

Validation of the Danish STOP-Bang obstructive sleep apnoea questionnaire in a public sleep clinic

Asbjørn Kørvel-Hanquist, Ida Gillberg Andersen, Sophie Elisabeth Krogh Lauritzen, Susanne Dahlgaard & Janko Moritz

ABSTRACT

INTRODUCTION: Obstructive sleep apnoea is common; a prevalence of 1-5% was previously reported. However, only few cases are diagnosed and receive treatment. The aim of this study was to validate the Danish translated version of the STOP-Bang screening tool for obstructive sleep apnoea (OSA) in a public sleep clinic.

METHODS: A study population of 208 patients who were referred to a public sleep clinic on suspicion of OSA were assessed with the STOP-Bang questionnaire and at-home cardiorespiratory monitoring in order to assess the quality of the questionnaire as an OSA screening tool.

RESULTS: In the study population, 73% were males, and 51% of the population had an Apnoea-Hypopnoea Index (AHI) ≥ 15 . The STOP-Bang screening tool had a sensitivity of 0.98 for detection of OSA with AHI ≥ 15 and a corresponding specificity of 0.09. Hence, the questionnaire is able to detect almost all patients suffering from OSA. However, using the tool will cause many healthy subjects to be falsely classified as having OSA.

CONCLUSIONS: The Danish version of the STOP-Bang screening tool does not seem useful for OSA screening of patients in a sleep clinic setup, but it may be useful in primary care.

FUNDING: The Zealand Research Foundation.

TRIAL REGISTRATION: not relevant.

Obstructive sleep apnoea (OSA) is characterised by multiple repetitive episodes of disturbed breathing and oxygen desaturation during sleep, and is the most prevalent sleeping disorder among adults. Although debated, a disease prevalence of 1-5% seems plausible in adults in developed countries [1]. Existing studies have associated OSA with a variety of diseases such as ischaemic heart disease, fatty liver, renal disease and psychiatric disorders [2-5]. OSA also seems to play a part in the metabolic syndrome, although the causal pathways between OSA and these comorbidities is not fully understood [6]. In addition, OSA is associated with an increased risk of traffic accidents and work-related injuries [7]. Altogether, OSA constitutes a considerable health risk, and efforts have been made to find and treat people with undiagnosed OSA in order to reduce future disability and to improve their quality of life.

Polysomnography (PSG), performed in a sleep labo-

ratory, is considered the gold standard for diagnosing OSA. However, this diagnostic test is expensive and time-consuming compared with the existing at-home portable cardiorespiratory monitoring (CRM) devices that have been developed in recent years. The NOX T3 is a device for at-home CRM. It measures respiratory effort, nasal airflow and also pulse and oxygen saturation during sleep. At-home CRM has been validated and used in multiple studies, and there is a strong correlation between the Apnoea-Hypopnoea Index (AHI) as determined by PSG and AHI as determined by the NOX T3 [8].

Different screening tools have been assessed with a view to detecting individuals at high risk of OSA. One such tool is the Canadian STOP-Bang questionnaire, which is designed to screen for OSA symptoms [9]. Compared with other screening tools such as the Berlin Questionnaire [10] it often goes undiagnosed in primary care encounters. **OBJECTIVE:** To test the Berlin Questionnaire as a means of identifying patients with sleep apnea. **DESIGN:** Survey followed by portable, unattended sleep studies in a subset of patients. **SETTING:** Five primary care sites in Cleveland, Ohio. **PATIENTS:** 744 adults (of 1008 surveyed [74%]) or the Sleep

ORIGINAL ARTICLE

Department of Otorhinolaryngology and Maxillofacial Surgery, Zealand University Hospital, Denmark

Dan Med J
2018;65(1):5434

! STOP-BANG

Sleep apnoea questionnaire

- S Do you Snore loudly
(louder than talking or loud enough to be heard through closed doors)?
- T Do you often feel Tired, fatigued, or sleepy during daytime?
- O Has anyone Observed you stop breathing during your sleep?
- P Do you have or are you being treated for high blood Pressure?
- B BMI more than 35kg/m²?
- A Age over 50 years old?
- N Neck circumference > 40cm (16 inches)?
- G Gender: Male?

	Yes	No
S	<input type="checkbox"/>	<input type="checkbox"/>
T	<input type="checkbox"/>	<input type="checkbox"/>
O	<input type="checkbox"/>	<input type="checkbox"/>
P	<input type="checkbox"/>	<input type="checkbox"/>
B	<input type="checkbox"/>	<input type="checkbox"/>
A	<input type="checkbox"/>	<input type="checkbox"/>
N	<input type="checkbox"/>	<input type="checkbox"/>
G	<input type="checkbox"/>	<input type="checkbox"/>

The STOP-Bang questionnaire [9].

TABLE 1

Descriptive characteristics of the study population divided in groups of obstructive sleep apnoea severity.

	Non-OSA: 0 ≤ AHI < 5 (N = 52)	Mild OSA: 5 ≤ AHI < 15 (N = 50)	Moderate OSA: 15 ≤ AHI < 30 (N = 35)	Severe OSA: AHI ≥ 30 (N = 71)	All (N = 208)	p-value
Male, n	30	35	29	58	152	0.054 ^a
Age, yrs, median (range)	48.72 (22.26-85.41)	54.60 (17.22-79.86)	50.69 (28.54-82.90)	55.51 (31.80-88.42)	53.37 (17.22-88.42)	0.024 ^b
Height, m, median (range)	1.76 (1.58-1.94)	1.76 (1.59-1.99)	1.80 (1.57-1.93)	1.78 (1.55-1.92)	1.78 (1.55-1.99)	0.329 ^b
Weight, kg, median (range)	85.00 (52.00-169.00)	90.00 (58.00-129.00)	95.00 (67.00-153.00)	107.00 (76.00-165.00)	95.00 (52.00-169.00)	< 0.001 ^b
BMI, kg/m ² , median (range)	27.70 (19.10-52.20)	29.40 (18.40-41.50)	29.40 (21.20-53.60)	33.60 (22.30-60.50)	30.45 (18.40-60.50)	< 0.001 ^b
STOP-Bang score, median (range)	4.00 (2.00-7.00)	5.00 (1.00-7.00)	5.00 (2.00-7.00)	6.00 (2.00-8.00)	5.00 (1.00-8.00)	< 0.001 ^b
ODI, /h, median (range)	2.20 (0.20-6.70)	8.20 (0.80-24.00)	23.00 (5.90-32.00)	46.50 (5.70-120.00)	17.00 (0.20-120.00)	-
AHI, /h, median (range)	2.00 (0.00-4.60)	10.00 (5.30-14.80)	22.90 (15.40-29.80)	47.90 (30.50-107.50)	15.90 (0.00-107.50)	-

AHI = Apnoea-Hypopnoea Index; Bang = BMI > 35 kg/m², Age > 50 yrs, Neck circumference > 40 cm, Gender: male; ODI = oxygen desaturation index; OSA = obstructive sleep apnoea; STOP = Snoring, Tiredness during daytime, Observed apnoea, and high blood Pressure.

a) Chi-square test between groups.

b) One-way ANOVA test of variance between group means.

Disorders Questionnaire [11] and treating OSA in epilepsy patients may reduce seizure frequency and improve daytime sleepiness. The SA-SDQ, a 12-item validated measure of sleep-related breathing disorders, may be a useful tool to screen epilepsy patients for OSA, although appropriate cutoff points have not been established in this population. Previously suggested SA-SDQ cutoff points for OSA in a non-epilepsy population were 32 for women and 36 for men.

PATIENTS AND METHODS: One hundred twenty-five subjects with epilepsy undergoing polysomnography completed a survey about their sleep, including the 12-item SA-SDQ scale. Receiver-operating characteristics curves were constructed to determine optimal sensitivity and specificity.

RESULTS: Sixty-nine of the 125 subjects (45%), the STOP-Bang questionnaire was found to be clinically applicable and to have a high sensitivity, although its specificity seems to be low [12]. Validation of these tools is still on-going and new tools are being developed.

The aim of the present study was to validate the Danish STOP-Bang screening tool for OSA in a public sleep clinic in a population of patients who were referred for CRM.

METHODS

The STOP-Bang sleep questionnaire was translated into Danish by Bille et al [13].

The study population consisted of patients referred to the public sleep clinic in Koege, Denmark, during a two-month period from March 2015. Patients were eligible if they had been referred for an overnight CRM on

suspicion of OSA and had not previously been diagnosed with OSA. Patients were primarily referred by private otorhinolaryngologists from the Zealand Region, Denmark. The NOX T3 device was used for at-home CRM. On the first visit to the clinic, the patients completed the STOP-Bang questionnaire. Information about weight, height and neck circumference was collected. Neck circumference was measured at the level of the laryngeal prominence. Patients were instructed in how to use the NOX T3 device at home, and they slept with the device for one night after the initial visit to the clinic in order to determine the presence of OSA. On the second visit to the clinic, patients were assessed by clinical examination including inspection of the nasal and oral cavity and fibre-optic rhinopharyngoscopy; and a trained sleep physician analysed their CRM. If the quality of the CRM was inappropriate or a technical error had occurred, the CRM was repeated. The sleep physicians were blinded to the patients' STOP-Bang test scores during the entire study period.

Statistical analyses were generated using SAS 9.4 software (SAS Institute Inc. SAS Cary, North Carolina, USA). The severity of OSA was defined as mild (AHI ≥ 5 and < 15), moderate (AHI ≥ 15 and < 30) or severe (AHI ≥ 30). The STOP-Bang score was calculated as the sum of items with a positive answer (Yes). The test was considered positive if the sum of positive answers was ≥ 3. If a question was not answered, it was registered as a negative answer (No). The chi-square test and ANOVA tests were used to identify variations in characteristics of the study population.

The ability of the STOP-Bang questionnaire to identify patients with OSA (i.e. sensitivity) and its ability to identify healthy individuals (i.e. specificity) were calculated by using 2×2 contingency tables. The questionnaire's positive predictive value (PPV) and negative predictive value (NPV) were calculated from these tables. The receiver-operating characteristic (ROC) curve was made from the STOP-Bang cut-off values and the respective sensitivity and specificity measures. Area under the ROC curve (AUC) was calculated using logistic regression modelling. The probability calculations for the various AHI cut-off values were generated by pooling the numbers of patients with positive tests for each of the AHI cut-off values (AHI ≥ 5 , ≥ 15 and ≥ 30) and for mild, moderate and severe disease (AHI ≥ 5 and < 15 , ≥ 15 and < 30 , and ≥ 30).

Trial registration: not relevant.

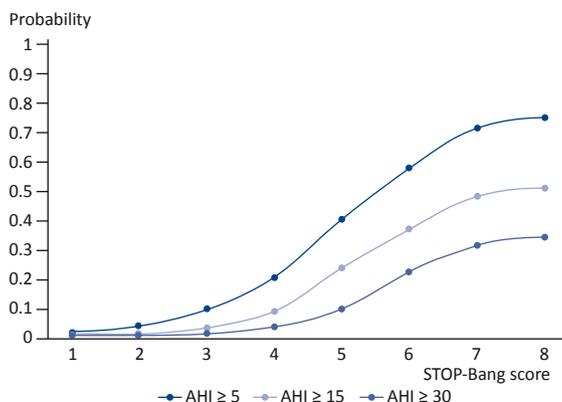
RESULTS

A population of 208 patients was studied and tested with the STOP-Bang screening tool and the CRM. In the study population, 73% were men, and the mean age was 53 years. The prevalence of OSA with AHI ≥ 15 was 51% (106 out of 208). The median AHI of the population was 15.90 ranging 0.00-107.50, and the median STOP-Bang score was 5.00 (range: 1.00-8.00). A quarter (52 of 208) of the study population had an AHI < 5 . Descriptive characteristics of the study population are displayed in **Table 1**.

When screening for OSA with AHI ≥ 15 and a STOP-Bang cut-off score ≥ 3 , the sensitivity of the questionnaire was 0.98 (95% confidence interval (CI): 0.93-1.00)

FIGURE 1

Plot of probabilities for predicting an Apnoea-Hypopnoea Index (AHI) at three different cut-off values, according to achieved STOP-Bang test score. Generated from study population frequencies.



Bang = BMI > 35 kg/m², Age > 50 yrs, Neck circumference > 40 cm, Gender: male; STOP = Snoring, Tiredness during daytime, Observed apnoea, and high blood Pressure.

TABLE 2

Test characteristics for different cut-off values in patient Apnoea-Hypopnoea Index measures and STOP-Bang cut-off value of 3. The values are mean (95% CI).

	AHI ≥ 5	AHI ≥ 15	AHI ≥ 30
Sensitivity, %	94.87 (90.15-97.76)	98.11 (93.35-99.77)	98.59 (92.40-99.96)
Specificity, %	6.52 (1.37-17.90)	8.82 (3.32-14.33)	7.30 (3.56-13.01)
PPV, %	77.49 (70.90-83.20)	52.79 (45.57-59.93)	35.53 (28.86-42.65)
NPV, %	27.27 (6.02-60.97)	81.82 (48.22-97.72)	90.91 (58.72-99.77)
Accuracy, %	74.75 (68.18-80.59)	54.33 (47.30-61.23)	38.46 (31.82-45.44)
LR+	1.01 (0.92-1.11)	1.08 (1.00-1.15)	1.06 (1.00-1.13)
LR-	0.79 (0.00-1.86) ^a	0.21 (0.00-0.56) ^a	0.19 (0.00-0.61) ^a
AUC	0.76 (0.69-0.83)	0.75 (0.69-0.82)	0.80 (0.73-0.86)

AHI = Apnoea-Hypopnoea Index; AUC = area under the receiver operating characteristic curve; Bang = BMI > 35 kg/m², Age > 50 yrs, Neck circumference > 40 cm, Gender: male; CI = confidence interval; LR = likelihood ratio; NPV = negative predictive value; PPV = positive predictive value; STOP = Snoring, Tiredness during daytime, Observed apnoea, and high blood Pressure. ^a The lower end of the CI was set at 0.00, although a negative lower end was initially calculated.

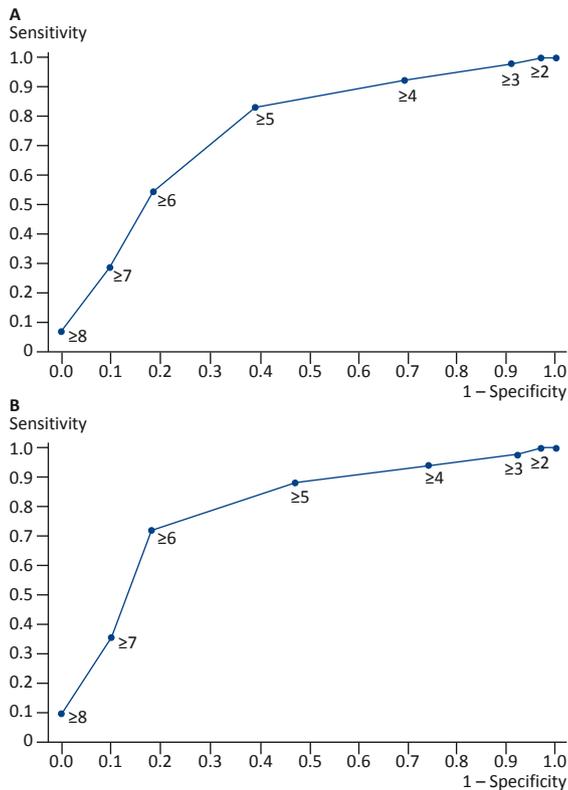
and the specificity was 0.09 (95% CI: 0.03-0.14). The accuracy of the test to detect OSA with an AHI ≥ 15 was estimated to 0.54 (95% CI: 0.47-0.61), and AUC of the ROC curve was 0.75 (95% CI: 0.69-0.82). The PPV was 0.53 (95% CI: 0.46-0.60) and the NPV was 0.82 (95% CI: 0.48-0.98). The positive likelihood ratio for OSA with AHI ≥ 15 was 1.08 and the negative likelihood ratio was 0.21. The corresponding values for OSA with AHI ≥ 5 and AHI ≥ 30 are displayed in **Table 2**. In addition, the probabilities for predicting an AHI of ≥ 5 , ≥ 15 and ≥ 30 in patients according to their achieved STOP-Bang score values are presented in **Figure 1**. The ROC curves for AHI ≥ 15 and AHI ≥ 30 , showing the sensitivity against 1-specificity at different STOP-Bang cut-off scores, are displayed in **Figure 2**. All patients with a score of 8 had severe OSA. Correspondingly, all patients with a STOP-Bang score of 1 had an AHI < 15 . The STOP-Bang scores were normally distributed around the score of 5, with a STOP-Bang score of 8 in six patients and a score of 1 in three patients. None of the patients had a STOP-Bang score of 0.

DISCUSSION

In this study, we tested the Danish STOP-Bang questionnaire with a cut-off score of ≥ 3 in patients referred to a sleep clinic. We found a high sensitivity of 0.98 for prediction of an AHI ≥ 15 , but a low specificity of only 0.09. The AUC for the ROC for AHI ≥ 15 was 0.75, corresponding to a fair diagnostic test. However, the low specificity indicates that 91% of the patients with an AHI < 15 had a false positive test result. Furthermore, the PPV of 0.53 indicates that the probability of having OSA if the test result is positive is only 53%. Hence, a positive test result is of minor relevance. The NPV was 0.82, but because of

FIGURE 2

Receiver operating characteristic curve with sensitivity and 1 – specificity according to STOP-Bang cut-off values of Apnoea-Hypopnoea Index (AHI) ≥ 15 (A) and AHI ≥ 30 (B).



Bang = BMI > 35 kg/m², Age > 50 yrs, Neck circumference > 40 cm, Gender: male; STOP = Snoring, Tiredness during daytime, Observed apnoea, and high blood Pressure.

the high rate of falsely positive screenings, the absolute number of patients who can be correctly identified as healthy is small and the probability of having OSA with a negative test result is still 18%. Taking into consideration that the CRM has no unwanted side effects for the patient, we would prefer a higher NPV. We therefore do not find the STOP-Bang questionnaire suitable for screening of OSA in a sleep clinic with high-risk patients. As is evident from Table 2, the specificity remains low considering AHI limits of ≥ 5 or ≥ 30 . The ROC curve for AHI ≥ 15 (Figure 2) indicates that a cut-off value of the STOP-Bang score of ≥ 5 increases the questionnaire's specificity; however, its sensitivity also decreases to around 80%, which we also consider unacceptable.

The estimated sensitivity of 0.98 seems high compared with results from a meta-analysis by Nagappa et al [14], who reported a pooled sensitivity of 0.94 for the STOP-Bang questionnaire to detect AHI ≥ 15 in a sleep clinic population. The specificity, estimated at 0.09 in the present study, was low compared with the pooled 0.34 specificity found in the meta-analysis. This differ-

ence may partly be explained by the relatively high median STOP-Bang score of 5.00 in the population together with a low median AHI of 15.90 in the study population. It cannot be determined whether these differences are due to differences in the characteristics of the study populations or differences between the questionnaire's properties due to the translation into Danish.

Comparing the results from the STOP-Bang questionnaire with the results from the Berlin questionnaire, which is also used as an OSA screening tool, the STOP-Bang questionnaire has a slightly higher sensitivity than the Berlin questionnaire with a sensitivity of 82% but a much lower specificity than the Berlin questionnaire with a specificity of around 37% [15]. Overall, the reliability of the mentioned questionnaires seems too low to justify their use in high-risk populations.

This study cannot conclude whether the STOP-Bang questionnaire can be used as a screening tool in a low-risk population, e.g., in primary care. Patients with symptoms of sleeping disorders would typically address a general practitioner or private ear, nose and throat specialist initially. It would be valuable to test the questionnaire in this setting in order to try to detect patients who need further referral to a specialised sleep clinic. In the primary care setting, a high sensitivity and specificity would still be required for the questionnaire to have clinical relevance.

As mentioned, the present study is limited by the fact that the study population consisted of high-risk individuals. This may have biased the estimates towards increased sensitivity and decreased specificity. Furthermore, the STOP-Bang questionnaire contains limitations. The binary score system only allows positive and negative answers to the eight questions. This seems to leave some patients with problems when answering the first four questions. In our population, 24 of 208 patients were unable to complete all the answers. However, this did not change the estimates because only one of the 24 patients would have changed group from a negative STOP-Bang test to a positive test if all the missing answers were classified as positive instead of negative. Also, the Danish questionnaire and the original questionnaire both miss information specifying which time period the questions refer to. Looking into the questions, there will logically always be a high rate of falsely positive test results as all males above 50 years only need one more point to screen positive when a cut-off value = 3 is used.

CONCLUSIONS

We tested the Danish STOP-Bang questionnaire as a screening tool for detection of OSA in adults who were referred to a sleep clinic, and found a high sensitivity but a very low specificity. The questionnaire cannot be rec-

ommended as a screening tool for OSA in a high-risk population. Nevertheless, the STOP-Bang questionnaire may prove to be useful in primary care in a low risk-population, but this warrants further investigation.

CORRESPONDENCE: *Asbjørn Kørvel-Hanquist.*

E-mail: asbp@regionsjaelland.dk

ACCEPTED: 20 November 2017

CONFLICTS OF INTEREST: none. Disclosure forms provided by the authors are available with the full text of this article at www.danmedj.dk

LITERATURE

1. Davies RJ, Stradling JR. The epidemiology of sleep apnoea. *Thorax* 1996;51(suppl 2):S65-S70.
2. Ahmed SB, Ronsley PE, Hemmelgarn BR et al. Nocturnal hypoxia and loss of kidney function. *PLoS ONE* 2011;6:e19029.
3. Gonzaga C, Bertolami A, Bertolami M et al. Obstructive sleep apnea, hypertension and cardiovascular diseases. *J Hum Hypertens* 2015;:705-12.
4. Sookoian S, Pirola CJ. Obstructive sleep apnea is associated with fatty liver and abnormal liver enzymes: a meta-analysis. *Obes Surg* 2013;23:1815-25.
5. Stubbs B, Vancampfort D, Veronese N et al. The prevalence and predictors of obstructive sleep apnea in major depressive disorder, bipolar disorder and schizophrenia: A systematic review and meta-analysis. *J Affect Disord* 2016;197:259-67.
6. Xu S, Wan Y, Xu M et al. The association between obstructive sleep apnea and metabolic syndrome: a systematic review and meta-analysis. *BMC Pulm Med* 2015;15:150.
7. Hirsch Allen AJM, Bansback N, Ayas NT. The effect of OSA on work disability and work-related injuries. *Chest* 2015;147:1422-8.
8. Cairns A, Wickwire E, Schaefer E et al. A pilot validation study for the NOX T3TM portable monitor for the detection of OSA. *Sleep Breath* 2014;18:609-14.
9. Chung F, Yegneswaran B, Liao P et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008;108:812-21.
10. 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.
11. Weatherwax KJ, Lin X, Marzec ML et al. Obstructive sleep apnea in epilepsy patients: the Sleep Apnea Scale of the Sleep Disorders Questionnaire (SA-SDQ) is a useful screening instrument for obstructive sleep apnea in a disease-specific population. *Sleep Med* 2003;4:517-21.
12. Boynton G, Vahabzadeh A, Hammoud S et al. Validation of the STOPBANG Questionnaire among patients referred for suspected obstructive sleep apnea. *J Sleep Disord Treat Care* 2013;2(4).
13. Bille J, Bille-Hasselstrøm C, Petersen CG. Translation and validation of the Stop-Bang Questionnaire for obstructive sleep apnoea into Danish. *Dan Med J* 2015;62(12):A5158.
14. Nagappa M, Liao P, Wong J et al. Validation of the STOP-Bang Questionnaire as a screening tool for obstructive sleep apnea among different populations: a systematic review and meta-analysis. *PLoS One* 2015;10:e0143697.
15. Senaratna CV, Perret JL, Matheson MC et al. Validity of the Berlin questionnaire in detecting obstructive sleep apnea: a systematic review and meta-analysis. *Sleep Med Rev* 2017;36:116-24.