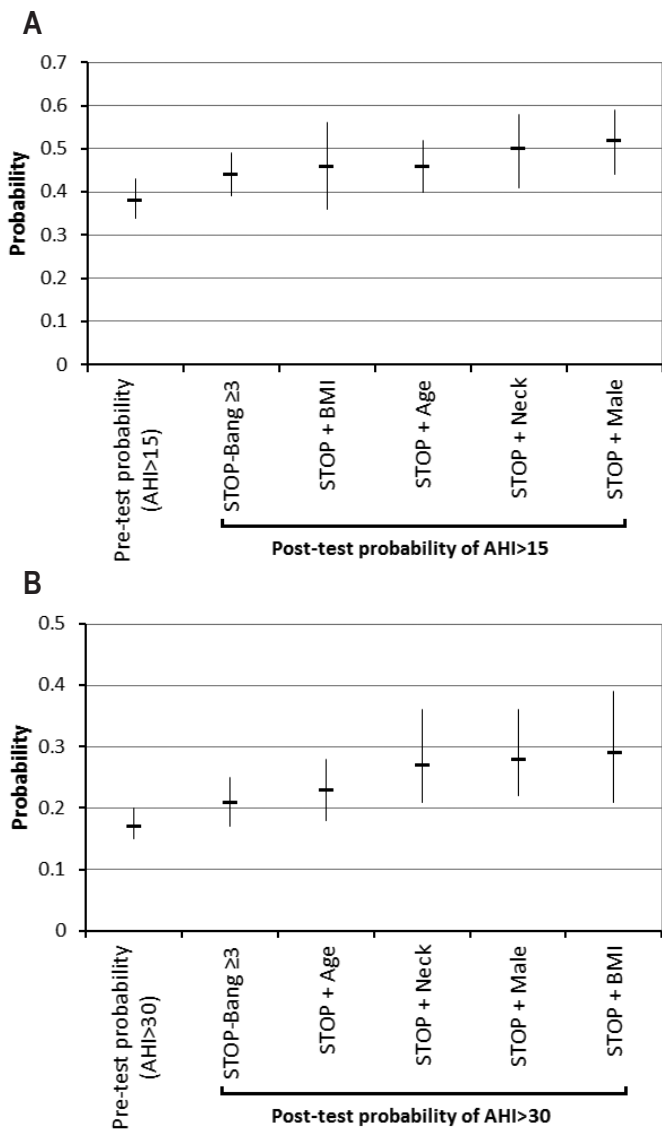
Data are presented as average (95% confidence interval). AHI, apnea-hypopnea index; NPV, negative predictive value; PPV, positive predictive value.

probability 4.1%). The patients with a STOP-Bang score 5 or greater are at high risk of severe OSA (probability 30.1%),¹⁰ and the patients with a STOP-Bang score 3-4 are at intermediate risk of severe OSA (probability 13.8%). By examining the specific item combinations of STOP-Bang questionnaire within the group of patients with a STOP-Bang score 3 or 4, we could further identify the patients with a higher risk of OSA from patients with an intermediate risk of OSA. Compared to the patients with an indiscriminate STOP-Bang score of 3 and 4, i.e., any 3 or 4 items positive, the probability of severe OSA in the patients with a STOP score 2 or greater + male gender and/or BMI > 35 kg/m² was increased by 64% (from 12.2% to 20%; **Figure 4**).

Since an elevated serum HCO₃ level also increased the specificity of STOP-Bang questionnaire,¹¹ the patients with an integration of a STOP-Bang score \geq 3 + an elevated serum HCO₃ level should be treated similarly to the patients with STOP-Bang score \geq 5. Although the incorporation of individual item weight into the score model detracts from the simplicity and ease of use,⁹ considering the predictive value of each component could be helpful in a two-step strategy for risk stratification.

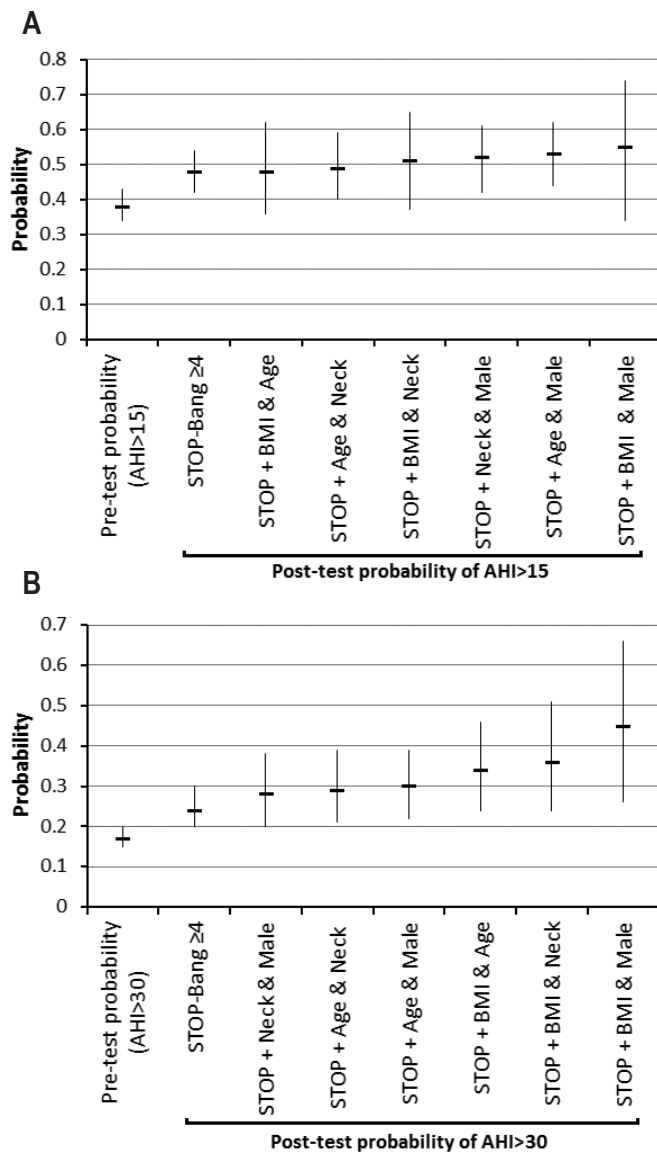
Although polysomnography remains the gold standard for the diagnosis of OSA, it is costly and time consuming.¹³ Given the diagnostic challenges associated with OSA, it remains essential for clinicians to properly identify patients at greater risk for major invasive procedures. At present, there is limited scientific literature on the preoperative preparation of patients with suspected OSA. The clinical decision rules on how to prepare these patients are mostly consensus based.^{21,22} The two-step application strategy of STOP-Bang questionnaire may help the perioperative care team to further stratify OSA risk in surgical patients. The appraisal of the STOP-Bang score, history of daytime somnolence, coexisting medical conditions, occupational hazards, invasiveness of surgery, and postoperative opioids requirement will help the perioperative care team determine whether patients should be referred to sleep medicine. Regardless of the risk, perioperative precautions should be practiced. This includes preparation for difficult intubation, use of short-acting anesthetic agents, adequate neuromuscular blockade reversal, and CPAP treatment if needed.²¹

Figure 2



(A) Post-test probability of moderate-severe OSA (AHI > 15 events/h) and (B) severe OSA (AHI > 30 events/h) by using a STOP-Bang 3 or greater and specific combination of 3 items as the cutoff value. Data were represented as probability with the 95% confidence interval for the specific combinations of STOP ≥ 2 + one item from Bang.

Figure 3

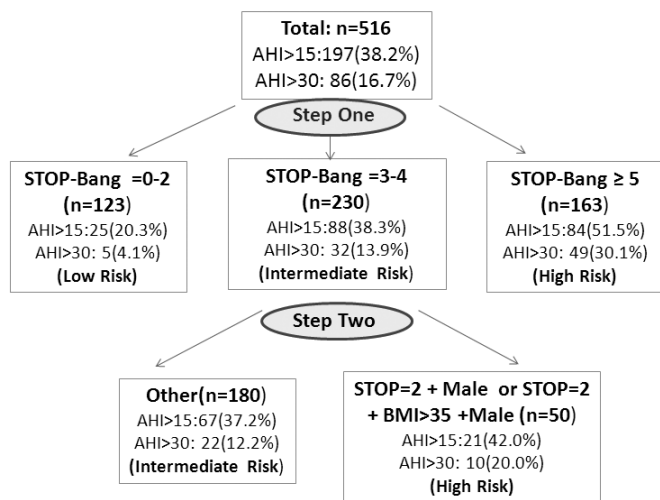


(A) Post-test probability of moderate-severe OSA (AHI > 15 events/h) and (B) severe OSA (AHI > 30 events/h) by using a STOP-Bang score 4 or greater and specific combination of 4 items as the cutoff value. Data were represented as probability with the 95% confidence interval for the specific combinations of STOP ≥ 2 + two various items from Bang.

The diagnostic challenges with OSA are not unique and are analogous to those encountered for diabetes, where stepwise screening has also emerged as part of a diagnostic strategy.^{23,24} Similar to OSA, diabetes is a common condition where a large proportion of patients remain undiagnosed, and up to 25% of new patients present with advanced comorbid disease. Thus effective screening strategies are required.^{23,24} Although the oral glucose tolerance test remains the diagnostic gold standard for diabetes, it poses as a major diagnostic barrier similar to polysomnography for the diagnosis of OSA.^{23,24} Thus randomly performed screening blood tests, such as hemoglobin A1c (HbA1c), have emerged as an acceptable alternative to accelerate diagnosis. The diagnostic value of HbA1c comes from the variable sensitivity and specificity at different cutoff thresholds,

much like the different STOP-Bang scores.^{10,23,24} For example, similar to a STOP-Bang score ≥ 3, HbA1c > 6.0 remains adequately sensitive (63% to 67%) to detect diabetes, but lacks specificity to provide any diagnostic value.²³ In contrast, when HbA1c > 7.0, although poorly sensitive (25.3%), it becomes highly specific (99.9%), much like a STOP-Bang score ≥ 5.^{10,23} As such, HbA1c has become part of the diagnostic criterion for diabetes but still requires a second test to confirm the diagnosis.^{23,24} Thus, screening for sleep disordered breathing might benefit from a strategy similar to diabetic screening, where patients are deemed to be at particularly high risk for moderate-severe OSA (STOP-Bang score ≥ 5, or a distinct amalgamation of STOP-Bang score 3 or 4), and have a second positive criterion/study, to confirm diagnosis.

Figure 4—Flow chart for two-step OSA screening strategy with STOP-Bang questionnaire.



Patient number, the probability of moderate-severe OSA (AHI > 15 events/h) and severe OSA (AHI > 30 events/h) were presented in each group.

The proper identification of OSA has important perioperative implications for surgical patients.^{6,7,25,26} High STOP-Bang scores have also been correlated with a greater increase in perioperative morbidity.²⁷ CPAP prevents postoperative exacerbation of OSA, confers survival benefits in general, improves comorbid diseases, and leads to significant reductions in medication in the months following surgery.^{7,8} Thus the proper identification of patients with OSA not only helps to improve perioperative outcome, but also help patients to initiate treatment of OSA, which may benefit them for long term, well beyond their surgical course.^{21,28}

There are notable limitations to our study. There was probably self-selection bias introduced by patients themselves during the recruitment process. Surgical patients with OSA related symptoms might be more likely to give consent to the sleep studies and artificially elevate the prevalence of OSA.² As a result, the pretest prevalence of moderate-severe OSA (38%) in this study population was higher than that in general population (10%),¹ which may lead to a higher probability of moderate-severe OSA (20%) in patients classified as low-risk for OSA by the STOP-Bang questionnaire (Figure 4). As in most sleep apnea screening studies, central apneas were not evaluated separately. The study data were collected in surgical patients, and the results may not be applicable to other populations; further validation is needed.

In summary, undiagnosed OSA remains an important issue for anesthesiologists, and these data further complete the 2-step OSA screening strategy with STOP-Bang questionnaire. The patients with a STOP-Bang score 0-2 are unlikely to have moderate-severe OSA and can be excluded from further investigation. The patients with a STOP-Bang score 5 or greater should be treated with a high index of suspicion for having undiagnosed OSA. For the patients with a STOP-Bang score of 3-4, a second step can be performed by examining the specific combination of factors (e.g., STOP score 2 or greater + BMI > 35 kg/m² and/or + male) or serum HCO₃ level to identify those at

high risk for moderate-severe OSA. This approach may help to exclude low risk patients, but will also help to identify those with a high likelihood of OSA. This would allow the health care team to facilitate the efficient allocation of healthcare resources such as sleep center referrals, as well as devise the optimal perioperative plan for patients.

CONCLUSION

Improvements in the specificity of STOP-Bang questionnaire can be made by considering the specific constellation of risk factors. For patients with STOP score 2 or greater, male and BMI are the more important predictive factors than age and neck circumference to improve the specificity and post-test probability of OSA identification.

REFERENCES

1. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol* 2013;177:1006-14.
2. Chung F, Yegneswaran B, Liao P, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008;108:812-21.
3. Finkel KJ, Searleman AC, Tymkew H, et al. Prevalence of undiagnosed obstructive sleep apnea among adult surgical patients in an academic medical center. *Sleep Med* 2009;10:753-8.
4. Young T, Evans L, Finn L, Palta M. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep* 1997;20:705-6.
5. Singh M, Liao P, Kobah S, Wijeyesundera DN, Shapiro C, Chung F. Proportion of surgical patients with undiagnosed obstructive sleep apnoea. *Br J Anaesth* 2013;110:629-36.
6. Kaw R, Chung F, Pasupuleti V, Mehta J, Gay PC, Hernandez AV. Meta-analysis of the association between obstructive sleep apnoea and postoperative outcome. *Br J Anaesth* 2012;109:897-906.
7. Mehta V, Subramanyam R, Shapiro CM, Chung F. Health effects of identifying patients with undiagnosed obstructive sleep apnea in the preoperative clinic: a follow-up study. *Can J Anaesth* 2012;59:544-55.
8. Liao P, Luo Q, Elsaid H, Kang W, Shapiro C, Chung F. Perioperative auto-titrated continuous positive airway pressure treatment in surgical patients with obstructive sleep apnea: a randomized controlled trial. *Anesthesiology* 2013;119:837-47.
9. 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;7:459-65.
10. Chung F, Subramanyam R, Liao P, Sasaki E, Shapiro C, Sun Y. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *Br J Anaesth* 2012;108:768-75.
11. Chung F, Chau E, Yang Y, Liao P, Hall R, Mokhlesi B. Serum bicarbonate level improves specificity of STOP-Bang screening for obstructive sleep apnea. *Chest* 2013;143:1284-93.
12. Ramachandran SK, Josephs LA. A meta-analysis of clinical screening tests for obstructive sleep apnea. *Anesthesiology* 2009;10:928-39.
13. Chung F, Liao P, Sun Y, Amirshahi B, Fazel H, Shapiro CM, Elsaid H. Perioperative practical experiences in using a level 2 portable polysomnography. *Sleep Breath* 2010;15:367-75.
14. Iber C, Ancoli-Israel S, Cheeson A, Quan SF. The AASM Manual for the Scoring of Sleep and Associated Events, Rules, Terminology and Technical Specifications. Westchester, IL: American Academy of Sleep Medicine, 2007.
15. Newman AB, Foster G, Givelber R, Nieto FJ, Redline S, Young T. Progression and regression of sleep-disordered breathing with changes in weight: the Sleep Heart Health Study. *Arch Intern Med* 2005;165:2408-13.
16. Young T, Peppard PE, Taheri S. Excess weight and sleep-disordered breathing. *J Appl Physiol* 2005;99:1592-9.
17. Dancy DR, Hanly PJ, Soong C, Lee B, Shepard J Jr., Hoffstein V. Gender differences in sleep apnea: the role of neck circumference. *Chest* 2003;123:1544-50.

18. Gabbay IE, Lavie P. Age- and gender-related characteristics of obstructive sleep apnea. *Sleep Breath* 2012;16:453-60.
19. Ralls FM, Grigg-Damberger M. Roles of gender, age, race/ethnicity, and residential socioeconomic in obstructive sleep apnea syndromes. *Curr Opin Pulm Med* 2012;18:568-73.
20. Plywaczewski R, Bielen P, Bednarek M, Jonczak L, Gorecka D, Sliwinski P. Influence of neck circumference and body mass index on obstructive sleep apnoea severity in males. *Pneumonol Alergol Pol* 2008;76:313-20.
21. Seet E, Chung F. Management of sleep apnea in adults - functional algorithms for the perioperative period: continuing professional development. *Can J Anaesth* 2010;57:849-64.
22. Gross JB, Apfelbaum JL, Caplan RA, et al. Practice guidelines for the perioperative management of patients with obstructive sleep apnea: an updated report by the American Society of Anesthesiologists Task Force on Perioperative Management of patients with obstructive sleep apnea. *Anesthesiology* 2014;120:268-86.
23. Saudek CD, Herman WH, Sacks DB, Bergenstal RM, Edelman D, Davidson MB. A new look at screening and diagnosing diabetes mellitus. *J Clin Endocrinol Metab* 2008;93:2447-53.
24. Pottie K, Jaramillo A, Lewin G, et al. Recommendations on screening for type 2 diabetes in adults. *CMAJ* 2012;184:1687-96.
25. Memtsoudis SG, Stundner O, Rasul R, et al. Sleep apnea and total joint arthroplasty under various types of anesthesia: a population-based study of perioperative outcomes. *Reg Anesth Pain Med* 2013;38:274-81.
26. Mokhlesi B, Hovda MD, Vekhter B, Arora VM, Chung F, Meltzer DO. Sleep-disordered breathing and postoperative outcomes after bariatric surgery: analysis of the nationwide inpatient sample. *Obes Surg* 2013;23:1842-51.
27. Vasu TS, Doghranji 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.
28. Swart P, Chung F, Fleetham J. An order-based approach to facilitate postoperative decision-making for patients with sleep apnea. *Can J Anaesth* 2013;60:321-4.

ACKNOWLEDGMENTS

Author contributions: Dr. Chung was responsible in designing and conducting the study, writing the manuscript and obtaining the funding for the trial. Dr. Yiliang Yang helped design the study, analyzed the data and write the manuscript. Dr. Russell Brown helped design the study and write the manuscript. Dr. Pu Liao helped design the study and write the manuscript.

The authors thank Hisham Elsaid, M.D. (Research Fellow, Department of Anesthesia, University Health Network, Toronto, Ontario, Canada); Babak Amirshahi, M.D. (Research Fellow, Department of Anesthesia, University Health Network, Toronto, Ontario, Canada); Hoda Fazel, M.D. (Research Fellow, Department of Anesthesia, University Health Network, Toronto, Ontario, Canada); Sazzadul Islam, M.Sc. (Research coordinator, Department of Anesthesia, University Health Network, Toronto, Ontario, Canada), and Santhira Vairavanathan, M.B.B.S. (Research coordinator, Department of Anesthesia, University Health Network, Toronto, Ontario, Canada) for collecting data and Yuming Sun, M.D. (Registered Polysomnographic Technologist, Sleep Research Unit & Department of Anesthesia, University Health Network, Toronto, Ontario, Canada) for scoring polysomnography recordings and Colin Shapiro, M.D. (Professor, Dept of Psychiatry, University of Toronto, Director, Sleep Research Unit, Toronto Western Hospital) for supervision of the scoring.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication September, 2013

Submitted in final revised form March, 2014

Accepted for publication April, 2014

Address correspondence to: Dr. Frances Chung, Room 405, 2McL wing, Department of Anesthesia, Toronto Western Hospital, University Health Network, 399 Bathurst Street, Toronto, Ontario, Canada M5T 2S8; Tel: (416) 603-5118; Fax: (416) 603-6494; Email: Frances.chung@uhn.ca

DISCLOSURE STATEMENT

This was not an industry supported study. Funding is from Physician Services Incorporated Foundation, University Health Network Foundation, and Department of Anesthesia, University Health Network-Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada. The department/institution to which the work is attributed is the Department of Anesthesiology and Pain Management, Toronto Western Hospital, University Health Network, University of Toronto. The authors have indicated no financial conflicts of interest.

APPENDIX

STOP-Bang Questionnaire

1. **Snoring:** Do you snore loudly (loud enough to be heard through closed doors)?
Yes No
2. **Tired:** Do you often feel tired, fatigued, or sleepy during daytime?
Yes No
3. **Observed:** Has anyone observed you stop breathing during your sleep?
Yes No
4. **Blood Pressure:** Do you have or are you being treated for high blood pressure?
Yes No
5. **BMI:** BMI more than 35 kg/m²?
Yes No
6. **Age:** Age over 50 years old?
Yes No
7. **Neck circumference:** Neck circumference greater than 40 cm?
Yes No
8. **Gender:** Male?
Yes No

High risk of OSA: Yes to 3 or more questions

Low risk of OSA: Yes to less than 3 questions

Chung F et al. *Anesthesiology* 2008;108:812-21.