Development and validation of a model for diagnosis of obstructive sleep apnoea in primary care

Chapman, Julia L.; Hoyos, Camilla M.; Killick, Roo; Sutherland, Kate; Cistulli, Peter A.; Zwar, Nick; Yee, Brendon J.; Marks, Guy; Grunstein, Ronald R.; Wong, Keith K.H.; The Sydney OSA-GP study Investigators

Published in:
Respirology

DOI:
10.1111/resp.14122

Licence:
Other

ORIGINAL ARTICLE:

TITLE: Development and validation of a model for diagnosis of obstructive sleep apnoea in primary care.

Authors' full names: Julia L Chapman[1,2], Camilla M Hoyos[1,2], Roo Killick[1], Kate Sutherland[3], Peter A Cistulli[3,4], Nick Zwar[5], Brendon J Yee[1,6], Guy Marks[7], Ronald R Grunstein[1,8], Keith KH Wong[1,6] on behalf of the Sydney OSA-GP study Investigators

Authors' affiliation(s):
1. Centre for Sleep and Chronobiology, Woolcock Institute of Medical Research, University of Sydney, Glebe, NSW, Australia
2. School of Psychology and Brain and Mind Centre, University of Sydney, Camperdown, Sydney
3. Charles Perkins Centre, University of Sydney, Sydney, Australia
4. Department of Respiratory and Sleep Medicine, Royal North Shore Hospital, St Leonards, NSW, Australia
5. Faculty of Health Sciences & Medicine, Bond University, QLD, Australia
6. Department of Respiratory and Sleep Medicine, Royal Prince Alfred Hospital, Camperdown, NSW, Australia
7. South Western Sydney Clinical School, University of New South Wales, Liverpool NSW Australia
8. Charles Perkins Centre - Royal Prince Alfred Clinic, Royal Prince Alfred Hospital, Camperdown, NSW, Australia

Sydney OSA-GP study investigators: Amanda Siebers, George Dungan II, Sarah Dennis, Lydia Makarie Rofail, *Liverpool Hospital* Hima Vedam, Peter Buchanan, Zinta Harrington, *Royal North Shore Hospital* Richard Lee, Andrew S. L. Chan, Mark Williams, *St George Hospital* Andrew Ng, Eric Wong, *Psychology Department Macquarie University* Michael Jones.

Corresponding author full contact details:
Name: Dr Keith Wong
Address: Woolcock Institute of Medical Research, PO Box M77
Post code:2050
City: Missenden Road
Country: Australia
Email: keith.wong@sydney.edu.au

SUMMARY AT A GLANCE: *(50/50 words).*
Obstructive sleep apnoea was reliably diagnosed using a simplified diagnostic technique and a home-worn device that can be applied in the primary care setting without polysomnography. Primary care physicians can identify patients at high suspicion of sleep apnoea, with little additional benefit from screening questionnaires despite recommendations for their use.

Manuscript submission template:

ABSTRACT:

Background and objective: Use of in-laboratory polysomnography (PSG) to diagnose obstructive sleep apnoea (OSA) is cost and resource intensive. Questionnaires, physical measurements and home monitors have been studied as potential simpler alternatives. This study aimed to develop a diagnostic model for OSA for use in primary care.

Methods: Primary care practitioners were trained to recognise symptoms of sleep apnoea and recruited patients based on clinical need to investigate OSA. Assessment was by symptom questionnaires, anthropomorphic measurements, digital facial photography, and a single channel nasal flow monitor (Flow Wizard©, DiagnoseIT, Sydney, Australia) worn at home for 3 nights. The in-laboratory PSG was the reference test, with OSA defined as apnoea-hypopnoea index (AHI) ≥10 events/hr.

Results: In the model development phase, 25 primary care practitioners studied 315 patients in whom they suspected OSA, of which 57% had AHI≥10 and 22% had AHI≥30. Published OSA questionnaires provided low to moderate prediction of OSA (AUC 0.53-0.73). The nasal flow monitor alone yielded high accuracy for predicting OSA with area under the curve (AUC) of 0.87. Sensitivity was 0.87 and specificity 0.77 at a threshold respiratory event index (REI) of 18 events/hr. A model adding age, gender, symptoms and BMI to the nasal flow monitor REI only modestly improved OSA prediction (AUC 0.89), with similar AUC (0.88) confirmed in the validation population of 114 patients.

Conclusion: Sleep apnea can be diagnosed in the primary care setting with a combination of clinical judgement and portable monitor test outcomes.

Short title:

Manuscript submission template:

Diagnosis of OSA in primary care

Keywords: *(5 words selected from the list recommended by the US National Library of Medicine's Medical Subject Headings (MeSH) at http://www.nlm.nih.gov/mesh/meshhome.html.)*

Sleep Apnea Syndromes

Diagnosis

Primary Health Care

Validation Study

Models, Statistical

INTRODUCTION:

Obstructive Sleep Apnoea (OSA) affects at least 2-4% of the adult population(1, 2) with more recent estimates showing prevalence may be as high as 50%.(3, 4) Untreated OSA is associated with increased cardiovascular risk, negative neurobehavioural outcomes and all-cause mortality.(5) It is therefore important to find and treat people with OSA.

In-laboratory polysomnography (PSG), the gold standard test to diagnose OSA recommended(6, 7) is resource-intensive and limited to sleep laboratory facilities, which are less accessible for those living outside of metropolitan areas.(8) In Australia, the rate of in-laboratory PSGs in 2017-2018 was 378 per 100,000 head of population. Even with the advent of government-funded at-home PSGs in 2008, adding an additional 351 sleep assessments per 100,000 in 2017-2018,(9) the number of investigations still falls short of the estimated 2,310 per 100,000 required to diagnose and treat OSA.(8) While at-home PSG is cheaper and there has been a rise in the quantity performed,(9) they are still costly. Guidelines still recommend full PSG where available,(7) but a simplified and more streamlined method of diagnosis is required to ensure that the demand for OSA diagnostic services is met. Some alternative models of care are already in use, despite limited evidence of their validity.(10) Moreover, accurate diagnosis of OSA is important given that it may impact on driver licensing, employment or life insurance.

Primary care is a setting where chronic diseases are often diagnosed and managed. Almost one-third of primary care practice attendees in Europe and North America were at high-risk for OSA.(11) A simple and validated test that can be implemented within the primary care setting would be useful to rapidly diagnose patients depending on their likelihood of OSA.(12) This would greatly reduce the burden on the healthcare system, especially as it has

been suggested that patients with uncomplicated OSA can be managed equally well by their primary care physician compared to a specialist centre.(13, 14)

There is limited research identifying simpler tests for OSA in primary care settings. Research, predominantly in sleep centre settings, has examined simplified methods for the diagnosis of OSA, including symptom questionnaires(15, 16) demographics and anthropometric measures,(17-19) and analysis of the physical features of the face.(20) Simplified sleep apnoea testing devices usually use one or more of the multiple signals employed with a PSG (e.g. pulse oximetry) but there is heterogeneity in the literature arising from differences in devices available and how OSA severity is classified.(21)

No study has attempted to build a diagnostic model for OSA that complements the clinical impressions of primary care physicians. A precise prediction model would streamline the diagnosis of OSA in primary care, allowing more patients to be better managed for this chronic disease in that setting.

This study aimed to develop a simplified diagnostic model using a combination of data from a number of OSA screening questionnaires, anthropometry, craniofacial photography and a portable nasal flow monitor for the simplified diagnosis of clinically significant OSA to complement the impressions of the primary care physician. This model was then prospectively validated using a second sample of patients.

METHODS:

This study was performed in Sydney, Australia at the Woolcock Institute of Medical Research, Liverpool Hospital, Royal North Shore Hospital and St George Hospital. This study was approved by the Human Research Ethics Committee of Sydney South West Area

Health Service (now Sydney Local Health District, Protocol no. X08-0125) as well as by the ethics committees of the other local sites/hospitals. The study was registered with the Australian and New Zealand Clinical Trial Registry (*ACTRN no. 12608000417381*).

Participants

Primary Care Practitioners (PCPs) were invited through advertising in medical publications. They were then provided with training (2x 2.5hour sessions) on study procedures and regarding the clinical presentation of OSA. See online supplement for further details.

Patients presenting to enrolled PCPs, who were identified as needing further evaluation for suspected OSA, were invited to participate in the study. Inclusion criteria were: aged 18-75; symptoms of sleep apnoea (e.g. snoring, choking, witnessed apnoeas or daytime sleepiness) or conditions associated with sleep apnoea (e.g. obesity, hypertension or metabolic syndrome). Patients were excluded if they were unable to comply with procedures, unable to wear the diagnostic device or if the suspected diagnosis was primarily a non-OSA sleep disorder. Written consent was obtained from all patients prior to data collection. Questionnaires were completed by the patient and PCP, sent to the coordinating site (Woolcock Institute) who randomised the participant to order of in-laboratory PSG and home sleep testing. All participants were offered follow-up with a sleep physician.

Assessments

The PCP completed: age, gender, weight, height, neck, waist, hip circumference, smoking and alcohol consumption, tonsillar size and palate position. The patient completed: Epworth Sleepiness Scale (ESS),(26) Multivariable Apnea Prediction Index (MAPI)(15) and Berlin Questionnaire.(11) Craniofacial phenotyping was undertaken using a previously developed digital photography technique.(20) PCPs received a monetary reimbursement for the paperwork and data collection involved with each patient referral ($100 AUD), and for each patient photographed ($30 AUD).

Home sleep testing was performed on three consecutive nights using a nasal flow monitor (Flow Wizard©, DiagnoseIT, Sydney, Australia), an Australian Therapeutic Goods Administration-approved, CE-marked device that records airflow by nasal pressure cannula.(27) It automatically scored respiratory event index (REI) and signal quality. The mean of the individual night REIs was used in the analysis.

All patients underwent an in-laboratory PSG as the reference standard, in one of five participating sleep centres. Scoring was performed manually by Registered Sleep Technologists, according to standard criteria(28, 29) and blinded to the results of the home nasal flow monitor. An apnea was defined as a complete cessation of breathing as shown in the nasal flow signal for at least 10 seconds, while a hypopnea was a clear amplitude reduction of airflow for 10 seconds or more accompanied by an arousal or 3% desaturation (see Supplement). The diagnosis of OSA was made based on an apnoea hypopnoea index (AHI) of ≥10 events per hour.

Statistical Methods

Data from the first 360 participants (development population) were used to develop the diagnostic model, while the remaining participants comprised the validation population, analysed only after analysis for the development population was completed.

Multivariable logistic regression models were built using variables chosen from univariate analyses based on the area under the curve (AUC) of the receiver operating characteristic (ROC), and prior knowledge of OSA risk factors, to develop a predictive model for diagnosis of OSA. Diagnostic test statistics (sensitivity, specificity, positive and negative predictive values) were calculated at the threshold closest to the top left corner of the ROC curve. The optimal model and threshold chosen from the development population was then tested on the validation population. Further detail can be seen in the supplementary methods.

The development population required 360 participants to achieve statistical power of 0.8 at a two-tailed alpha of 0.05 given an odds ratio of 2.0 assuming a correlation (r^2) between variables of 0.2. For the validation population we required 210 participants, to ensure the upper 95% confidence limit of the negative likelihood ratio was at least 1/3 and the lower confidence limit of the positive likelihood ratio was at least 3, assuming a prevalence of 20%.

RESULTS:

Participants

Twenty-nine PCPs were trained in the study procedures, of which 25 referred at least one patient into the study (See Supplemental Table 1). A total of 575 patients were referred between September 2008 and December 2010, of which 500 were randomised, and 429 patients completed both the home flow monitor and the PSG (Figure 1). The major reasons for non-completion were that the participant withdrew consent prior to testing (n=41) or that they were unable to be contacted for follow up (n=21). Figure 1 shows that due to test failures, despite reattempts, there was data for 315 participants available for the model development and 114 participants for the model validation phases of the study.

Table 1 shows the participant characteristics, which showed a typical population at risk of OSA; predominantly male, middle aged and overweight. Average AHI was in the mild to moderate range of OSA. There were no differences in these characteristics between the development and validation populations.

Development population

On univariate analysis (Table 2), the traditional risk factors for OSA (age, gender, BMI), and sleep apnoea symptoms (as measured by the 3 questions concerning snoring, gasping, and breathing stops that comprise index 1 of the MAP questionnaire) were not highly predictive of OSA individually (sensitivity and specificity both below 75% for PSG AHI≥10, AUC<0.73 for all). The Berlin questionnaire had a sensitivity of 86%, but only a specificity

of 21% at this threshold. The flow monitor was the best single measure (AUC 0.87) for predicting PSG diagnosed OSA. Photographic variables yielded AUCs of 0.69 or below, with the best predictor, nose width, included in multivariable models. Candidate multivariable logistic regression models comprising various combinations of predictor variables are shown in Table 3. The addition of age, male, BMI, and symptoms slightly increases the AUC observed with the flow monitor alone (table 3 and figure 2, model C). Using the same threshold with Model C, sensitivity was 0.99, specificity 0.61, PPV 0.42 and NPV 0.99 for a PSG AHI of ≥30 (Figure 3).

Validation population

As Models C and D were similar, but Model C contained fewer variables that are easier to collect in primary care, Model C was chosen to be evaluated with the validation population, comprising n=101 with complete data for the six variables, and yielded an AUC of 0.88. Using the same threshold obtained from the development population, the model showed a reduced sensitivity of 0.71 (95% CI 0.59 – 0.81), specificity of 0.80 (0.63 – 0.91), positive predictive value (PPV) of 0.87 (0.75 – 0.94), a negative predictive value (NPV) of 0.60 (0.44 – 0.73), positive likelihood ratio of 3.6 (1.8 – 7.0), negative likelihood ratio of 0.36 (0.24 – 0.54), and diagnostic odds ratio of 9.9 (3.7- 26.5). The model therefore cannot reliably rule out the presence of OSA in this high prevalence population, however those determined to be low risk are unlikely to have severe disease, with only 4% having severe OSA. With a PSG AHI of ≥30 as the reference, using the same threshold with Model C, sensitivity was 0.92, specificity 0.60, PPV 0.44 and NPV 0.96.

DISCUSSION:

This is the first study to develop a diagnostic model for OSA which considers the clinical impressions of the PCP, closely emulating a routine clinical referral pathway. A single

channel device, along with demographic and clinical characteristics easily obtained during a primary care consultation, can predict the presence of in-laboratory PSG determined OSA. After basic training, when PCPs refer patients on clinical suspicion of OSA, they are likely to have the disease (58% with AHI≥10). Given this high pre-test probability of OSA, when basic clinical and demographic information (age, gender, BMI, questionnaire for breathing symptoms) is added, 74% of those considered high risk will actually have OSA, but 36% considered low risk will have OSA. With the addition of a simple, single channel at-home nasal flow monitor, 85% of patients considered high-risk will have OSA, but 22% of those with a low risk result on flow monitor may still have mild OSA. Our results show that a simple flow monitor used at home can enhance PCP assessment in the evaluation of OSA. Established patient questionnaires, anthropometrics and craniofacial photography do not add value to the assessment of OSA in this setting. The ability of our model to predict OSA was largely maintained between the development and validation phases of the study.

Previous studies using limited channel devices have also found that simple testing may rule-in OSA given a high pre-test probability but may have limited capacity to rule-out milder disease.(31) It is unclear whether the addition of these simple screening tools would add anything further where the pre-test probability is even higher than our sample, for example where 84% of those referred for sleep investigations had an AHI≥10.(25) Therefore our study adds value to this and other prior research which called for studies taking into account different pre-test probabilities of OSA within different patient populations including primary care settings.(25, 32) Severe OSA (AHI≥30) can be confidently ruled out using our model that combines basic clinical and demographic information with the nasal flow monitor (NPV 96%, Figure 3), which is important as it has the highest potential to benefit from treatment. Recent suggestions that 1 billion people globally have OSA(4) emphasise the need focus diagnosis and treatment on severe disease.(33) It may be that with clinical information and

basic at-home testing, severe disease could be ruled out by a PCP alone, hence reducing the number of PSGs required. In addition, we have shown that our model is able to predict milder OSA (AHI≥10) which has been omitted from most previous studies.(12) Milder OSA may still be associated with negative outcomes including daytime sleepiness and its consequences and patients benefit from CPAP treatment.(34, 35) Accurate prediction of a range of OSA severity is also important because of increasing attention regarding employment, life insurance and driver licensing in patients.(36)

Interestingly, without direct measurement of airflow, basic clinical assessments and questionnaires did not increase predictive value much more than the PCPs' clinical impressions. This is important due to recent changes in medical regulations for sleep apnoea testing in some countries e.g. in Australia, requiring that PCPs wishing to refer patients directly for home sleep testing must first demonstrate a patient's positive result on questionnaire (STOP-BANG, OSA-50 or Berlin). In a post-hoc analysis (Supplement Table 3a), we estimated the STOP-BANG and OSA-50 scores, and found that 22-39% of those classified as low probability of OSA could still have moderate-severe OSA, on the basis of negative predictive values of 0.61-0.78 for AHI≥15. Our data has shown that when screening questionnaires are given after the PCP has decided the patient requires testing, further evaluation may still be indicated even if the screening questionnaire indicates low risk, and in the Australian context referral to a specialist sleep physician should be considered.

Assessment of facial morphology by craniofacial photography did not improve the ability of clinical assessments and the nasal flow monitor to identify disease in this primary care context. Previous research found that visual assessment of photographs was not useful in the diagnosis of OSA,(37) but did not use the validated photographic assessment technique as used in our study.(20) There is evidence to show that craniofacial measurements relate to

OSA differently across different ethnic groups,(38) and this adds complexity as a screening tool.

Other studies have developed OSA prediction models based on screening the general population,(19, 39) based on screening all patients coming through primary care practice,(12, 40) or in a sleep centre.(15, 17, 20, 25) None of these truly represent the most common referral pathway for those with OSA which depends on PCPs' clinical judgement. In this study, the PCPs were instructed to refer those who they thought may need further sleep investigations. Therefore, our predictive model is the first to have been developed taking this clinical impression into account making it relevant to how a simplified diagnostic model could be utilised in the primary care practice setting.

Recently the US Taskforce for Preventative Services recommended against screening for OSA in the general population,(41) meaning that an approach like ours is likely to continue to be a common referral pathway for OSA. Our focus was on developing a validated model for primary care diagnosis of OSA. While there have been recent studies that have shown that once diagnosed, OSA can be managed optimally in primary care settings,(13, 40, 42, 43) these studies did not attempt to provide improved diagnosis for OSA in this setting.

Some limitations do need to be acknowledged. The a priori power calculation required the analysis of data from 575 participants, but due to screening failures, dropouts and data loss the study may be underpowered. Additionally, the power calculation underestimated the prevalence of OSA in our study population as 20%, rather than 58%. Although we have defined the presence of OSA by the AHI, this metric is a composite index that may include central as well as obstructive respiratory events.

There may be patients with OSA who the PCPs missed so we do not know the prevalence of OSA in those who are not referred for testing. As PCPs were not randomly selected, but responded to advertising about the study, they may have had a special interest in sleep

disorders, thus introducing performance bias. However, all participating PCPs had relatively minimal training compared to previous studies in primary care.(13, 40) Additionally, their clinical judgement may have improved over the course of the two-year trial, after receiving results from previously recruited patients. This may lead to a difference in the spectrum of patients between the development and validation datasets. Additionally, as some PCPs in our study referred a larger number of the patients, and our data may have been influenced by these PCPs who had a particular interest in OSA (See Supplemental Table 1). Finally, our post-hoc analysis examining the newer OSA predictive questionnaires used estimated responses based on the other collected questionnaire responses and were not collected prospectively.

These models have shown that OSA may be identified in primary care, but this model needs to be assessed first in terms of its ability to improve patient outcomes through referral to treatment and secondly with cost-effectiveness analysis. Future studies should focus on the implementation of such a model in a study to analyse the decision to treat OSA, which can be managed wholly within primary care in many clinical situations.

There are three major findings to arise from this study. Firstly, with limited training, PCPs can identify a population at high risk of OSA purely on clinical impression. Secondly, basic clinical data and questionnaires perform poorly when trying to rule-out OSA. Finally, a direct measure of airflow or a validated surrogate of this is required to rule-in OSA or rule-out severe disease, and hence diagnosis is unlikely to be possible without the use of a device. Future research will need to focus on the implementation of this simple diagnostic model and its effect on patient and economic outcomes.

ACKNOWLEDGEMENTS: *(if applicable)*
The authors were supported by the following funding from the Australian National Health and Medical Research Council (NHMRC): C. Hoyos- NHMRC-ARC Dementia Research Development Fellowship (1104003), G. Marks- Research Fellowship (1060614), R. Grunstein- Investigator Grant (1197439), Senior Principal Research Fellowship (1106974). And the NHMRC Centre of Clinical Research Excellence in Respiratory and Sleep Medicine (264598), the Centre of Clinical Research Excellence in Interdisciplinary Sleep Health, CIRUS (571421) and the Centre of Research Excellence NeuroSLEEP: The Centre for Translational Sleep and Circadian Neurobiology (1060992).

DISCLOSURE STATEMENT: *(if applicable)*
Supported by a NHMRC Project Grant: No 512497. KW had full access to all the data in the study and had final responsibility for the decision to submit for publication.

TABLES: *include Tables in the main document, one Table per page.*

Table 1: Participant characteristics

	n	Development population	n	Validation population
Age, years	315	48 (13)	114	49 (12)
Men, n (%)	315	200 (63%)	114	76 (67%)
BMI, kg/m2	315	30.3 (5.9)	114	30.0 (6.2)
Neck circumference, mm	313	396 (43)	114	397 (41)
Waist circumference, cm	313	101 (15)	114	102 (18)
Hip circumference, cm	313	110 (13)	114	108 (13)
ESS total score /24	300	9.7 (4.9)	106	9.6 (4.6)
MAPI score	294	0.57 (0.22)	101	0.58 (0.20)
Berlin high risk	315	261 (83%)	114	87 (76%)
Craniofacial photography, Nose width, cm	302	3.95 (0.41)	105	3.91 (0.38)
Flow monitor REI, events/hr	315	26 (17)	114	27 (18)
Total AHI, events/hr	315	19.3 (19.5)	114	22.1 (20.8)
OSA severity	315		114	
no OSA, AHI<5/h, n (%)		75 (24%)		22 (19%)
mild, AHI 5-15/h, n (%)		98 (31%)		35 (31%)
moderate, AHI 15-30/h n (%)		74 (23%)		28 (25%)
severe, AHI>30/h, n (%)		68 (22%)		29 (25%)
Smoking status	315		114	
Current smoker, n (%)		38 (12%)		18 (16%)
Ex-smoker, n (%)		89 (28%)		39 (34%)
Non-smoker, n (%)		188 (60%)		57 (50%)

Values are mean (SD) or n (%).

Table 2: Diagnostic characteristics of individual testing components with respect to PSG AHI≥10

	AUC	Sensitivity	Specificity	PPV	NPV
Male Gender	0.60 (0.55 – 0.65)	0.72 (0.65 – 0.78)	0.48 (0.39 – 0.57)	0.65 (0.57 – 0.71)	0.57 (0.47 – 0.66)
BMI, ≥ 29.7§	0.62 (0.55 – 0.68)	0.57 (0.49 – 0.64)	0.63 (0.54 – 0.71)	0.67 (0.59 – 0.74)	0.53 (0.45 – 0.60)
Age, ≥ 53	0.62 (0.56 – 0.68)	0.58 (0.51 – 0.65)	0.62 (0.54 – 0.71)	0.67 (0.59 – 0.74)	0.53 (0.45 – 0.61)
Symptoms (Index 1 of MAP), ≥ 2.5	0.60 (0.54 – 0.67)	0.60 (0.52 – 0.67)	0.58 (0.49 – 0.66)	0.65 (0.57 – 0.72)	0.53 (0.44 – 0.61)
MAPI, ≥ 0.6	0.73 (0.67 – 0.79)	0.69 (0.61 – 0.76)	0.67 (0.58 – 0.75)	0.73 (0.65 – 0.80)	0.62 (0.54 – 0.70)
Berlin Questionnaire, high risk	0.53 (0.49 – 0.57)	0.86 (0.79 – 0.90)	0.21 (0.14 – 0.29)	0.59 (0.52 – 0.65)	0.52 (0.38 – 0.66)
Estimated OSA50 score*, ≥ 5	0.62 (0.55 – 0.68)	0.93 (0.87 – 0.96)	0.23 (0.16 – 0.32)	0.62 (0.55 – 0.68)	0.70 (0.53 – 0.84)
Estimated STOPBANG score*, high risk	0.70 (0.65 – 0.76)	0.40 (0.33 – 0.48)	0.78 (0.70 – 0.84)	0.70 (0.60 – 0.79)	0.50 (0.43 – 0.57)
Estimated NoSAS score*, ≥ 9	0.66 (0.60 – 0.72)	0.75 (0.68 – 0.81)	0.52 (0.43 – 0.60)	0.67 (0.60 – 0.73)	0.61 (0.52 – 0.70)
Mallampati score, ≥ 4	0.57 (0.51 – 0.63)	0.54 (0.46 – 0.61)	0.59 (0.50 – 0.67)	0.63 (0.55 – 0.71)	0.49 (0.41 – 0.57)
Flow monitor REI, ≥ 18	0.87 (0.83 – 0.91)	0.87 (0.81 – 0.91)	0.77 (0.69 – 0.84)	0.83 (0.77 – 0.88)	0.81 (0.73 – 0.88)
Photography (nose width)**, ≥ 3.9	0.69 (0.63 - 0.75)	0.65 (0.57 – 0.72)	0.69 (0.60 – 0.77)	0.74 (0.66 – 0.80)	0.60 (0.52 – 0.68)

AUC: area under the curve, BMI: Body Mass Index, NPV: negative predictive value; PPV: positive predictive value, MAPI: multivariable apnea prediction index.

§Thresholds used to calculate diagnostic test statistics.

*OSA50, STOPBANG, NoSAS scores estimated from responses given to other questionnaires.

** Nose width (photographic variable L68) was chosen via CART model as the most predictive

Table 3: Exploration of alternative logistics regression models with respect to polysomnography AHI≥10.

Model	Variables	n	AUC	Sensitivity	Specificity	PPV	NPV
A	Age Male Gender BMI Symptoms	294	0.74 (0.68 – 0.80)	0.78 (0.71 – 0.84)	0.60 (0.51 – 0.69)	0.72 (0.65 – 0.78)	0.68 (0.59 – 0.76)
B	Age Male Gender BMI Symptoms Photography	282	0.76 (0.70 – 0.81)	0.78 (0.71 – 0.84)	0.62 (0.52 – 0.70	0.73 (0.65 – 0.79)	0.68 (0.59 - 0.77)
C	Age Male Gender BMI Symptoms Flow Monitor	294	0.89 (0.85 – 0.93)	0.80 (0.73 – 0.56)	0.84 (0.77 – 0.90)	0.87 (0.80 – 0.92)	0.77 (0.69 – 0.83)
D	Age Male Gender BMI Symptoms Flow Monitor Photography	282	0.89 (0.85 – 0.92)	0.84 (0.78 – 0.89)	0.81 (0.73 - 0.87)	0.85 (0.79 – 0.90)	0.80 (0.72 – 0.87)

AUC: area under the curve, BMI: Body Mass Index. NPV: negative predictive value; PPV: positive predictive value

FIGURE LEGEND: *The description in the legend should be sufficient for the reader to interpret the figure without reference to the text.*
Figure 1. Participant flow diagram

Figure 2. Receiver operating characteristic plot for Model C with respect to the development population (blue solid line) and validation population (red dashed line).

Figure 3. Model C AUC (solid line), PPV (blue dashed line) and NPV (red dotted line) at four thresholds (5, 10, 15, 30) for the PSG AHI, in the development population.
Legend: AUC=Area under the curve, PPV=positive predictive value, NPV=negative predictive value, PSG=polysomnography, AHI=apnoea hypopnoea index.

REFERENCES

1. Bearpark H, Elliott L, Grunstein R, Cullen S, Schneider H, Althaus W, et al. Snoring and sleep apnea. A population study in Australian men. Am J Respir Crit Care Med. 1995;151(5):1459-65.
2. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med. 1993;328(17):1230-5.
3. Heinzer R, Vat S, Marques-Vidal P, Marti-Soler H, Andries D, Tobback N, et al. Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study. The Lancet Respiratory medicine. 2015;3(4):310-8.
4. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MSM, Morrell MJ, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. The Lancet Respiratory medicine. 2019;7(8):687-98.
5. Kendzerska T, Mollayeva T, Gershon AS, Leung RS, Hawker G, Tomlinson G. Untreated obstructive sleep apnea and the risk for serious long-term adverse outcomes: A systematic review. Sleep Med Rev. 2014;18(1):49-59.
6. Douglas JA, Chai-Coetzer CL, McEvoy D, Naughton MT, Neill AM, Rochford P, et al. Guidelines for sleep studies in adults - a position statement of the Australasian Sleep Association. Sleep Med. 2017;36 Suppl 1:S2-S22.
7. Qaseem A, Dallas P, Owens DK, Starkey M, Holty JE, Shekelle P, et al. Diagnosis of obstructive sleep apnea in adults: a clinical practice guideline from theAmerican College of Physicians. Ann Intern Med. 2014;161(3):210-20.
8. Flemons WW, Douglas NJ, Kuna ST, Rodenstein DO, Wheatley J. Access to diagnosis and treatment of patients with suspected sleep apnea. Am J Respir Crit Care Med. 2004;169(6):668-72.
9. Statistics M. Medicare Statistics. medicarestatisticshumanservicesgovau.
10. Hanes CA, Wong KKH, Saini B. Diagnostic pathways for obstructive sleep apnoea in the Australian community: observations from pharmacy-based CPAP providers. Sleep and Breathing. 2015;19(4):1241-8.
11. Netzer NC, Hoegel JJ, Loube D, Netzer CM, Hay B, Alvarez-Sala R, et al. Prevalence of symptoms and risk of sleep apnea in primary care. Chest. 2003;124(4):1406-14.
12. Chai-Coetzer CL, Antic NA, Rowland LS, Catcheside PG, Esterman A, Reed RL, et al. A simplified model of screening questionnaire and home monitoring for obstructive sleep apnoea in primary care. Thorax. 2011;66(3):213-9.
13. Chai-Coetzer CL, Antic NA, Rowland LS, Reed RL, Esterman A, Catcheside PG, et al. Primary care vs specialist sleep center management of obstructive sleep apnea and daytime sleepiness and quality of life: a randomized trial. JAMA. 2013;309(10):997-1004.
14. Chai-Coetzer CL, Antic NA, McEvoy RD. Identifying and managing sleep disorders in primary care. The Lancet Respiratory medicine. 2015;3(5):337-9.
15. Maislin G, Pack AI, Kribbs NB, Smith PL, Schwartz AR, Kline LR, et al. A survey screen for prediction of apnea. Sleep. 1995;18(3):158-66.
16. Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. Ann Intern Med. 1999;131(7):485-91.
17. Rowley JA, Aboussouan LS, Badr MS. The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. Sleep. 2000;23(7):929-38.
18. Pang KP, Terris DJ. Screening for obstructive sleep apnea: an evidence-based analysis. Am J Otolaryngol. 2006;27(2):112-8.
19. Marti-Soler H, Hirotsu C, Marques-Vidal P, Vollenweider P, Waeber G, Preisig M, et al. The NoSAS score for screening of sleep-disordered breathing: a derivation and validation study. The Lancet Respiratory medicine. 2016;4(9):742-8.

20. Lee RW, Chan AS, Grunstein RR, Cistulli PA. Craniofacial phenotyping in obstructive sleep apnea--a novel quantitative photographic approach. Sleep. 2009;32(1):37-45.
21. Ross SD, Sheinhait IA, Harrison KJ, Kvasz M, Connelly JE, Shea SA, et al. Systematic review and meta-analysis of the literature regarding the diagnosis of sleep apnea. Sleep. 2000;23(4):519-32.
22. Rofail LM, Wong KK, Unger G, Marks GB, Grunstein RR. The utility of single-channel nasal airflow pressure transducer in the diagnosis of OSA at home. Sleep. 2010;33(8):1097-105.
23. Flemons WW, Littner MR, Rowley JA, Gay P, Anderson WM, Hudgel DW, et al. Home diagnosis of sleep apnea: a systematic review of the literature. An evidence review cosponsored by the American Academy of Sleep Medicine, the American College of Chest Physicians, and the American Thoracic Society. Chest. 2003;124(4):1543-79.
24. Chesson AL, Jr., Berry RB, Pack A, American Academy of Sleep M, American Thoracic S, American College of Chest P. Practice parameters for the use of portable monitoring devices in the investigation of suspected obstructive sleep apnea in adults. Sleep. 2003;26(7):907-13.
25. Masa JF, Duran-Cantolla J, Capote F, Cabello M, Abad J, Garcia-Rio F, et al. Effectiveness of home single-channel nasal pressure for sleep apnea diagnosis. Sleep. 2014;37(12):1953-61.
26. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. 1991;14(6):540-5.
27. Wong KK, Jankelson D, Reid A, Unger G, Dungan G, Hedner JA, et al. Diagnostic test evaluation of a nasal flow monitor for obstructive sleep apnea detection in sleep apnea research. Behav Res Methods. 2008;40(1):360-6.
28. EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. Sleep. 1992;15(2):173-84.
29. 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(5):667-89.
30. Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. Anesthesiology. 2008;108(5):812-21.
31. Collop NA, Anderson WM, Boehlecke B, Claman D, Goldberg R, Gottlieb DJ, et al. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. J Clin Sleep Med. 2007;3(7):737-47.
32. Kunisaki KM, Greer N, Khalil W, Koffel E, Koeller E, MacDonald R, et al. Provider Types and Outcomes in Obstructive Sleep Apnea Case Finding and Treatment: A Systematic Review. Ann Intern Med. 2018;168(3):195-202.
33. McNicholas WT, Bonsignore MR, Levy P, Ryan S. Mild obstructive sleep apnoea: clinical relevance and approaches to management. The Lancet Respiratory medicine. 2016;4(10):826-34.
34. Chowdhuri S, Quan SF, Almeida F, Ayappa I, Batool-Anwar S, Budhiraja R, et al. An Official American Thoracic Society Research Statement: Impact of Mild Obstructive Sleep Apnea in Adults. Am J Respir Crit Care Med. 2016;193(9):e37-54.
35. McNicholas WT. Active management of mild obstructive sleep apnoea: the evidence grows. The Lancet Respiratory medicine. 2020;8(4):322-3.

36. Bonsignore MR, Randerath W, Riha R, Smyth D, Gratziou C, Goncalves M, et al. New rules on driver licensing for patients with obstructive sleep apnoea: EU Directive 2014/85/EU. Eur Respir J. 2016;47(1):39-41.
37. Cheung K, Ishman SL, Benke JR, Collop N, Tron L, Moy N, et al. Prediction of obstructive sleep apnea using visual photographic analysis. J Clin Anesth. 2016;32:40-6.
38. Sutherland K, Lee RWW, Chan TO, Ng S, Hui DS, Cistulli PA. Craniofacial Phenotyping in Chinese and Caucasian Patients With Sleep Apnea: Influence of Ethnicity and Sex. J Clin Sleep Med. 2018;14(7):1143-51.
39. Senaratna CV, Perret JL, Lowe A, Bowatte G, Abramson MJ, Thompson B, et al. Detecting sleep apnoea syndrome in primary care with screening questionnaires and the Epworth sleepiness scale. Med J Aust. 2019.
40. Peñacoba P, Llauger MA, Fortuna AM, Flor X, Sampol G, Pedro-Pijoan AM, et al. Primary care and sleep unit agreement in management decisions for sleep apnea: a prospective study in Spain. J Clin Sleep Med. 2020.
41. Force USPST. Screening for Obstructive Sleep Apnea in Adults: US Preventive Services Task Force Recommendation Statement. JAMA. 2017;317(4):407-14.
42. Sanchez-Quiroga MA, Corral J, Gomez-de-Terreros FJ, Carmona-Bernal C, Asensio-Cruz MI, Cabello M, et al. Primary Care Physicians Can Comprehensively Manage Sleep Apnea Patients: A Non-inferiority Randomized Controlled Trial. Am J Respir Crit Care Med. 2018.
43. Tarraubella N, Sanchez-de-la-Torre M, Nadal N, De Batlle J, Benitez I, Cortijo A, et al. Management of obstructive sleep apnoea in a primary care vs sleep unit setting: a randomised controlled trial. Thorax. 2018;73(12):1152-60.

Supplementary Material to *Development and validation of a model for diagnosis of obstructive sleep apnoea in primary care.*

Methods

Recruitment of primary care practitioners and patients

We invited Primary Care Practitioners (PCPs) to be part of our study by advertising in local medical publications. Responding PCPs were trained both on study procedures and on recognising the clinical presentation of OSA (two 2.5 hour sessions). Training comprised of a combination of reading materials and interactive face-to-face discussion that made them eligible to accrue continuing professional development points with their professional college. Following the training, PCPs were asked to invite any patients who presented to them with suspected or possible OSA to participate in the study. A subset of PCPs (n=11) agreed to additional training in craniofacial photography and were given a camera to use in their clinic rooms. PCPs received a monetary reimbursement for the paperwork and data collection involved with each patient referral ($100 AUD), and for each patient photographed ($30 AUD).

Procedures and materials

A questionnaire booklet comprising items to be completed by the patient, as well as their PCP was completed. The consent and questionnaire booklet were then posted to research staff at the main site (Woolcock Institute), who coordinated testing with an in-laboratory PSG and home sleep testing with a nasal flow monitor. The order of testing was randomised. All patients underwent a simple craniofacial photography procedure at their PSG visit, and a sub-group of patients also had photography at the initial PCP visit. PSG results were sent to the referring PCP after all tests were completed, and a follow-up appointment with a sleep physician was booked for all patients to discuss PSG results and subsequent management.

Home sleep testing was performed using a nasal flow monitor (Flow Wizard©, DiagnoseIT, Sydney, Australia), an Australian Therapeutic Goods Administration-approved, CE-marked device that records airflow by nasal pressure cannula.(1) All patients were posted a device with instructions and asked to wear it during sleep for three consecutive nights). They were asked to return the device via a provided postage-paid satchel. On return of the device, data was downloaded; producing an automatically scored respiratory event index (REI) and a report on signal quality. The mean of the individual night REIs was used in the analysis.

All patients underwent an in-laboratory PSG as the reference standard, in one of five participating sleep centres. Scoring was done manually by Registered Sleep Technologists, according to standard criteria(2, 3). An obstructive apnea was defined as a complete cessation of breathing as shown in the nasal flow signal for at least 10 seconds, accompanied by effort shown in the thoracic or abdominal effort bands or diaphragm EMG. A hypopnea was defined as a clear amplitude reduction of airflow for 10 seconds or more accompanied by an arousal or 3% desaturation, where the baseline was the mean amplitude of stable breathing and oxygenation in the two minutes preceding onset of the event (in individuals who have a stable breathing pattern during sleep) or the mean amplitude of the three largest breaths in the two minutes preceding onset of the event). A central apnea was defined as a complete cessation of breathing in the absence of any effort as recorded by thoracic or abdominal bands, or EMG diaphragm for 10 seconds or longer. and blinded to the results of the home nasal flow monitor. PSG scorers were blinded to the results of the home nasal flow monitor. The diagnosis of OSA was made based on an apnoea hypopnoea index (AHI) of ≥10 events per hour.

Supplementary Material to *Development and validation of a model for diagnosis of obstructive sleep apnoea in primary care.*

Craniofacial phenotyping was undertaken using a previously developed digital photography technique (4). All patients had frontal and right profile digital photographs taken when they attended the sleep laboratory for their PSG and, in a subset of patients, additionally at the initial GP visit. The photographs were rated for image quality on a subjective 1-5 scale. Photograph analysis involved manual marking of craniofacial landmarks on the digital images. Craniofacial geometry was calculated using a custom-spreadsheet producing quantitative facial measurements.(5) In cases where a sleep laboratory and GP photograph was available, the image with the higher rated quality was used, or in the case of a tie, the sleep laboratory photograph was used in the analysis. Candidate craniofacial variables were chosen with the assistance of classification and regression tree (CART) models.(4)

Statistical methods

Firstly, univariate logistic regression was used to determine the association between individual variables and the presence of OSA as defined by AHI≥10 on the reference standard test (PSG). No transformation or categorisation was conducted on continuous predictor variables. Secondly multivariable logistic regression models were built manually, using variables chosen from the univariate analyses based on the area under the curve (AUC) of the receiver operating characteristic (ROC), and prior knowledge of OSA risk factors, to develop a predictive model for diagnosis of OSA. Diagnostic test statistics (sensitivity, specificity, positive and negative predictive values) were calculated at the threshold closest to the top left corner of the ROC curve. The optimal model and threshold chosen from the development population was then tested on the validation population. Diagnostic test statistics for the model, using alternative PSG AHI thresholds (5, 15, 30) were also reported graphically. The overall dataset was restricted to participants with complete AHI and Flow Wizard data. Individual models used complete data for the variables they included. We did not perform imputation for missing data. A post hoc analysis was also included using data already collected to derive estimated scores for questionnaires published after commencement of this study (STOP-BANG,(6) OSA50,(7) NoSAS(8)). These estimated questionnaires were used for comparison alongside data from variables and questionnaire scores collected directly in this study (Table 2) but were not used to build the OSA predictive model.

The development population required 360 participants to achieve statistical power of 0.8 at a two-tailed alpha of 0.05 given an odds ratio of 2.0 assuming a correlation (r^2) between variables of 0.2. For the validation population we required 210 participants, to ensure the upper 95% confidence limit of the negative likelihood ratio was at least 1/3 and the lower confidence limit of the positive likelihood ratio was at least 3, assuming a prevalence of 20%.

Supplementary Material to *Development and validation of a model for diagnosis of obstructive sleep apnoea in primary care.*

Results/Discussion

Supplementary Table 1: List of General Practitioners and number of participants they referred to study

GP Code	Number referred	Percentage rounded
2	117	20
5	99	17
11	46	8
23	40	7
7	28	5
15	28	5
22	21	4
20	21	4
9	20	3
19	20	3
8	18	3
17	15	3
10	14	2
16	13	2
12	13	2
29	12	2
6	11	2
24	10	2
3	9	2
27	7	1
31	5	1
28	5	1
14	1	0.2
1	1	0.2
18	1	0.2
	574	100

Supplementary Material to *Development and validation of a model for diagnosis of obstructive sleep apnoea in primary care.*

Supplementary Table 2: Exploration of alternative logistics regression models with respect to polysomnography AHI≥10.

Model	Reference diagnosis	n	AUC	Sensitivity	Specificity	PPV	NPV
C	AHI >=10	101	0.88 (0.82 – 0.95)	0.71 (0.59 – 0.81)	0.80 (0.63 – 0.91)	0.87 (0.75-0.94)	0.60 (0.44 - 0.73)
C	AHI >= 15	101	0.88 (0.82 – 0.94)	0.80 (0.67 – 0.90)	0.74 (0.59 – 0.85)	0.76 (0.62 – 0.86)	0.79 (0.64 – 0.89)
C	AHI >= 30	101	0.92 (0.86 – 0.99)	0.92 (0.73 – 0.99)	0.60 (0.48 – 0.71)	0.44 (0.31 – 0.59)	0.96 (0.84 – 0.99)

AUC: area under the curve, BMI: Body Mass Index. NPV: negative predictive value; PPV: positive predictive value

Supplementary Table 3a: Diagnostic characteristics of OSA screening questionnaires with respect to PSG AHI≥15

	AUC	Sensitivity	Specificity	PPV	NPV
Berlin Questionnaire, high risk	0.55 (0.51 – 0.59)	0.88 (0.81 – 0.93)	0.21 (0.16 – 0.28)	0.48 (0.42 – 0.28)	0.69 (0.54 – 0.80)
Estimated OSA50 score*, ≥ 5	0.62 (0.55 – 0.69)	0.94 (0.87 – 0.97)	0.20 (0.14 – 0.28)	0.50 (0.44 – 0.56)	0.78 (0.0.61 – 0.90)
Estimated STOPBANG score*, high risk	0.71 (0.56 – 0.76)	0.41 (0.33 – 0.49)	0.75 (0.67 – 0.81)	0.56 (0.46 – 0.66)	0.61 (0.54 – 0.67)
Estimated NoSAS score*, ≥ 9	0.68 (0.62 – 0.74)	0.80 (0.72 – 0.86)	0.50 (0.42 – 0.57)	0.57 (0.49 – 0.63)	0.75 (0.66 – 0.83)

AUC: area under the curve, NPV: negative predictive value; PPV: positive predictive value.

*OSA50, STOPBANG, NoSAS scores estimated from responses given to other questionnaires.

Supplementary Material to *Development and validation of a model for diagnosis of obstructive sleep apnoea in primary care.*

Supplementary Table 3b: Diagnostic characteristics of OSA screening questionnaires with respect to PSG AHI≥30

	AUC	Sensitivity	Specificity	PPV	NPV
Berlin Questionnaire, high risk	0.58 (0.55 – 0.62)	0.96 (0.87 – 0.99)	0.21 (0.16 – 0.26)	0.25 (0.20 – 0.31)	0.94 (0.84 – 0.99)
Estimated OSA50 score*, ≥ 5	0.65 (0.57 – 0.73)	0.93 (0.82 – 0-.98)	0.16 (0.11 – 0.21)	0.23 (0.18 – 0.29)	0.89 (0.74 – 0.97)
Estimated STOPBANG score*, high risk	0.73 (0.66 – 0.80)	0.53 (0.40 – 0.65)	0.73 (0.67 – 0.79)	0.34 (0.26 – 0.45)	0.85 (0.80 – 0.90)
Estimated NoSAS score*, ≥ 9	0.70 (0.63 – 0.77)	0.85 (0.74 – 0.92)	0.42 (0.36 – 0.49)	0.29 (0.23 – 0.35)	0.91 (0.84 – 0.96)

AUC: area under the curve, NPV: negative predictive value; PPV: positive predictive value.

*OSA50, STOPBANG, NoSAS scores estimated from responses given to other questionnaires.

References

1. Wong KK, Jankelson D, Reid A, Unger G, Dungan G, Hedner JA, et al. Diagnostic test evaluation of a nasal flow monitor for obstructive sleep apnea detection in sleep apnea research. Behav Res Methods. 2008;40(1):360-6.
2. EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association. Sleep. 1992;15(2):173-84.
3. 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(5):667-89.
4. Lee RW, Chan AS, Grunstein RR, Cistulli PA. Craniofacial phenotyping in obstructive sleep apnea--a novel quantitative photographic approach. Sleep. 2009;32(1):37-45.
5. Maislin G, Pack AI, Kribbs NB, Smith PL, Schwartz AR, Kline LR, et al. A survey screen for prediction of apnea. Sleep. 1995;18(3):158-66.
6. Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. Anesthesiology. 2008;108(5):812-21.
7. Chai-Coetzer CL, Antic NA, Rowland LS, Catcheside PG, Esterman A, Reed RL, et al. A simplified model of screening questionnaire and home monitoring for obstructive sleep apnoea in primary care. Thorax. 2011;66(3):213-9.
8. Marti-Soler H, Hirotsu C, Marques-Vidal P, Vollenweider P, Waeber G, Preisig M, et al. The NoSAS score for screening of sleep-disordered breathing: a derivation and validation study. The Lancet Respiratory medicine. 2016;4(9):742-8.