BMJ Open

Predicting sleep disordered breathing in outpatients with suspected OSA

Douglas C Cowan,[1] Gwen Allardice,[2] Duncan MacFarlane,[1] Darren Ramsay,[1] Heather Ambler,[1] Stephen Banham,[1] Eric Livingston,[1] Christopher Carlin[1]

[1]North Glasgow Sleep Service, Gartnavel General Hospital and Glasgow Royal Infirmary, Glasgow, UK
[2]West of Scotland Cancer Surveillance Unit, University of Glasgow, Glasgow, UK

Correspondence to
Dr Douglas C Cowan;
douglas.cowan@otago.ac.nz

ABSTRACT

Objective: To validate the utilities of Berlin, STOP and STOP-BANG Questionnaires, other patient characteristics, comorbidities, Epworth Sleepiness Scale (ESS), fractional exhaled nitric oxide (F_ENO) and blood markers for the prediction of sleep disordered breathing (SDB) on limited polygraphy.

Setting: North Glasgow Sleep Service (a tertiary referral centre).

Participants: 129 consecutive patients, aged ≥16 years, referred to the sleep clinic for assessment of possible obstructive sleep apnoea.

Interventions: We selected cut-points of apnoea hypopnoea index (AHI) of ≥5 and ≥15/h from their home polygraphy and determined associations of these with individual symptoms, questionnaire scores and other results. Receiver operating characteristic analysis and univariate and multivariate logistic regression were used to explore these.

Primary and secondary outcomes measures:
Primary: The utility of STOP, STOP-BANG and Berlin Questionnaires for prediction of SDB. *Secondary*: The utility of other measures for prediction of SDB.

Results: AHI was ≥5 in 97 patients and ≥15 in 56 patients. STOP and STOP-BANG scores were associated with both AHI cut-points but results with ESS and Berlin Questionnaire scores were negative. STOP-BANG had a negative predictive value 1.00 (0.77–1.00) for an AHI ≥15 with a score ≥3 predicting AHI ≥5 with sensitivity 0.93 (95% CI 0.84 to 0.98) and accuracy 79%, while a score ≥6 predicted AHI ≥15 with specificity 0.78 (0.65 to 0.88) and accuracy 72%. Neck circumference ≥17 inch and presence of witnessed apnoeas were independent predictors of SDB.

Conclusions: STOP and STOP-BANG Questionnaires have utility for the prediction of SDB in the sleep clinic population. Modification of the STOP-BANG Questionnaire merits further study in this and other patient groups.

INTRODUCTION

Obstructive sleep apnoea syndrome (OSAS) is common with prevalence of approximately 4% in middle-aged men and 2% in middle-aged women.[1] Frequent partial (hypopnoea) or complete (apnoea) upper airway collapse

Strengths and limitations of this study

- This is the first study to prospectively evaluate the utility of the Berlin, STOP and STOP-BANG Questionnaires in the prediction of sleep disordered breathing in the population referred to a sleep service for assessment of possible obstructive sleep apnoea (OSA).
- The results of this study show that the STOP and STOP-BANG, but not the Berlin Questionnaire, have utility for prediction of sleep disordered breathing in the sleep clinic population.
- This study uses home unattended limited sleep studies rather than in-hospital attended full polysomnography; however, this is considered standard clinical practice in the UK and is considered an acceptable method for diagnosis of OSA by the American Academy of Sleep Medicine.
- The sample size limits the conclusions that can be drawn from the multivariate analysis; however, this was a secondary objective of the study.

during sleep leads to oxygen desaturation, increased respiratory effort, arousal and sleep fragmentation.[2] Patients typically present with witnessed apnoeas, loud snoring and excessive daytime somnolence.[3] The syndrome is associated with impaired quality of life,[4] cognitive functioning and work performance,[5] and with increased risk of road traffic accidents.[6] OSAS is considered an independent risk factor for hypertension,[7] and has associations with coronary disease, stroke, heart failure, arrhythmias,[8] metabolic syndrome[9] and type 2 diabetes.[10]

Despite the substantial burden of this disease, it is under-recognised. One study estimated that 93% of women and 82% of men with moderate-to-severe OSAS were not clinically diagnosed,[11] and more recent data support this finding.[12] Sleep studies are required for OSAS diagnosis but are expensive and not widely available.[3] Given the recent increases in childhood[13] and adulthood obesity,[14] the workload for sleep clinics

and sleep laboratories will increase. Predictors of sleep disordered breathing (SDB) are required to allow recognition of OSAS and prioritisation of investigations.

Several questionnaires have been designed to screen for SDB in different populations. The Berlin Questionnaire was first validated in primary care against portable unattended sleep studies and a 'high risk' score predicted a respiratory disturbance index >5 with sensitivity 0.86, specificity 0.77, positive predictive value (PPV) 0.89 and likelihood ratio 3.79.[15] Its utilisation in other populations has been assessed with variable success.[16–22] The STOP and STOP-BANG Questionnaires were originally validated in surgical patients using in-hospital attended polysomnography.[23] For prediction of apnoea hypopnoea index (AHI) greater than 5, 15 and 30, sensitivities for the STOP and STOP-BANG Questionnaires were 65.6%, 74.3% and 79.5%, and 83.9%, 92.9% and 100%, respectively. The Berlin and STOP Questionnaires have been compared in a cohort of surgical patients[24] and the STOP and STOP-BANG Questionnaires have been compared in a large study involving several distinct cardiovascular and respiratory disease cohorts.[25] No study has, however, compared these screening tools in a sleep service-referred population. Finally, because of rising obesity rates, there is the potential for increasing recognition of SDB in primary care, and in the face of this evolution in sleep clinic practice it is therefore necessary to update and re-evaluate established assessment tools.

The objective of this study was, first, to compare the utility of Berlin, STOP and STOP-BANG Questionnaires for prediction of SDB in a population referred to the sleep clinic for assessment of possible OSA. Second, we sought to identify the most important variables from these questionnaires and routine sleep clinic assessment that might be utilised in the development of a composite predictive score for future use in this population.

METHODS
This was a prospective observational study conducted during May–December 2012. Study participants received an information sheet and provided informed consent.

Participants
Consecutive patients aged ≥16 years referred to the North Glasgow Sleep Service (a tertiary centre) for assessment of possible OSA were invited to participate.

Measurements
Height, weight, body mass index (BMI), neck circumference, blood pressure and Epworth Sleepiness Scale (ESS)[26] were completed at the Sleep Clinic. Participants attended the Sleep Laboratory on a separate day so that a Sleep Physiologist could provide, and instruct on fitting, a sleep study device. On that occasion, relevant symptoms and comorbidities were recorded, Mallampati score was assessed and the Berlin and STOP-BANG Questionnaires were completed. Blood samples

including a non-fasting lipid profile, glycated haemoglobin (HbA1c) and C reactive protein were taken. Two fractional exhaled nitric oxide (F_ENO) measurements were taken using the NIOX MINO (Aerocrine, Solna, Sweden), and the mean calculated.

Sleep studies
Unattended home limited polygraphy sleep studies were performed using the SOMNOmedics SOMNOscreen kit (Randersacker , Germany) with channels that recorded body position, thoracoabdominal movements, oronasal airflow, heart rate, pulse oximetry and snoring. Sleep study scoring by experienced Sleep Physiologists was in accordance with accepted guidelines.[27] An apnoea was defined as cessation of nasal flow for ≥10 s, while a hypnoea was defined as 50% reduction in nasal flow for ≥10 s, or lesser reduction in flow associated with oxygen desaturation of ≥4%.

The ESS, Berlin, STOP and STOP-BANG Questionnaires
The ESS is a validated measure of daytime sleepiness including eight questions, each with four possible responses, that assess the likelihood of dozing in different situations; a score of ≥11/24 denotes excessive daytime somnolence.[26] The Berlin Questionnaire includes questions in three categories that relate, first, to snoring and witnessed apnoeas, second, to tiredness, fatigue and sleepiness, and third, to hypertension and obesity.[15] High risk of OSA is defined by scoring positively in ≥2 categories. The STOP Questionnaire includes four yes/no questions that relate to Snoring, Tiredness, Observed apnoeas and high blood Pressure.[23] High risk of OSA is defined as a score of ≥2. The STOP-BANG Questionnaire includes four additional questions relating to BMI, age, neck circumference and gender, and high risk of OSA is defined as a score of ≥3.[23]

Statistical analyses
Statistical analyses were carried out using GraphPad Prism 5, IBM SPSS Statistics V.19 and STATA V.12. Normality of data was checked using the D'Agostino & Pearson omnibus normality test. A priori, two cut-points were chosen for AHI ≥5 events/h (the standard cutpoint for the diagnosis of OSA)[28] and ≥15 events/h, to predict significant SDB (the standard cut-point for initiating continuous positive airway pressure (CPAP) therapy).[28] Groups were compared using unpaired t tests, Mann-Whitney tests and Fisher's exact tests as appropriate. Sensitivities, specificities, PPV and negative predictive value (NPV), positive and negative likelihood ratios and overall accuracies were calculated for each of the questionnaires for prediction of SDB as defined by AHI cut-points ≥5 and ≥15. Associations between individual variables and each of the cut-points for AHI were explored using univariate and multivariate logistic regression. For multivariate analysis, in a few cases where BMI was known but neck circumference was not known, a value for the neck circumference was imputed using

linear regression with BMI as the independent value. This allowed for a dataset of 116 cases with all of the variables known or imputed to be built to identify independent variables for inclusion in a composite score. Receiver operating characteristic (ROC) curve analysis was used to assess predictive value and an area under the curve (AUC) >0.7 was considered clinically significant. Data are presented as mean (SD), median (IQR) and proportion (percentage), unless stated otherwise. A p value <0.05 was considered statistically significant.

RESULTS
In total 150 participants participated in this study, of which 129 had adequate sleep study data and were included in the analysis. AHI was ≥5 in 97/129 (75%) and ≥15 in 56/129 (43%). Overall, 82 (64%) were male, mean (SD) age was 49 (11) years, and median (IQR) BMI was 32 (29–39) kg/m^2.

Predicting SDB: patient characteristics
An AHI <5 ('rule-out measurement') was associated with female sex, younger age, lower weight and neck circumference, less frequently reported witnessed apnoeas, higher high-density lipoprotein (HDL) cholesterol, and lower triglycerides, cholesterol/HDL and HbA1c (see table 1). An AHI ≥15 ('rule-in measurement') was associated with male sex, obesity, higher weight, BMI and neck circumference, more frequently reported hypertension and witnessed apnoeas, lower HDL cholesterol, and higher triglycerides, cholesterol/HDL and HbA1c.

Predicting SDB: ESS, Berlin, STOP and STOP-BANG
The ESS and Berlin questionnaire outcomes were not associated with either AHI cut-point. An AHI <5 was associated with lower STOP and STOP-BANG scores and fewer participants being classified as 'high risk' for OSA by both STOP and STOP-BANG Questionnaires (see tables 2–4). An AHI ≥15 was associated with higher STOP and STOP-BANG scores and more participants being classified as 'high risk' for OSA by the STOP-BANG Questionnaire but not by the STOP questionnaire.

For the AHI cut-point of ≥5, the Berlin, STOP and STOP-BANG Questionnaires had high sensitivities, moderate PPVs and poor specificities and NPVs for prediction of SDB. The STOP-BANG Questionnaire performed best with an overall accuracy of 79%. For the AHI cut-point of ≥15, the Berlin Questionnaire had high sensitivity but otherwise performed poorly. The STOP and STOP-BANG Questionnaires had high sensitivities and NPVs. Again, the STOP-BANG Questionnaire performed best, but with a low overall accuracy of 56%. The low negative likelihood ratios for the STOP and STOP-BANG Questionnaires at both cut-points indicate that these questionnaires have value in excluding disease. As shown in table 4, the cut-points for the STOP-BANG score that was associated with best overall

accuracy were ≥3 and ≥6 for prediction of AHI ≥5 and ≥15, respectively.

SDB versus no SDB: predictors and a composite score
For the cut-point of AHI of ≥5, univariate logistic regression showed significant associations for age, gender, weight, neck circumference, witnessed apnoeas, triglycerides and cholesterol/HDL (p<0.05; see tables 5 and 6 and figure 1). For the cut-point of ≥15, significant associations were found for gender, weight, BMI, neck circumference, witnessed apnoeas, obesity, hypertension, F$_E$NO and cholesterol/HDL (p<0.05). Multivariate logistic regression based on the significant variables from univariate logistic regression showed that for both cut-points neck circumference and witnessed apnoeas were independent predictors of SDB. For the cut-point of AHI of ≥5, in a model incorporating neck circumference and witnessed apnoeas, the probability of SDB was 0.94 for individuals with neck circumference ≥17 inch and witnessed apnoeas (sensitivity 84%, overall accuracy 77%, ROC AUC 0.768, p<0.001). For the cut-point of AHI ≥15, the probability of SDB was 0.69 for individuals with neck circumference ≥17 inch and witnessed apnoeas (specificity 80%, overall accuracy 69%, ROC AUC 0.722, p<0.001).

DISCUSSION
This is the first study to prospectively evaluate the utility of the Berlin, STOP and STOP-BANG Questionnaires in prediction of SDB in a population referred to a tertiary sleep service for assessment of possible OSA. We found that in this population the Berlin Questionnaire had no significant association with cut-points of ≥5 or ≥15 for AHI, but that both the STOP and STOP-BANG scores were significantly associated with both cut-points. The STOP-BANG Questionnaire had better performance for the prediction of OSA on home sleep study, and different cut-points for the STOP-BANG score could be selected depending on the preference to exclude SDB (score <3) or predict SDB (score ≥6). In addition, we found notable associations between sleep study results and several patient characteristics. In particular, neck circumference and witnessed apnoeas were found to be independent predictors of SDB in our population.

In our study, the Berlin Questionnaire was almost ubiquitously positive (116 of 125 participants had a positive result) and the positivity rate did not differ between those with and without SDB. This was expected as this questionnaire was designed for primary care assessment and our study population consisted of individuals referred from primary care with symptoms suggestive of SDB. Our results indicate that the Berlin Questionnaire is not useful in the prediction of SDB in the sleep clinic referral population and this is consistent with previous reports.[19] The high sensitivities obtained for both AHI cut-points support previous findings that the Berlin Questionnaire may have a role as a 'rule-out'

Table 1 Comparison of patient characteristics including anthropometric measurements, comorbidities, symptoms, inflammatory markers, lipid profile and oximetry between groups without and with SDB as defined by AHI <5 and ≥15 events/h

	AHI <5	AHI ≥5	p Value	AHI <15	AHI ≥15	p Value
Male gender	*15/32 (47%)*	*67/97 (69%)*	*0.034*	*38/73 (52%)*	*44/56 (79%)*	*0.003*
Age (years)	*44 (12)*	*51 (11)*	*0.004*	48 (13)	51 (9)	0.103
Weight (kg)	*89 (19)*	*101 (22)*	*0.022*	*92 (21)*	*107 (20)*	*0.001*
BMI (kg/m^2)	31 (28–36)	33 (29–40)	0.118	*31 (27–36)*	*34 (31–41)*	*0.009*
Obesity (BMI ≥30 kg/m^2)	18/28 (64%)	61/85 (72%)	0.482	*38/63 (60%)*	*41/50 (82%)*	*0.014*
Neck circumference (inch)	*15 (2)*	*17 (2)*	*<0.001*	*16 (2)*	*17 (1)*	*<0.001*
Neck circumference ≥17 inch	*4/22 (18%)*	*45/76 (59%)*	*0.001*	*18/54 (33%)*	*31/44 (70%)*	*<0.001*
Mallampati	2 (1)	2 (1)	0.192	2 (1)	2 (1)	0.900
SBP (mm Hg)	134 (17)	136 (16)	0.480	134 (15)	138 (18)	0.194
DBP (mm Hg)	82 (11)	83 (9)	0.528	81 (10)	85 (10)	0.086
Diabetes	2/32 (6%)	6/97 (6%)	1.000	4/73 (5%)	4/56 (7%)	0.727
Hypertension	7/32 (22%)	36/97 (37%)	0.134	*18/73 (25%)*	*25/56 (45%)*	*0.024*
Hyperlipidaemia	2/32 (6%)	18/97 (19%)	0.156	10/73 (14%)	10/56 (18%)	0.625
Loud snorer	28/32 (88%)	92/97 (95%)	0.224	65/73 (89%)	55/56 (98%)	0.077
Witnessed apnoeas	*15/32 (47%)*	*72/97 (74%)*	*0.008*	*40/73 (55%)*	*47/56 (84%)*	*<0.001*
Nocturia*	11/32 (34%)	38/97 (39%)	0.679	26/73 (36%)	23/56 (41%)	0.585
Nocturnal wakenings*	18/32 (56%)	65/97 (67%)	0.293	46/73 (63%)	37/56 (66%)	0.853
Nocturnal choking	15/32 (47%)	35/97 (36%)	0.301	28/73 (38%)	22/56 (39%)	1.000
Nocturnal gasping	11/32 (34%)	38/97 (39%)	0.679	26/73 (36%)	23/56 (41%)	0.585
F$_E$NO (ppb)	15 (12–25)	18 (12–26)	0.595	15 (11–24)	19 (12–27)	0.050
Cholesterol (mmol/L)	5.2 (1.1)	5.4 (1.0)	0.431	5.3 (1.1)	5.4 (0.9)	0.674
HDL cholesterol (mmol/L)	*1.3 (1.0–1.5)*	*1.1 (1.0–1.3)*	*0.008*	*1.2 (1.0–1.4)*	*1.1 (1.0–1.2)*	*0.016*
Triglycerides (mmol/L)	*1.9 (1.0–2.5)*	*2.2 (1.5–3.2)*	*0.015*	*1.9 (1.3–2.9)*	*2.3 (1.6–3.2)*	*0.042*
Cholesterol/HDL	*4.1 (3.3–5.0)*	*4.7 (3.8–5.7)*	*0.011*	*4.3 (3.5–5.6)*	*4.8 (4.0–5.8)*	*0.022*
HbA1c (mmol/mol)	*34 (32–37)*	*38 (36–41)*	*0.001*	*36 (33–39)*	*38 (36–42)*	*0.002*
CRP	3.0 (1.2–8.5)	3.8 (1.4–7.5)	0.608	3.0 (1.3–7.6)	4.3 (1.4–8.6)	0.173

Data presented as mean (SD), median (IQR) or proportion (percentage) as appropriate. Significant differences in italics.
*≥2/Night.
AHI, apnoea hypopnoea index; BMI, body mass index; CRP, C reactive protein; DBP, diastolic blood pressure; F$_E$NO, fraction of exhaled nitric oxide; HbA1c, glycated haemoglobin; HDL, high-density lipoprotein; ODI, oxygen desaturation index; SBP, systolic blood pressure; SDB, sleep disordered breathing.

measurement in the primary care or screening setting,[15 17 20 24] though there have been some conflicting results, suggesting that it does not have adequate discriminatory power.[16 22]

In our study, ESS data indicated that two-thirds of participants had excessive daytime somnolence (ESS ≥11); however, scores were similar in individuals with or without SDB. Therefore, at least in the sleep clinic population, the ESS is not useful for the prediction of SDB. It may be of value perhaps if it is combined with other measures, including those highlighted in this study, in prediction of compliance with and benefit from OSA treatment. Further research is required to address this question. The exhaled nitric oxide levels were not significantly different between individuals with or without SDB, whether defined by an AHI cut-point of ≥5 or ≥15. There are conflicting data in the literature regarding whether F$_E$NO is associated with SDB;[29–32] however,

Table 2 Comparison of results of ESS and Berlin, STOP and STOP-BANG Questionnaires between groups without and with SDB as defined by AHI <5 and ≥15 events/h

	AHI <5	AHI ≥5	p Value	AHI <15	AHI ≥15	p Value
ESS score	13 (8–16)	13 (7–17)	0.845	13 (7–16)	13 (9–18)	0.476
ESS positive (≥11/24)	17/28 (61%)	63/92 (68%)	0.496	42/66 (64%)	38/54 (70%)	0.560
Berlin positive	29/31 (94%)	87/94 (93%)	1.000	65/71 (92%)	51/54 (94%)	0.731
STOP score	*2.5 (2–3)*	*3 (2–3)*	*0.011*	*3 (2–3)*	*3 (3–3.5)*	*<0.001*
STOP positive (≥2/4)	*27/32 (84%)*	*93/96 (97%)*	*0.023*	66/73 (90%)	54/55 (98%)	0.137
STOP-BANG score	*4 (2–5)*	*5 (5–6)*	*<0.001*	*5 (2–5)*	*6 (5–6)*	*<0.001*
STOP-BANG positive (≥3/8)	*21/30 (70%)*	*88/93 (95%)*	*<0.001*	*54/68 (79%)*	*55/55 (100%)*	*<0.001*

Data presented as median (IQR) or proportion (percentage) as appropriate. Significant differences in italics.
ESS, Epworth Sleepiness Scale; SDB, sleep disordered breathing.

Table 3 Sensitivities, specificities, PPV, NPV, LR+, LR– and overall accuracies of the Berlin, STOP and STOP-BANG Questionnaires for prediction of significant SDB as defined by: model A: AHI ≥5 events/h; model B: ≥15 events/h

	Sensitivity	Specificity	PPV	NPV	LR+	LR–	Accuracy (%)
Model A							
Berlin	0.93 (0.85 to 0.97)	0.06 (0.01 to 0.21)	0.75 (0.66 to 0.83)	0.22 (0.03 to 0.60)	0.99	1.15	71
STOP	0.97 (0.91 to 0.99)	0.16 (0.05 to 0.33)	0.78 (0.69 to 0.85)	0.62 (0.24 to 0.91)	1.15	0.20	77
STOP-BANG	0.95 (0.88 to 0.98)	0.30 (0.15 to 0.49)	0.81 (0.72 to 0.88)	0.64 (0.35 to 0.87)	1.35	0.18	79
Model B							
Berlin	0.94 (0.85 to 0.99)	0.08 (0.03 to 0.17)	0.44 (0.35 to 0.53)	0.67 (0.30 to 0.92)	1.03	0.66	46
STOP	0.98 (0.90 to 1.00)	0.10 (0.04 to 0.19)	0.45 (0.36 to 0.54)	0.88 (0.47 to 1.00)	1.09	0.19	48
STOP-BANG	1.00 (0.94 to 1.00)	0.21 (0.12 to 0.32)	0.50 (0.41 to 0.60)	1.00 (0.77 to 1.00)	1.26	0.00	56

Data presented with 95% CIs.
AHI, apnoea hypopnoea index; LR+, likelihood ratios positive; LR–, likelihood ratios negative; NPV, negative predictive values; PPV, positive predictive values; SDB, sleep disordered breathing.

our results suggest that it does not have utility in prediction of SDB; further work is required to clarify this.

We found that the STOP and STOP-BANG Questionnaires have utility in the prediction of SDB in the sleep clinic population, and that STOP-BANG was superior, with higher overall predictive accuracy. The STOP and STOP-BANG Questionnaires were developed and validated in a surgical population using in-laboratory polysomnography[23] and have subsequently been studied in a cardiovascular disease population.[25] Our results are in agreement with these two earlier studies as regards the increased predictive value of STOP-BANG over STOP. In contrast to these earlier studies, however, we found sensitivities to be higher and specificities to be lower for both cut-points of AHI. We suggest that of the two AHI cut-points, ≥15 events/h is the more important, being diagnostic of at least moderate SDB and also an indication for CPAP treatment. At this cut-point, STOP and STOP-BANG performed with high sensitivities and negative predictive values (STOP-BANG being superior to STOP), indicating that these questionnaires are more useful in excluding significant SDB. This is further corroborated by the negative likelihood ratios of <0.2 obtained for STOP and STOP-BANG that also indicate that these questionnaires are most useful in ruling out SDB. STOP-BANG may be of value in the primary care setting, perhaps if combined with type IV portable monitoring sleep studies, to determine requirement for sleep clinic review and more detailed polygraphy.

At the AHI cut-point of ≥15, STOP-BANG had sensitivity and NPV of 100%, and since this is the standard cut-point conventionally used to determine need for CPAP,[28] we suggest that the STOP-BANG Questionnaire is the preferred tool for prediction of SDB in the sleep clinic setting of those currently available. STOP-BANG, perhaps with modifications, merits further evaluation for the prediction of SDB in the sleep clinic population and, more importantly, its utility in prediction of clinical outcomes including treatment success should be assessed.

The original STOP-BANG Questionnaire uses a cut-point of ≥3 to predict SDB.[23] However, in our study, we show that different cut-points can be selected depending on the preference to rule-in or rule-out SDB. A score of ≥3 had the highest overall accuracy and a sensitivity of 0.93 for the AHI cut-point of ≥5, whereas a score of ≥6 had the highest overall accuracy and a specificity of 0.78 for the AHI cut-point of ≥15. Two other studies have examined the usefulness of different cut-points for the STOP-BANG score.[33 34] In the obese, a score of ≥3 was associated with a sensitivity of 0.90 for predicting an AHI >5, while a score of ≥6 had a specificity of 0.88 for predicting an AHI >15 and similar results have been obtained in the morbidly obese[33] and in a surgical population.[34] Thus, in the sleep clinic setting where the ultimate goal is to identify patients requiring CPAP, a higher cut-point for STOP-BANG may be preferred, whereas in a primary care setting where the priority is not to miss disease, a lower cut-point may be chosen.

The STOP-BANG Questionnaire is, however, still an imperfect tool for prediction of results on home polygraphy. Accordingly, the secondary objective of our study was to identify variables for inclusion in a locally developed composite score for future validation in the sleep clinic and potentially wider population. Univariate analysis showed several significant, expected associations for both cut-points of AHI. Using multivariate analysis, neck circumference ≥17 inch and the presence of witnessed apnoeas were independent predictors of SDB. This is not a novel finding, but it does support the robustness of our data. Particularly when SDB was defined by an AHI cut-point of ≥5, the regression model derived indicated a high probability of SDB of 0.94 if both factors were present. The STOP-BANG Questionnaire, of course, includes both these variables, and it is possible that adjustment of the inclusion variables, or their weighting, might improve its performance. In future work, we aim to validate a simple composite score based on these two variables in a modification of STOP-BANG, to determine utility for predicting sleep study data and outcomes with treatment.

Table 4 Sensitivities, specificities, PPV, NPV, LR+, LR− and overall accuracies of different cut-points of the STOP-BANG questionnaire for prediction of significant SDB as defined by: model A: AHI ≥5 events/h; model B: ≥15 events/h

STOP-BANG cut-point (/8)	Proportion positive SB	Sensitivity	Specificity	PPV	NPV	LR+	LR−	Accuracy (%)
Model A								
≥2	92/96	0.99 (0.92 to 1.00)	0.12 (0.03 to 0.32)	0.77 (0.67 to 0.85)	0.75 (0.19 to 0.99)	1.13	0.11	77
≥3	82/96	0.93 (0.84 to 0.98)	0.38 (0.19 to 0.59)	0.82 (0.72 to 0.89)	0.64 (0.35 to 0.87)	1.49	0.18	79
≥4	79/96	0.90 (0.81 to 0.96)	0.42 (0.22 to 0.63)	0.82 (0.72 to 0.90)	0.59 (0.33 to 0.82)	1.55	0.23	78
≥5	67/96	0.81 (0.70 to 0.89)	0.62 (0.41 to 0.81)	0.87 (0.76 to 0.94)	0.52 (0.32 to 0.71)	2.15	0.31	76
≥6	38/96	0.46 (0.34 to 0.58)	0.79 (0.58 to 0.93)	0.87 (0.72 to 0.96)	0.33 (0.21 to 0.46)	2.20	0.68	54
≥7	15/96	0.18 (0.10 to 0.29)	0.92 (0.73 to 0.99)	0.87 (0.60 to 0.98)	0.27 (0.18 to 0.38)	2.17	0.89	36
8	5/96	0.07 (0.02 to 0.15)	1.00 (0.86 to 1.00)	1.00 (0.48 to 1.00)	0.26 (0.18 to 0.37)	−	0.93	30
Model B								
≥2	92/96	1.00 (0.91 to 1.00)	0.07 (0.02 to 0.18)	0.45 (0.34 to 0.55)	1.00 (0.40 to 1.00)	1.08	0	47
≥3	82/96	1.00 (0.91 to 1.00)	0.25 (0.15 to 0.39)	0.50 (0.39 to 0.61)	1.00 (0.77 to 1.00)	1.34	0	57
≥4	79/96	1.00 (0.91 to 1.00)	0.31 (0.19 to 0.45)	0.52 (0.40 to 0.63)	1.00 (0.80 to 1.00)	1.45	0	60
≥5	67/96	0.93 (0.80 to 0.98)	0.47 (0.34 to 0.61)	0.57 (0.44 to 0.69)	0.90 (0.73 to 0.98)	1.76	0.15	67
≥6	38/96	0.63 (0.47 to 0.78)	0.78 (0.65 to 0.88)	0.68 (0.51 to 0.82)	0.74 (0.61 to 0.85)	2.91	0.47	72
≥7	15/96	0.22 (0.11 to 0.38)	0.89 (0.78 to 0.96)	0.60 (0.32 to 0.84)	0.60 (0.49 to 0.71)	2.01	0.88	60
8	5/96	0.10 (0.03 to 0.23)	0.98 (0.90 to 1.00)	0.80 (0.28 to 0.99)	0.59 (0.48 to 0.70)	5.36	0.92	60

Data presented with 95% CIs.
AHI, apnoea hypopnoea index; LR+, likelihood ratios positive; LR−, likelihood ratios negative; NPV, negative predictive values; PPV, positive predictive values; SDB, sleep disordered breathing.

Table 5 Probabilities of SDB using a composite score based on neck circumference (≥17 or <17 inch) and presence (1) or absence (0) of witnessed apnoeas derived from logistic regression models

Neck circumference (inch)	Witnessed apnoeas	
Model A	0	1
<17	0.47	0.75
≥17	0.83	0.94
Model B	0	1
<17	0.17	0.40
≥17	0.40	0.69

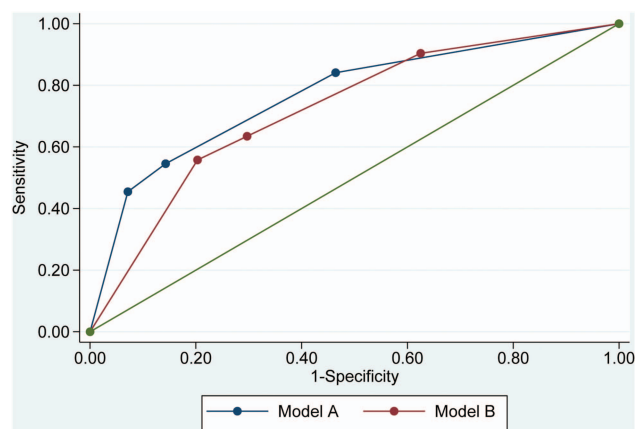SDB defined by: A: AHI ≥5 events/h; B: ≥15 events/h.

Figure 1 Receiver operating characteristic (ROC) curves of logistic models for prediction of sleep disordered breathing (SDB) as defined by: model A: apnoea hypopnoea index (AHI) ≥5 events/h; predict SDB *unless* neck <17 inch *and* witnessed apnoeas absent; model B: AHI ≥15 events/h; predict SDB *only if* neck ≥17 inch *and* witnessed apnoeas present.

Ultimately, a predictive tool that can be utilised in primary care is the goal. Our results indicate low specificity of STOP-BANG, and therefore in its current form, if used in primary care to identify patients requiring referral for further assessment, it is likely to result in a significant percentage of patients being referred unnecessarily (false positives). It is hoped that a modified STOP-BANG with improved specificity, while not compromising sensitivity, may be developed that can be used safely in primary care for identification of patients requiring referral to sleep services. Of upmost importance too is the prediction of treatment outcome. Non-adherence to CPAP treatment occurs between 46% and 83%.[35] [36] Prediction of poor adherence by STOP-BANG or other similar tools would allow greater attention to interventions to improve adherence in patients more likely to default from treatment. The authors are not aware of any studies investigating this question and future research should explore this important issue.

A possible limitation of our study was that SDB was characterised using home unattended limited sleep studies rather than in-hospital attended full polysomnography. The latter is considered the gold standard for diagnosis of SDB but is more expensive, less easily accessed and potentially unrepresentative with sleep in an unfamiliar environment. Home unattended and in-hospital attended sleep studies have previously been shown to produce similar results.[37] Accordingly, home testing with portable monitors is standard clinical practice in the UK, and is now considered an acceptable

method for diagnosis of OSA by the American Academy of Sleep Medicine.[28] The sample size limits the conclusions that can be drawn from multivariate analysis; however, this was a secondary objective of the current study. It is possible that variables predictive of SDB on univariate analysis in this cohort would have been identified as independently predictive in multivariate models in a larger population. The results of this study allow us, and potentially others, to focus future work to validate more extensively the results obtained until now. We chose AHI cut-points of ≥5 and ≥15 to define significant SDB. This was based on the consensus guideline produced by the Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine that states that diagnosis of OSA is based on a cut-point of >15 events/h or >5 events/h with relevant symptoms, and that CPAP is indicated for treatment of moderate-to-severe OSA with ≥15 events/h.[28] Although the cut-point of >30 events/h is consistent with severe OSA, we suggest that this cut-point is less relevant clinically from a diagnostic perspective or from that of determining treatment. Finally, owing to the prospective design of our study, we cannot comment on the relative value of other tools developed for prediction of OSA such as the

Table 6 Sensitivities, specificities, PPV, NPV and overall accuracies of logistic models for prediction of SDB as defined by: A: AHI ≥5 events/h; predict SDB *unless* neck <17 inch *and* witnessed apnoeas absent; B: AHI ≥15 events/h; predict SDB *only if* neck ≥17 inch *and* witnessed apnoeas present

Model	Sensitivity	Specificity	PPV	NPV	Accuracy (%)
A	0.84 (0.75 to 0.91)	0.54 (0.34 to 0.72)	0.85 (0.76 to 0.92)	0.52 (0.32 to 0.71)	77
B	0.56 (0.41 to 0.70)	0.80 (0.68 to 0.89)	0.69 (0.53 to 0.82)	0.69 (0.57 to 0.79)	69

Data presented with 95% CIs.
AHI, apnoea hypopnoea index; LR+, likelihood ratios positive; LR−, likelihood ratios negative; NPV, negative predictive values; PPV, positive predictive values; SDB, sleep disordered breathing.

Sleep Apnea Clinical Score[38] and American Society of Anesthesiologists Checklist.[39] To compare their utility with that of the Berlin, STOP and STOP-BANG Questionnaires in the population referred to, the sleep service would require a further study.

In conclusion, the Berlin Questionnaire was not useful in the prediction of SDB within our sleep clinic population. The STOP-BANG Questionnaire had superior predictive performance to the STOP Questionnaire at both cut-points of AHI (≥5 and ≥15). A STOP-BANG score of ≥3 had the highest overall accuracy and a sensitivity of 0.93 for the prediction of an AHI ≥5, while a score of ≥6 had the highest overall accuracy and a specificity of 0.78 for the prediction of an AHI ≥15. Future work will validate a composite score including neck circumference ≥17 inch and the presence of witnessed apnoeas for the prediction of SDB in the sleep clinic referral population. An optimised composite score could then be evaluated in primary care and against treatment outcomes, with our overall aim being to provide required tools for use in the expanded and consolidated sleep services that are now necessary given the current obesity and OSA epidemics.

REFERENCES

1. Young T, Palta M, Dempsey J, *et al*. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;328:1230–5.
2. Jordan AS, McSharry DG, Malhotra A. Adult obstructive sleep apnoea. *Lancet* 2014;383:736–47.
3. McNicholas WT. Diagnosis of obstructive sleep apnea in adults. *Proc Am Thorac Soc* 2008;5:154–60.
4. Lacasse Y, Godbout C, Series F. Health-related quality of life in obstructive sleep apnoea. *Eur Respir J* 2002;19:499–503.
5. Engleman HM, Douglas NJ. Sleep. 4: sleepiness, cognitive function, and quality of life in obstructive sleep apnoea/hypopnoea syndrome. *Thorax* 2004;59:618–22.
6. Teran-Santos J, Jimenez-Gomez A, Cordero-Guevara J. The association between sleep apnea and the risk of traffic accidents.

7. Robinson GV, Stradling JR, Davies RJ. Sleep 6: obstructive sleep apnoea/hypopnoea syndrome and hypertension. *Thorax* 2004;59:1089–94.
8. Lattimore JD, Celermajer DS, Wilcox I. Obstructive sleep apnea and cardiovascular disease. *J Am Coll Cardiol* 2003;41:1429–37.
9. Svatikova A, Wolk R, Gami AS, *et al*. Interactions between obstructive sleep apnea and the metabolic syndrome. *Curr Diab Rep* 2005;5:53–8.
10. Meslier N, Gagnadoux F, Giraud P, *et al*. Impaired glucose-insulin metabolism in males with obstructive sleep apnoea syndrome. *Eur Respir J* 2003;22:156–60.
11. Young T, Evans L, Finn L, *et al*. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep* 1997;20:705–6.
12. Heffner JE, Rozenfeld Y, Kai M, *et al*. Prevalence of diagnosed sleep apnea among patients with type 2 diabetes in primary care. *Chest* 2012;141:1414–21.
13. Wang Y, Lobstein T. Worldwide trends in childhood overweight and obesity. *Int J Pediatr Obes* 2006;1:11–25.
14. Flegal KM, Carroll MD, Ogden CL, *et al*. Prevalence and trends in obesity among US adults, 1999–2000. *JAMA* 2002;288:1723–7.
15. 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.
16. Sert Kuniyoshi FH, Zellmer MR, Calvin AD, *et al*. Diagnostic accuracy of the Berlin Questionnaire in detecting sleep-disordered breathing in patients with a recent myocardial infarction. *Chest* 2011;140:1192–7.
17. Best MW, Fitzpatrick M, Milev R, *et al*. Utility of the Berlin questionnaire for predicting obstructive sleep apnea in individuals with treatment-resistant depression. *Sleep Breath* 2013;17:1221–7.
18. Kang K, Park KS, Kim JE, *et al*. Usefulness of the Berlin Questionnaire to identify patients at high risk for obstructive sleep apnea: a population-based door-to-door study. *Sleep Breath* 2013;17:803–10.
19. Ahmadi N, Chung SA, Gibbs A, *et al*. The Berlin questionnaire for sleep apnea in a sleep clinic population: relationship to polysomnographic measurement of respiratory disturbance. *Sleep Breath* 2008;12:39–45.
20. Bouloukaki I, Komninos ID, Mermigkis C, *et al*. Translation and validation of Berlin questionnaire in primary health care in Greece. *BMC Pulm Med* 2013;13:6.
21. Saleh AB, Ahmad MA, Awadalla NJ. Development of Arabic version of Berlin questionnaire to identify obstructive sleep apnea at risk patients. *Ann Thorac Med* 2011;6:212–16.
22. Sforza E, Chouchou F, Pichot V, *et al*. Is the Berlin questionnaire a useful tool to diagnose obstructive sleep apnea in the elderly? *Sleep Med* 2011;12:142–6.
23. Chung F, Yegneswaran B, Liao P, *et al*. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008;108:812–21.
24. 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.
25. Silva GE, Vana KD, Goodwin JL, *et al*. Identification of patients with sleep disordered breathing: comparing the four-variable screening tool, STOP, STOP-Bang, and Epworth Sleepiness Scales. *J Clin Sleep Med* 2011;7:467–72.
26. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991;14:540–5.
27. 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.
28. Epstein LJ, Kristo D, Strollo PJ, *et al*. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med* 2009;5:263–76.
29. Chua AP, Aboussouan LS, Minai OA, *et al*. Long-term continuous positive airway pressure therapy normalizes high exhaled nitric oxide levels in obstructive sleep apnea. *J Clin Sleep Med* 2013;9:529–35.
30. Fortuna AM, Miralda R, Calaf N, *et al*. Airway and alveolar nitric oxide measurements in obstructive sleep apnea syndrome. *Respir Med* 2011;105:630–6.
31. Depalo A, Carpagnano GE, Spanevello A, *et al*. Exhaled NO and iNOS expression in sputum cells of healthy, obese and OSA subjects. *J Intern Med* 2008;263:70–8.
32. Agusti AG, Barbe F, Togores B. Exhaled nitric oxide in patients with sleep apnea. *Sleep* 1999;22:231–5.

Cooperative Group Burgos-Santander. *N Engl J Med* 1999;340:847–51.

33. Chung F, Yang Y, Liao P. Predictive performance of the STOP-Bang score for identifying obstructive sleep apnea in obese patients. *Obes Surg* 2013;23:2050–7.

34. Chung F, Subramanyam R, Liao P, *et al*. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. *Br J Anaesth* 2012;108:768–75.

35. Weaver TE, Grunstein RR. Adherence to continuous positive airway pressure therapy: the challenge to effective treatment. *Proc Am Thorac Soc* 2008;5:173–8.

36. Wolkove N, Baltzan M, Kamel H, *et al*. Long-term compliance with continuous positive airway pressure in patients with obstructive sleep apnea. *Can Respir J* 2008;15:365–9.

37. Iber C, Redline S, Kaplan Gilpin AM, *et al*. Polysomnography performed in the unattended home versus the attended laboratory setting—Sleep Heart Health Study methodology. *Sleep* 2004;27:536–40.

38. 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.

39. Gross JB, Bachenberg KL, Benumof JL, *et al*. Practice guidelines for the perioperative management of patients with obstructive sleep apnea: a report by the American Society of Anesthesiologists Task Force on Perioperative Management of patients with obstructive sleep apnea. *Anesthesiology* 2006;104:1081–93.