

Cost Minimization Using an Artificial Neural Network Sleep Apnea Prediction Tool for Sleep Studies

Rahel A. Teferra¹, Brydon J. B. Grant^{2,3}, Jesse W. Mindel¹, Tauseef A. Siddiqi¹, Imran H. Iftikhar^{1,4}, Fatima Ajaz¹, Jose P. Aliling¹, Meena S. Khan¹, Stephen P. Hoffmann¹, and Ulysses J. Magalang¹

¹Division of Pulmonary, Allergy, Critical Care, and Sleep Medicine, Ohio State University Wexner Medical Center, Columbus, Ohio; ²Division of Pulmonary, Critical Care, and Sleep Medicine, University at Buffalo, State University of New York, Buffalo, New York; ³Department of Medicine, Veteran Affairs Medical Center, Fayetteville, North Carolina; and ⁴Division of Pulmonary, Critical Care, and Sleep Medicine, University of South Carolina School of Medicine, Columbia, South Carolina

Abstract

Rationale: More than a million polysomnograms (PSGs) are performed annually in the United States to diagnose obstructive sleep apnea (OSA). Third-party payers now advocate a home sleep test (HST), rather than an in-laboratory PSG, as the diagnostic study for OSA regardless of clinical probability, but the economic benefit of this approach is not known.

Objectives: We determined the diagnostic performance of OSA prediction tools including the newly developed OSUNet, based on an artificial neural network, and performed a cost-minimization analysis when the prediction tools are used to identify patients who should undergo HST.

Methods: The OSUNet was trained to predict the presence of OSA in a derivation group of patients who underwent an in-laboratory PSG (n = 383). Validation group 1 consisted of in-laboratory PSG patients (n = 149). The network was trained further in 33 patients who underwent HST and then was validated in a separate group of 100 HST patients (validation group 2). Likelihood ratios (LRs) were compared with two previously published prediction tools. The total costs from the use of the three

prediction tools and the third-party approach within a clinical algorithm were compared.

Measurements and Main Results: The OSUNet had a higher +LR in all groups compared with the STOP-BANG and the modified neck circumference (MNC) prediction tools. The +LRs for STOP-BANG, MNC, and OSUNet in validation group 1 were 1.1 (1.0–1.2), 1.3 (1.1–1.5), and 2.1 (1.4–3.1); and in validation group 2 they were 1.4 (1.1–1.7), 1.7 (1.3–2.2), and 3.4 (1.8–6.1), respectively. With an OSA prevalence less than 52%, the use of all three clinical prediction tools resulted in cost savings compared with the third-party approach.

Conclusions: The routine requirement of an HST to diagnose OSA regardless of clinical probability is more costly compared with the use of OSA clinical prediction tools that identify patients who should undergo this procedure when OSA is expected to be present in less than half of the population. With OSA prevalence less than 40%, the OSUNet offers the greatest savings, which are substantial when the number of sleep studies done annually is considered.

Keywords: cost analysis; neural network models; clinical prediction rule; sleep-disordered breathing

(Received in original form April 16, 2014; accepted in final form July 10, 2014)

Supported by NHLBI HL093463 and the Tzagournis Medical Research Endowment Funds of Ohio State University (both to U.J.M.).

Author Contributions: Study concept and design: R.A.T., B.J.B.G., J.W.M., U.J.M.; acquisition of data: R.A.T., J.W.M., T.A.S., I.H.I., F.A., J.P.A., M.S.K., U.J.M.; analysis and interpretation of data: R.A.T., B.J.B.G., J.W.M., S.P.H., U.J.M.; drafting the manuscript: R.A.T., B.J.B.G., J.W.M., U.J.M.; manuscript revision and other important intellectual content: T.A.S., I.H.I., F.A., J.P.A., M.S.K., S.P.H.

Correspondence and requests for reprints should be to Ulysses J. Magalang, M.D., Division of Pulmonary, Allergy, Critical Care, and Sleep Medicine, Ohio State University Wexner Medical Center, 201 Davis Heart and Lung Research Institute, 473 West 12th Avenue, Columbus, OH 43210. E-mail: magalang.1@osu.edu

This article has an online supplement, which is accessible from this issue's table of contents online at www.atsjournals.org

Ann Am Thorac Soc Vol 11, No 7, pp 1064–1074, Sep 2014

Copyright © 2014 by the American Thoracic Society

DOI: 10.1513/AnnalsATS.201404-161OC

Internet address: www.atsjournals.org

Polysomnography (PSG), the current “gold standard” for defining the presence and severity of obstructive sleep apnea (OSA), is a common medical procedure in the

United States. According to a conservative estimate, about 1.17 million PSGs were performed in the United States in 2001 (1).

A home sleep test (HST) most often involves portable monitoring of respiration during the usual sleep hours; it is unattended by a technologist and without recording

of sleep stages. Many third-party payers now advocate an HST as the initial diagnostic test to diagnose OSA in all patients regardless of clinical probability because of the lower cost of this procedure compared with an in-laboratory PSG (2). This process transformation, implemented through prior authorization for in-laboratory PSGs, has resulted in abrupt closure of a number of sleep centers last year in Boston, Massachusetts, leaving more than 30,000 patients without a sleep medicine provider (3). However, the dollar cost of the third-party policy as compared with other approaches is unclear because most clinicians have a low threshold to proceed with an in-laboratory study in symptomatic patients with a negative HST (4, 5), and because a significant proportion of patients with a negative HST may have significant OSA when studied in the laboratory (6, 7). Duplication of tests in these patients may negate any cost savings derived from the lower cost of the HST (4). When a decision has been made by the clinician to perform a PSG, ideally patients who have a high likelihood of having OSA should undergo an HST (5). Indeed, in a large multicenter study of patients suspected of having OSA, the therapeutic decision performed after the diagnosis by HST was effective only for patients with severe OSA, defined as an apnea-hypopnea index (AHI) greater than 30/hour, suggesting that only those patients predicted to have an increased risk for the condition should undergo the in-home test as the initial diagnostic procedure (8).

The overall clinical impression of sleep medicine physicians of the likelihood of OSA in individual patients has been reported to be of limited usefulness (9, 10). Several clinical prediction tools for the presence of OSA have been proposed to help clinicians in their decision to pursue a sleep study (10–13). However, an economic evaluation of these clinical prediction tools when used in clinical practice, including whether the study should be done at home, has not been previously performed. The use of an artificial neural network (ANN) has several advantages in clinical prediction as it is not affected by problems associated with multicollinearity, it automatically models nonlinear relations, and it detects implicitly all possible interaction terms (14). In particular situations, ANNs have been reported to outperform physician prediction (15). We sought to leverage

these advantages of ANN and hypothesized that a newly developed prediction tool based on ANN will outperform previously published clinical predictions and will offer cost savings when incorporated within a clinical pathway for deciding which patients should be referred for an HST.

Methods

The detailed METHODS are located in the online supplement. Consecutive patients over the age of 18 years, referred for in-laboratory PSG for suspicion of OSA from April 2010 to March 2011 to the Ohio State University (OSU) Sleep Disorders Center, were identified and included in the study. Subjects with total sleep time less than 2 hours, those who required oxygen supplementation during the PSG, and those with incomplete sleep questionnaires or without anthropometric data were excluded from the study. The HST program at our center started at a later time and consecutive patients referred for HST from April 2011 to July 2012 were included in the study, using the same inclusion and exclusion criteria as previously described except for the total sleep time. Medical records and sleep questionnaires that the patients completed as part of their clinical evaluations were reviewed and the data were collected on a spreadsheet. The research protocol was approved by the Institutional Review Board of the Ohio State University Wexner Medical Center (Columbus, OH).

Polysomnography

All in-laboratory PSGs included the recommended standard channels, and each 30-second epoch was manually scored for sleep stage, apneas, and hypopneas (16). The total sleep time was used as the denominator to calculate the AHI for in-laboratory PSGs, and the recording time from patient-reported lights-off to lights-on was used to calculate the AHI for HSTs.

Clinical Prediction Rules

Modified neck circumference. Modified neck circumference (MNC) was measured in centimeters and adjusted if the patient had hypertension (4 cm added), was a habitual snorer (3 cm added), or was reported to choke or gasp most nights (3 cm added). An MNC score greater than 43 cm indicated an increased risk of having OSA, whereas an

MNC not greater than 43 indicated low risk (6, 10, 13).

STOP-BANG. The score for the STOP-BANG was calculated by assigning a score of 1 for each positive answer to Snoring, Tiredness, being Observed to stop breathing during sleep, high blood Pressure, Body mass index (BMI) greater than 35 kg/m², Age greater than 50 years, Neck circumference greater than 40 cm, and male Gender. An increased risk for OSA was defined as three or more affirmative answers to the eight STOP-BANG items, whereas patients with two or fewer affirmative answers were considered to have low risk (11).

OSUNet. A generalized regression neural network based on the cascade correlation algorithm of S. Fahlman (NeuroShell Classifier; Ward Systems Group Inc., Frederick, MD) was used to generate the OSUNet, which produced the desired output of whether the subject is at increased risk for OSA or not (17, 18). The variables chosen as final inputs for the ANN were selected from an initial list of 19 items (see Table E1 in the online supplement) obtained from patients' questionnaire as well as anthropometric measurements using regression analysis with the AHI as the dependent variable. A *P* value not exceeding 0.10 was arbitrarily chosen as the cutoff for the continued inclusion of a given variable in the final ANN inputs (19). The final nine ANN inputs were as follows: age, history of hypertension, history of diabetes mellitus, BMI, neck circumference, and responses to four questions: "I am told I snore in my sleep," "I am told I stop breathing in my sleep," "I have or have been told that I have restless legs," and "My desire or interest in sex is less than it used to be." The four questions had a six-item Likert response. Missing values and responses marked "not applicable" were substituted with the class mean (20).

The ANN was initially trained to predict the presence or absence of OSA in a derivation group (*n* = 383) of patients who underwent an in-laboratory PSG. The ANN was then validated in another group of consecutive patients (validation group 1) who also underwent an in-laboratory PSG (*n* = 149). The ANN was trained again, using 33 patients who underwent in-home PSG. The resulting network was used to predict the presence of OSA in 100 consecutive patients who underwent in-home PSG (validation group 2).

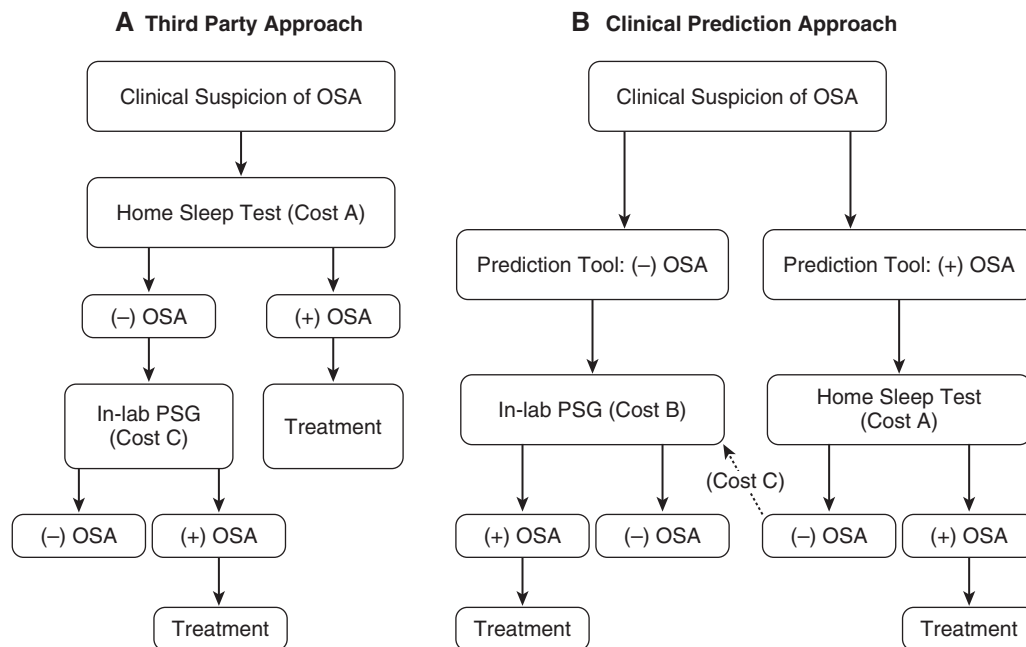


Figure 1. Clinical algorithm for sleep studies. (A) The total cost of the third-party approach wherein all patients undergo a home sleep test (HST) as the initial procedure and (B) the costs of the approaches using the three prediction tools were calculated on the basis of the algorithm shown. In (B), patients identified as having an increased risk of having obstructive sleep apnea (OSA) by the prediction tool go on to have an HST as the initial diagnostic procedure whereas those without an increased risk have an in-laboratory polysomnogram (PSG) as the initial diagnostic procedure.

Cost-Minimization Analysis

The total cost of the third-party approach wherein all patients undergo an HST regardless of clinical probability and the costs of the approach using each of the three prediction tools (MNC, STOP-BANG, and OSUNet) were calculated on the basis of a clinical algorithm shown in Figures 1A and 1B, respectively (4, 5). Patients identified as having a high risk of having OSA by the prediction tool (Figure 1B) go on to have an HST as the initial diagnostic procedure whereas those with less than a high risk have an in-laboratory PSG as the initial diagnostic procedure (5). The

algorithm includes the following assumptions: (1) a decision to pursue PSG has already been made by a clinician, and (2) patients who have negative results when an HST is used are referred for an in-laboratory PSG. The stakeholders for the cost-minimization analysis include the patients who are being evaluated for OSA, the health care providers who order the PSGs, and the third-party payers. Because the outcomes of the various approaches (third party, MNC, STOP-BANG, and OSUNet) are similar as shown in Figure 1, that is, all patients diagnosed with OSA will undergo treatment and those without OSA

will not have any treatment for OSA, we performed a cost-minimization analysis rather than other types of economic evaluations such as cost-benefit or cost-effectiveness analysis (21, 22).

The template for cost-minimization analysis using the clinical pathway described previously is shown in Table 1 and is based on the sensitivity (Sen) and specificity (Spec) of the prediction tools, as well as the prevalence (Prev) of OSA. The costs listed in Table 1 are also shown in Figures 1A and 1B. The cost of the sleep studies was based on 2013 Centers for Medicare & Medicaid Services global reimbursement rates:

Table 1. Template for cost-minimization analysis based on various approaches

| | Third-Party Approach | OSA Prediction Tool |
|--|---|---|
| Cost A: HST (US\$) | 174.00 | $[\text{Sen}(\text{Prev}) + (1 - \text{Spec})(1 - \text{Prev})] \times 174$ |
| Cost B: In-lab PSG (US\$) | — | $[\text{Prev}(1 - \text{Sen}) + \text{Spec}(1 - \text{Prev})] \times 607$ |
| Cost C: In-lab PSG for patients with negative HST (US\$) | $(1 - \text{Prev}) \times 607.00$ | $[(1 - \text{Prev}) \times (1 - \text{Spec})] \times 607.00$ |
| Total cost (US\$) | Cost A + cost B + cost C | Cost A + cost B + cost C |
| Savings per patient with prediction tool (US\$) | Total cost of OSA prediction tool less total cost of third-party approach | Savings per patient $\times 1,340,413^*$ |
| Total savings per year extrapolated to U.S. population | | |

Definition of abbreviations: HST, home sleep test; OSA, obstructive sleep apnea; PPV, positive predictive value for a given population was calculated as $(\text{Sen} \times \text{Prev}) / [(\text{Sen} \times \text{Prev}) + (1 - \text{Spec}) \times (1 - \text{Prev})]$; Prev, prevalence of OSA; PSG, polysomnography; Sen, sensitivity; Spec, specificity.

*Based on 427 PSGs performed per 10,000 population/year.

\$174.00 for HST and \$607.00 for in-laboratory PSG (23).

The cost of the third-party approach was calculated as \$174 for each HST in all patients suspected of OSA (cost A). There will be an additional cost of performing the in-laboratory PSG on those who do not have OSA calculated as $[(1 - \text{Prev}) \times \$607]$ (cost C). Therefore, the total cost for the third-party approach was calculated as $[\$174 + (1 - \text{Prev}) \times \$607]$.

The derivations of the various costs included in Table 1 are shown in Appendix E1. Briefly, the cost resulting from the use of the OSA prediction tool for the HSTs performed on the proportion of those patients with a positive prediction (the positive test rate of the tool) was calculated as $[\text{Sen}(\text{Prev}) + (1 - \text{Spec})(1 - \text{Prev})] \times 174$ (cost A). The proportion of those patients who were predicted not to have sleep apnea, the negative test rate of the tool, will have an in-laboratory PSG performed and the cost calculated as $[\text{Prev}(1 - \text{Sen}) + \text{Spec}(1 - \text{Prev})] \times 607$ (cost B). In addition, in-laboratory PSGs will be performed on the proportion of those patients who turn out to have a negative HST, but were predicted to have OSA by the tool: the false positives $[(1 - \text{Prev}) \times (1 - \text{Spec})] \times 607$. Therefore, the total cost of the approach using the OSA prediction tools was calculated as cost A + cost B + cost C = $[\{\text{Sen}(\text{Prev}) + (1 - \text{Spec})(1 - \text{Prev})\} \times 174] + \{\text{Prev}(1 - \text{Sen}) + \text{Spec}(1 - \text{Prev})\} \times 607 + \{[(1 - \text{Prev}) \times (1 - \text{Spec})] \times 607\}$.

The cost savings per patient from each of the prediction tools, if any, was then extrapolated to the entire U.S. population based on 427 PSGs performed/10,000 population/year (1). The U.S. population (313,914,040) was based on the 2012 U.S. Census Bureau estimate (<http://quickfacts.census.gov/qfd/states/00000.html>). The total costs from the four approaches described previously depend on the prevalence of OSA in the population where the clinical pathway is applied. Therefore, we tabulated the cost-minimization analysis based on the 49% prevalence of OSA in our clinic population as well as the 18–23% prevalence reported in primary care clinics (24, 25).

Sensitivity Analysis

Impact of OSA prevalence. We used a bootstrap approach (26) to obtain the confidence interval of cost savings at different levels of prevalence ranging from

Table 2. Demographics of clinic patients

| | In-Laboratory PSG (n = 532) | HST (n = 133) | P Value |
|------------------------|-----------------------------|---------------|---------|
| Age, yr | 48.5 ± 0.6 | 50.7 ± 1.1 | 0.06 |
| Sex, % male | 49.0 | 51.0 | 0.80 |
| BMI, kg/m ² | 35.8 ± 0.4 | 36.7 ± 0.9 | 0.31 |
| AHI | 25.3 ± 1.3 | 22.6 ± 2.1 | 0.30 |

Definition of abbreviations: AHI, apnea-hypopnea index; BMI, body mass index; HST, home sleep test; PSG, polysomnography.

Values represent means ± SEM unless otherwise indicated.

0.10 to 0.90. We implemented the bootstrap by generating 100,000 random samples with replacement from the 100 patients who underwent HST and in whom OSA status was confirmed for each of the three clinical prediction tools. Each random sample provided an estimate of sensitivity and specificity. We used these data to obtain 100,000 estimates of cost savings per patient at each level of prevalence from which the mean and 95% confidence limits were calculated with a computer program written in Fortran.

Impact of the proportion of HST used as the initial diagnostic test for OSA. In the cost minimization analysis described previously, the third-party approach (as the comparator vs. the clinical prediction rules) assumed that all patients suspected of OSA undergo an HST first. Although third-party payers in the United States are increasingly adopting a policy of requiring an HST, the exact proportion of patients who actually undergo HST in the United States as the initial diagnostic procedure is not known and will vary according to the preauthorization criteria established by

various third-party payers. Therefore, we performed a second sensitivity analysis wherein the relative proportions of HSTs and in-laboratory PSGs as the initial diagnostic procedure were varied from 1.0 to 0 in the third-party approach and the total expenditures compared with that using the three clinical prediction approaches.

Statistical Analysis

An AHI cutoff of at least 15/hour (based on an in-laboratory or HST) was used to define the presence of OSA (6, 7, 27). To test the equality of multiple sensitivities and specificities of the prediction tools, we used Cochran's Q applied to patients with or without OSA, respectively. The alternative hypothesis was that at least two sensitivities (or specificities) were different. The rejection of the Cochran's Q null hypothesis was followed by pairwise comparisons using McNemar's test. The comparisons of the positive and negative predictive values were based on the method of Moskowitz and Pepe (28). The positive (+) and negative (−) likelihood ratios (LRs) were compared using the method of Nofuentes

Table 3. Diagnostic characteristics of obstructive sleep apnea prediction tools with 95% confidence intervals in derivation group

| | OSUNet | MNC | STOP-BANG |
|-------------|-------------------|-------------------------------|------------------|
| Sensitivity | 0.74 (0.67–0.80)* | 0.87 (0.80–0.91) [†] | 0.95 (0.91–0.98) |
| Specificity | 0.78 (0.72–0.84)* | 0.42 (0.35–0.50) [†] | 0.21 (0.16–0.28) |
| PPV | 0.76 (0.69–0.82)* | 0.58 (0.52–0.64) [†] | 0.53 (0.48–0.59) |
| NPV | 0.76 (0.70–0.82) | 0.77 (0.68–0.85) | 0.82 (0.69–0.92) |
| Positive LR | 3.40 (2.56–4.47)* | 1.50 (1.31–1.72) [†] | 1.20 (1.12–1.31) |
| Negative LR | 0.34 (0.26–0.44) | 0.32 (0.21–0.48) | 0.23 (0.12–0.46) |

Definition of abbreviations: LR, likelihood ratio; MNC, modified neck circumference; NPV, negative predictive value; PPV, positive predictive value; STOP-BANG, Snoring, Tiredness, being Observed to stop breathing during sleep, high blood Pressure, Body mass index greater than 35 kg/m², Age greater than 50 years, Neck circumference greater than 40 cm, and male Gender.

Note: n = 383.

*Significantly different compared with MNC and STOP-BANG.

[†]Significantly different compared with STOP-BANG.

and Luna del Castillo when more than two binary diagnostic tests were applied to the same sample (29). All comparisons of the diagnostic performance of the three prediction tools using the previously described methodologies were accomplished with BDTcomparator software (<http://code.google.com/p/bdtcomparator/>) (30–32). All *P* values were adjusted for multiple comparisons, using Holm's procedure (33). Comparisons of the diagnostic performance of the prediction tools as well as the cost-minimization analysis were repeated with an AHI cutoff of at least 5/hour to define the presence of OSA.

Results

A total of 665 study participants were included in the study with an overall OSA prevalence of 49%; 532 study participants had in-laboratory PSG and 133 patients underwent HST. The demographic characteristics are presented in Table 2. There were no significant differences in age, sex, BMI, and severity of OSA between those patients who underwent an in-laboratory PSG compared with those who had an HST, although the latter tended to be older.

Performance Characteristics of OSA Prediction Tools

For the derivation group (Table 3), both the STOP-BANG and MNC had higher sensitivities compared with the OSUNet (95, 87, and 74%, respectively), with the STOP-BANG having the highest sensitivity among the three clinical prediction tools. However, the STOP-BANG also had the lowest specificity compared with the MNC and the OSUNet (21, 42, and 78%, respectively). The OSUNet had the highest positive predictive value (PPV) compared with the MNC and the STOP-BANG (76, 58, and 53%, respectively). There were no significant differences in the negative predictive values (NPVs).

For validation group 1 (Table 4), both the STOP-BANG and MNC had higher sensitivities compared with the OSUNet (97, 89, and 63%, respectively), with the STOP-BANG having the highest sensitivity among the three clinical prediction tools. The STOP-BANG also had the lowest specificity compared with the MNC and the OSUNet (12, 30, and 70%, respectively).

Table 4. Diagnostic characteristics of obstructive sleep apnea prediction tools with 95% confidence intervals in validation group 1

| | OSUNet | MNC | STOP-BANG |
|-------------|-------------------|-------------------------------|------------------|
| Sensitivity | 0.63 (0.51–0.74)* | 0.89 (0.80–0.95) [†] | 0.97 (0.91–1.00) |
| Specificity | 0.70 (0.58–0.80)* | 0.30 (0.20–0.42) [†] | 0.12 (0.06–0.21) |
| PPV | 0.67 (0.54–0.78)* | 0.55 (0.46–0.64) | 0.51 (0.43–0.60) |
| NPV | 0.66 (0.55–0.76) | 0.74 (0.55–0.88) | 0.82 (0.48–0.98) |
| Positive LR | 2.08 (1.42–3.06)* | 1.28 (1.08–1.51) | 1.10 (1.01–1.20) |
| Negative LR | 0.53 (0.38–0.74) | 0.36 (0.17–0.76) | 0.23 (0.05–1.04) |

Definition of abbreviations: LR, likelihood ratio; MNC, modified neck circumference; NPV, negative predictive value; PPV, positive predictive value; STOP-BANG, Snoring, Tiredness, being Observed to stop breathing during sleep, high blood Pressure, Body mass index greater than 35 kg/m², Age greater than 50 years, Neck circumference greater than 40 cm, and male Gender.

Note: n = 149.

*Significantly different compared with MNC and STOP-BANG.

[†]Significantly different compared with STOP-BANG.

The OSUNet had the highest PPV compared with the MNC and the STOP-BANG (67, 55, and 51%, respectively). There were no significant differences in the NPV.

For validation group 2 (Table 5), both the STOP-BANG and MNC had higher sensitivities compared with the OSUNet (97, 93, and 70%, respectively). The STOP-BANG and the MNC had lower specificities compared with the OSUNet (30, 44, and 79%, respectively). The OSUNet had the highest PPV compared with the MNC and the STOP-BANG (82, 69, and 65%, respectively). There were no significant differences in the NPV.

Comparison of the +LRs is shown in Figure 2 with the *P* values. In all groups, the OSUNet had significantly higher +LR compared with the other prediction tools. The MNC had a higher +LR compared with the STOP-BANG in the derivation group (*P* < 0.001) and also tended to be

higher in validation group 1 (*P* = 0.056) and validation group 2 (*P* = 0.087) compared with the STOP-BANG. There were no significant differences in the –LRs in all groups among the three prediction tools.

The diagnostic performance characteristics of the OSA prediction tools using an AHI cutoff of at least 5/hour to define OSA in the derivation group, validation group 1, and validation group 2 are shown in the online supplement (Tables E2, E3, and E4, respectively). Using the lower AHI cutoff, the +LRs of the OSUNet and MNC were higher compared with the STOP-BANG in both validation groups, but the differences did not achieve statistical significance.

Cost-Minimization Analysis

Results of the cost-minimization analysis for the various clinical pathways are shown in Table 6 for an OSA prevalence of 49% and

Table 5. Diagnostic characteristics of obstructive sleep apnea prediction tools with 95% confidence intervals in validation group 2

| | OSUNet | MNC | STOP-BANG |
|-------------|-------------------|------------------|------------------|
| Sensitivity | 0.70 (0.57–0.82)* | 0.93 (0.83–0.98) | 0.97 (0.88–0.99) |
| Specificity | 0.79 (0.64–0.90)* | 0.44 (0.29–0.60) | 0.30 (0.17–0.46) |
| PPV | 0.82 (0.68–0.91)* | 0.69 (0.57–0.79) | 0.65 (0.54–0.75) |
| NPV | 0.68 (0.52–0.79) | 0.83 (0.61–0.95) | 0.87 (0.60–0.98) |
| Positive LR | 3.35 (1.83–6.14)* | 1.67 (1.27–2.19) | 1.38 (1.13–1.69) |
| Negative LR | 0.38 (0.25–0.58) | 0.16 (0.06–0.43) | 0.12 (0.03–0.49) |

Definition of abbreviations: LR, likelihood ratio; MNC, modified neck circumference; NPV, negative predictive value; PPV, positive predictive value; STOP-BANG, Snoring, Tiredness, being Observed to stop breathing during sleep, high blood Pressure, Body mass index greater than 35 kg/m², Age greater than 50 years, Neck circumference greater than 40 cm, and male Gender.

Note: n = 100.

*Significantly different compared with MNC and STOP-BANG.

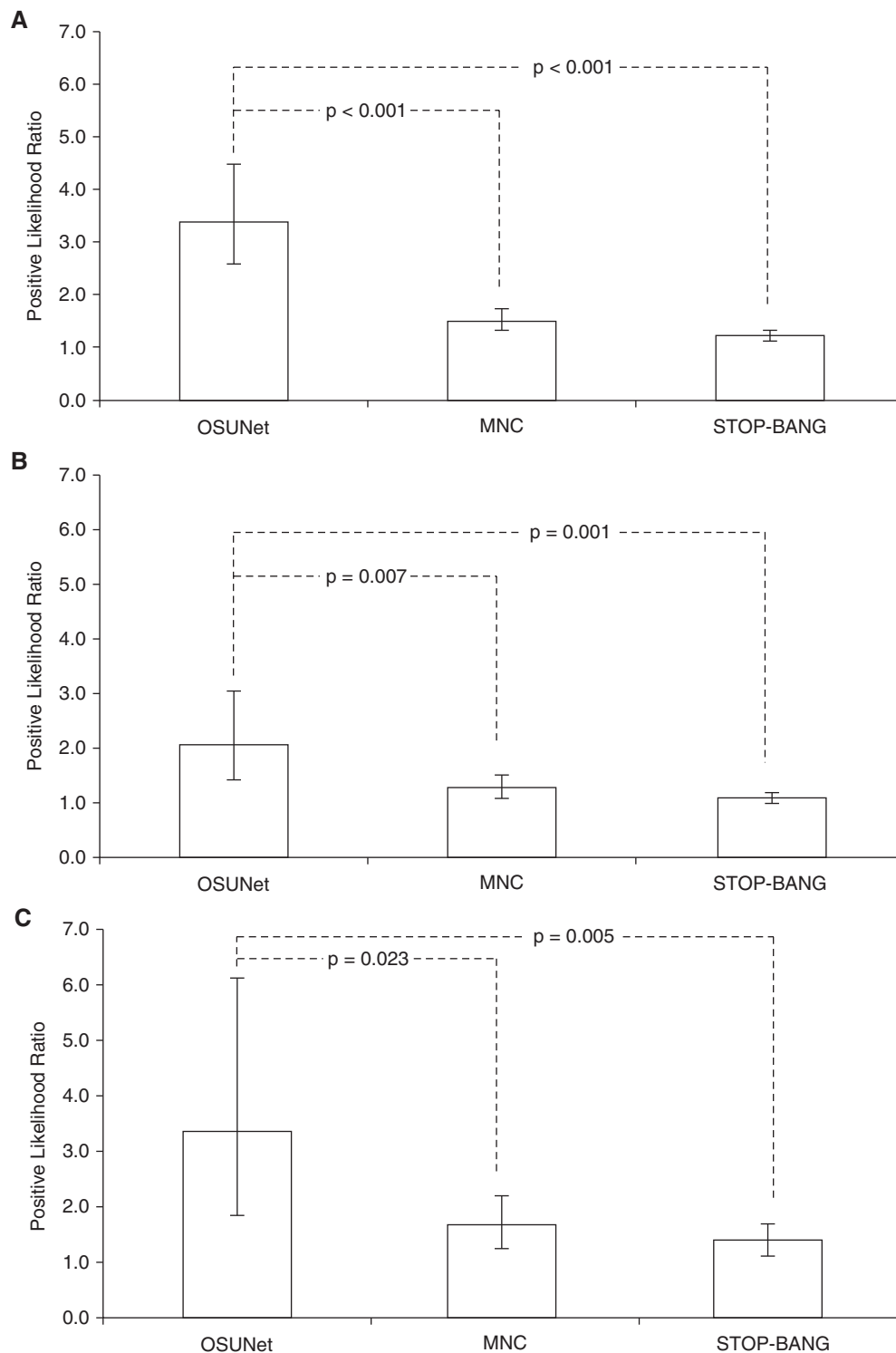


Figure 2. Comparison of positive likelihood ratios of clinical prediction tools. (A) Derivation group, (B) validation group 1, and (C) validation group 2. The error bars represent 95% confidence intervals. MNC = modified neck circumference; STOP-BANG, Snoring, Tiredness, being Observed to stop breathing during sleep, high blood Pressure, Body mass index greater than 35 kg/m², Age greater than 50 years, Neck circumference greater than 40 cm, and male Gender.

20%, respectively (calculated in 1,000 patients). The use of all clinical prediction tools in the algorithm had positive cost

savings compared with the third-party approach. The cost savings were greater with the use of the OSUNet when the

prevalence of OSA was lower, at 20%, such as what may be encountered in primary care clinics. In the latter situation, the cost

savings for every 1,000 patients when the OSUNet was used to select those who should undergo an HST as the initial diagnostic procedure compared with the third-party approach were \$84,237. When extrapolated to the entire U.S. population, the cost savings were \$112,912,432 per year. As shown in Table 6, the use of either the STOP-BANG or MNC as a guide to selecting patients who should undergo an HST as the initial diagnostic test also offered cost savings compared with the third-party approach; the savings from both these tools were also greater with lower OSA prevalence.

Sensitivity Analysis

The sensitivity analysis of the effect of OSA prevalence for the use of the three clinical prediction tools is shown in Figure 3. The cost savings and the 95% confidence interval of the cost savings using the OSUNet, MNC, and STOP-BANG according to varying prevalence of OSA is shown in Figures 3A–3C, respectively. The cost savings increased if the population where the clinical pathway was applied had a lower prevalence of OSA, with the use of the OSUNet resulting in the biggest savings when the OSA prevalence was less than 40%. Conversely, if the population being studied had an extremely high OSA prevalence the savings from using the clinical prediction tools were reduced. Use of the OSUNet, MNC, and STOP-BANG

compared with the third-party approach resulted in positive savings only when the OSA prevalence was less than 52, 72, and 78%, respectively. Similar findings were achieved when the presence of OSA was defined as an AHI of at least 5/hour instead of at least 15/hour. Use of the OSUNet, MNC, and STOP-BANG compared with the third-party approach resulted in positive savings only when the OSA prevalence was less than 58, 66, and 74%, respectively (Figure 4).

The second sensitivity analysis for use of the OSA clinical prediction tools is shown in Figure E1, using an OSA prevalence of 49%, the prevalence in our clinic population. The cost savings from use of the OSA clinical prediction tools were calculated according to varying proportions of HST from 1.0 to 0 used as the initial test to diagnose OSA in the third-party approach. The cost savings from use of the OSA clinical prediction tools were greater if the proportion of HSTs used as the initial diagnostic procedure by the third-party payer was lower than 1.0. As expected, a strategy of using only in-laboratory PSG as the initial diagnostic procedure (0% HST and 100% in-laboratory PSG) had the highest expense in the third-party approach.

Discussion

This study investigated the diagnostic performance characteristics of two widely

used clinical prediction tools for OSA (STOP-BANG and MNC) and the newly developed OSUNet. We then performed a cost-minimization analysis in which the three prediction tools were used to select patients who should undergo an HST as the initial diagnostic procedure compared with an approach now being adopted by third-party payers to use HSTs in all patients regardless of clinical probability. The major findings of this study are (1) the newly developed OSUNet, which is based on ANN technology, had a higher +LR compared with the STOP-BANG and MNC; (2) all three prediction tools have equivalent –LRs, but these are not sufficient to exclude the need for a sleep study; (3) the routine requirement of an HST to diagnose OSA regardless of clinical probability is more costly compared with the use of OSA clinical prediction tools that identify patients who should undergo this procedure when OSA is expected to be present in less than half of the population; and (4) the savings from the clinical prediction tools were greater when used in a population with lower OSA prevalence, with the biggest savings derived from the use of the OSUNet when the OSA prevalence was less than 40%. The use of the clinical prediction tools also resulted in net savings regardless of the proportion of HSTs and in-laboratory PSGs that third-party payers would authorize as the initial diagnostic test for patients suspected of OSA. In settings where the prevalence of OSA is expected to be substantially lower than in sleep centers, such as in primary care clinics (24, 25), we believe that use of the OSUNet as a guide to which patients should undergo an HST versus an in-laboratory PSG to establish the diagnosis of OSA offers substantial benefits to all stakeholders involved in the process—the patients who are being evaluated for OSA (because of convenience if they are good candidates for the portable HST and avoidance of repetitive testing), the health care providers who order the sleep studies (because it facilitates decision-making on what type of test to order), and the third-party payers who reimburse for the diagnostic testing (because of substantial cost savings). To our knowledge, this is the first study that has performed a cost-minimization analysis that can be derived from the use of OSA clinical prediction tools.

Physicians are faced with two questions during the evaluation of an individual with

Table 6. Results of cost-minimization analysis in 1,000 patients

| | Third-Party Approach | OSUNet | MNC | STOP-BANG |
|--|----------------------|---------|---------|-----------|
| OSA Prevalence = 0.49 | | | | |
| Cost A: HST (US\$) | 174,000 | 78,405 | 128,806 | 144,180 |
| Cost B: In-lab PSG (US\$) | | 333,483 | 157,659 | 104,027 |
| Cost C: In-lab PSG for patients with negative HST (US\$) | 309,570 | 64,794 | 172,783 | 215,979 |
| Total cost (US\$) | 483,570 | 476,682 | 459,248 | 464,186 |
| Savings per 10 ³ patients with prediction tool (US\$) | | 6,888 | 24,322 | 19,384 |
| OSA Prevalence = 0.20 | | | | |
| Cost A: HST (US\$) | 174,000 | 53,556 | 110,051 | 130,695 |
| Cost B: In-lab PSG (US\$) | | 420,170 | 223,087 | 151,069 |
| Cost C: In-lab PSG for patients with negative HST (US\$) | 485,600 | 101,637 | 271,032 | 338,791 |
| Total cost (US\$) | 659,600 | 575,363 | 604,170 | 620,555 |
| Savings per 10 ³ patients with prediction tool (US\$) | | 84,237 | 55,430 | 39,045 |

Definition of abbreviations: AHI, apnea–hypopnea index; BMI, body mass index; HST, home sleep test; MNC, modified neck circumference; OSA, obstructive sleep apnea; PSG, polysomnography; STOP-BANG, Snoring, Tiredness, being Observed to stop breathing during sleep, high blood Pressure, Body mass index greater than 35 kg/m², Age greater than 50 years, Neck circumference greater than 40 cm, and male Gender.

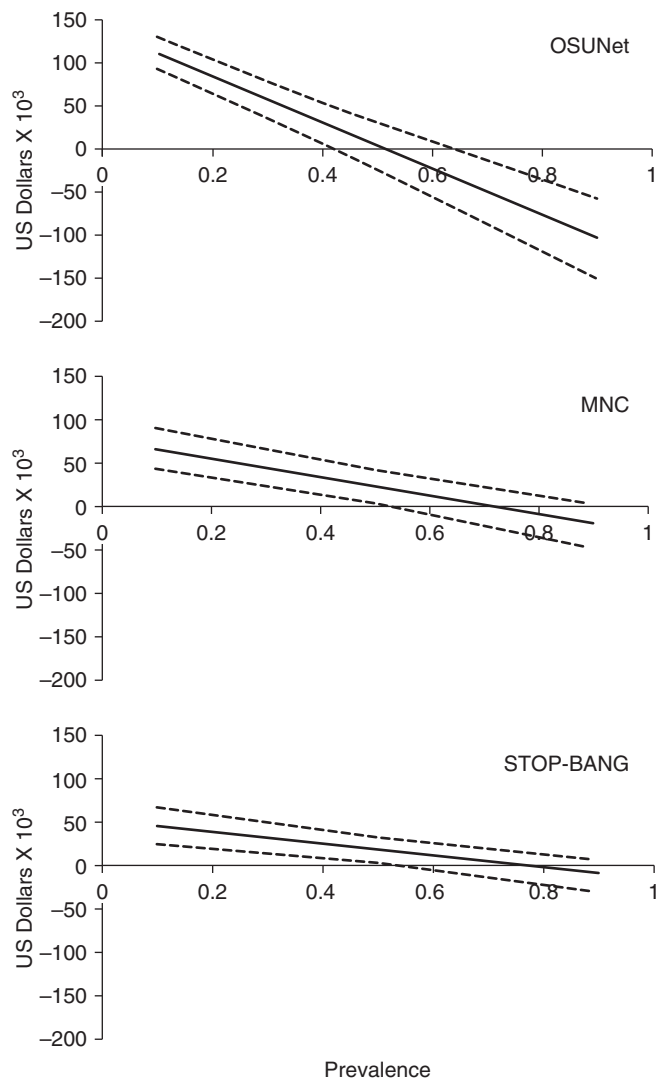


Figure 3. Cost savings from the use of prediction tools using an apnea–hypopnea index (AHI) of at least 15/hour to define the presence of obstructive sleep apnea (OSA). (A) Savings per 1,000 patients using the OSUNet to select patients to undergo a home sleep test (HST) as the initial diagnostic procedure according to prevalence of OSA. (B) Savings per 1,000 patients using the modified neck circumference (MNC) to select patients to undergo HST as the initial diagnostic procedure according to prevalence of OSA. (C) Savings per 1,000 patients using the STOP-BANG to select patients to undergo HST as the initial diagnostic procedure according to prevalence of OSA. The *dashed lines* represent the corresponding 95% confidence intervals. STOP-BANG, Snoring, Tiredness, being Observed to stop breathing during sleep, high blood Pressure, Body mass index greater than 35 kg/m², Age greater than 50 years, Neck circumference greater than 40 cm, and male Gender.

symptoms consistent with OSA. The first question is whether the patient warrants a PSG. Unfortunately, the assignment of OSA probability via clinician impression has been shown to be inferior to clinical prediction rules (10). This has resulted in the use of OSA clinical prediction tools. The majority of these prediction tools are based on regression analysis. Two OSA clinical prediction rules that have been

widely used are the STOP-BANG (11) and MNC (6, 10, 13), likely because of their simplicity and ease of calculation. Consistent with prior studies on these tools, we found that they both indeed have high sensitivity, but suffer from low specificity (12, 34, 35), particularly the STOP-BANG questionnaire. Their +LRs were less than 2 in all three of our patient groups and therefore use of these tools did not result in

any substantial change in the probability of having OSA in any particular patient (36). Our results are consistent with a prior study that examined the diagnostic characteristics of the STOP-BANG in a large community sample ($n = 4,770$) included in the Sleep Heart Health Study (34). In this community sample, the STOP-BANG was found to have a sensitivity of 87%, specificity of 43%, +LR of 1.5, and –LR of 0.30. A systematic review also reported the same characteristics of this tool with a sensitivity of 93%, specificity of 35%, +LR of 1.4, and –LR of 0.20 (9). More recently, the authors of the STOP-BANG questionnaire have advocated a change in the criteria of having a high probability of moderate to severe OSA to at least 5 instead of at least 3 (37). With this change in cutoff, they reported a sensitivity of 23%, specificity of 56%, PPV of 31%, and NPV of 45%. Therefore, the +LR will be even lower and would not improve the results even if we had used this new cutoff in our groups of patients.

ANN technology mimics the human brain's own problem-solving process (38). Just as humans apply knowledge gained from past experience to new problems, a neural network takes previously solved examples to build a system of neurons (nodes) that makes new decisions and classifications. It provides an alternative to logistic regression, the most commonly used method for developing predictive models for dichotomous outcomes in medicine (14). Compared with regression, ANN offers several advantages in prediction including the ability to implicitly detect complex nonlinear relations between dependent and independent variables, the ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms (14). We found that the newly developed OSUNet, a prediction tool based on ANN, had higher +LR compared with the other OSA prediction tools. We have used the +LR to compare the three clinical predictions because it constitutes one of the best ways to measure diagnostic accuracy (39). In our clinic population, the OSUNet provided a better prediction of the presence of OSA (changing the probability by about 20%) compared with the STOP-BANG or MNC. We also found that none of the three clinical prediction tools have sufficiently low –LR to exclude the possibility of OSA, and therefore they would have limited

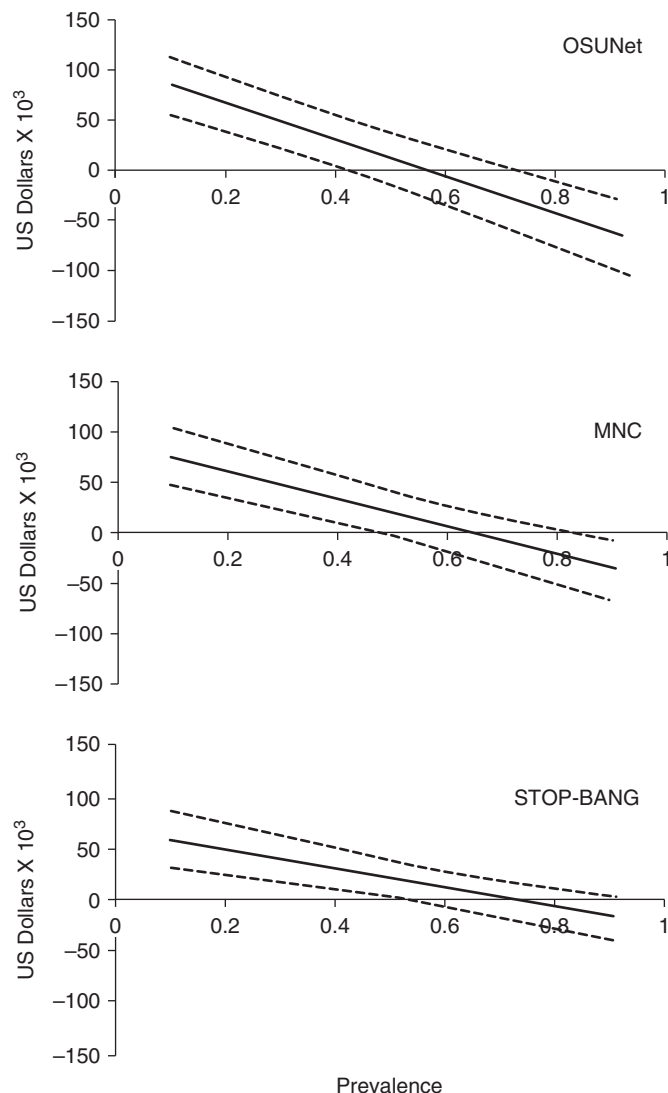


Figure 4. Cost savings from the use of prediction tools using an apnea–hypopnea index (AHI) of at least 5/hour to define the presence of obstructive sleep apnea (OSA). (A) Savings per 1,000 patients using the OSUNet to select patients to undergo a home sleep test (HST) as the initial diagnostic procedure according to prevalence of OSA. (B) Savings per 1,000 patients using the modified neck circumference (MNC) to select patients to undergo HST as the initial diagnostic procedure according to prevalence of OSA. (C) Savings per 1,000 patients using the STOP-BANG to select patients to undergo HST as the initial diagnostic procedure according to prevalence of OSA. The *dashed lines* represent the corresponding 95% confidence intervals. STOP-BANG, Snoring, Tiredness, being Observed to stop breathing during sleep, high blood Pressure, Body mass index greater than 35 kg/m², Age greater than 50 years, Neck circumference greater than 40 cm, and male Gender.

usefulness in negating the need for a sleep study (36).

The second question that a clinician is faced with is the type of sleep study to order—whether an in-laboratory PSG or an HST. Increasingly, this decision is out of the hands of practitioners because many third-party payers now require preauthorization for an in-laboratory PSG, but not for an HST.

That HSTs have a significant role in the diagnosis of OSA is now indeed supported by several studies (6, 7). This change has resulted in third-party payers adopting a policy of requiring the portable HST as the initial diagnostic procedure for all patients regardless of clinical probability (40). However, most of the studies supporting a role for HSTs were done in populations where the prevalence of OSA is

high. For example, in a study that showed the noninferiority of ambulatory management of OSA, the prevalence of the condition was 88% (7). Indeed, our study suggests that when the prevalence of OSA is that high, the cost savings derived from the clinical prediction tools as a guide to PSG are nonexistent. However, this is not the situation in the majority of clinical settings where the prevalence of OSA is expected to be much lower, for example, in primary care offices where the prevalence has been reported to be 18–23% (24, 25). Interestingly, these previously reported OSA prevalences in the primary care clinics are similar to the most recent estimate of 15% in the general population in the Wisconsin Sleep Cohort when OSA is defined as an AHI of at least 5/hour in symptomatic patients (Epworth Sleepiness Scale score > 10) (41). We used an AHI cutoff of at least 15/hour to define the presence of OSA in our primary analysis mainly because treatment of these individuals is considered standard of practice, whereas treatment of those with an AHI of 5–14/hour is considered an option and remains controversial (42). Nonetheless, when an AHI cutoff of at least 5/hour was used to define the presence of OSA in our study, we also found that the third-party approach is more costly compared with the use of OSA clinical prediction tools when OSA is present in less than half of the population. When the volume of sleep studies performed in the United States is taken into consideration (1), the amount of cost savings per year from using a simple tool such as the OSUNet is significant. Furthermore, use of the OSUNet within the clinical algorithm in Figure 1 also offers an objective tool in helping the practitioner make consistent decisions as to which patients should go on to have an HST versus an in-laboratory PSG.

Three prior studies have reported the use of ANN to predict the presence and severity of OSA based on demographic characteristics and answers to sleep questionnaires (19, 20, 43). However, none of these studies compared their diagnostic performance with the widely used STOP-BANG and MNC prediction tools in the same group of patients. In addition, these prior studies also did not involve HSTs and none performed a cost-minimization analysis regarding use of the prediction tools to guide clinicians to identify those

who will be ideal candidates for in-home tests.

Our study had limitations. First, we assumed that all patients who had a negative HST will be referred by their practitioner for an in-laboratory PSG. As stated previously, clinicians have a low threshold to proceed with an in-laboratory study for symptomatic patients with a negative HST, and because a significant proportion of patients with a negative HST may have significant OSA when studied in the laboratory (6, 7). Therefore, we believe that our assumption would be reasonable. However, the proportion of patients with a negative HST who would really need to be referred for an in-laboratory PSG (as opposed to other nontesting strategies such as clinical follow-up) is unknown. This matter should be the subject of future studies because additional significant savings could be anticipated if it could be predicted reliably which patients with a negative in-home test will also turn out to have a negative in-laboratory PSG. It is entirely possible that patients predicted by the OSUNet to have a low probability of having OSA with a negative HST do not need to have an in-laboratory test. Second,

our study participants are clinic patients with a relatively high prevalence (49%) of OSA. Further studies will be needed in the primary care setting, where the prevalence of OSA is expected to be lower, to determine the test characteristics and cost savings from the OSUNet. Our sensitivity analysis, however, suggests that even greater cost savings from use of the OSUNet could be achieved when the prevalence of OSA is lower. Third, our study involved only a single sleep disorders center. Multicenter studies involving a larger number of patients will be needed to determine the generalizability of our findings. Fourth, it was beyond the scope of our study to take into account the treatment phase in our cost-minimization analysis. However, two studies show that the use of automatic continuous positive airway pressure (CPAP) titration at home has equivalent outcomes in those OSA patients without significant comorbidities compared with an in-laboratory CPAP titration. Therefore, this approach would be a reasonable strategy for treatment after a diagnosis of OSA has been confirmed by PSG in the majority of cases. Finally, our study applies to the evaluation of OSA only

and we recognize that sleep clinicians evaluate patients with other sleep disorders that may require an in-laboratory PSG.

In summary, when the prevalence of OSA is less than 52%, our study strongly suggests that an approach requiring an HST regardless of clinical probability that is increasingly being adopted by third-party payers results in greater costs compared with a clinical pathway incorporating the OSA clinical prediction tools to identify patients who should undergo an HST. The OSUNet, as compared with the MNC and STOP-BANG, offers greater savings when the OSA prevalence is less than 40%, such as what may be expected in the primary care setting. The savings are substantial when extrapolated to the number of sleep studies done in the United States. The savings from use of the OSA clinical prediction tools are more pronounced the lower the prevalence of OSA in the population to which it is applied. Future studies involving other sleep disorders center as well as primary care clinics are needed to determine the generalizability of our findings. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

- 1 Tachibana N, Ayas NT, White DP. A quantitative assessment of sleep laboratory activity in the United States. *J Clin Sleep Med* 2005;1:23–26.
- 2 Pack AI. What can sleep medicine do? *J Clin Sleep Med* 2013;9:629.
- 3 Quan SF, Epstein LJ. A warning shot across the bow: the changing face of sleep medicine. *J Clin Sleep Med* 2013;9:301–302.
- 4 Ayas NT, Fox J, Epstein L, Ryan CF, Fleetham JA. Initial use of portable monitoring versus polysomnography to confirm obstructive sleep apnea in symptomatic patients: an economic decision model. *Sleep Med* 2010;11:320–324.
- 5 Collop NA, Anderson WM, Boehlecke B, Claman D, Goldberg R, Gottlieb DJ, Hudgel D, Sateia M, Schwab R; Portable Monitoring Task Force of the American Academy of Sleep Medicine. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. *J Clin Sleep Med* 2007;3:737–747.
- 6 Rosen CL, Auckley D, Benca R, Foldvary-Schaefer N, Iber C, Kapur V, Rueschman M, Zee P, Redline S. A multisite randomized trial of portable sleep studies and positive airway pressure autotitration versus laboratory-based polysomnography for the diagnosis and treatment of obstructive sleep apnea: the HomePAP study. *Sleep* 2012;35:757–767.
- 7 Kuna ST, Gurubhagavatula I, Maislin G, Hin S, Hartwig KC, McCloskey S, Hachadoorian R, Hurley S, Gupta R, Staley B, et al. Noninferiority of functional outcome in ambulatory management of obstructive sleep apnea. *Am J Respir Crit Care Med* 2011;183:1238–1244.
- 8 Masa JF, Corral J, Pereira R, Duran-Cantolla J, Cabello M, Hernández-Blasco L, Monasterio C, Alonso A, Chiner E, Zamorano J, et al.; Spanish Sleep Network. Therapeutic decision-making for sleep apnea and hypopnea syndrome using home respiratory polygraphy: a large multicentric study. *Am J Respir Crit Care Med* 2011;184:964–971.
- 9 Myers KAMM, Mrkobrada M, Simel DL. Does this patient have obstructive sleep apnea? The Rational Clinical Examination systematic review. *JAMA* 2013;310:731–741.
- 10 Flemons WW, Whitelaw WA, Brant R, Remmers JE. Likelihood ratios for a sleep apnea clinical prediction rule. *Am J Respir Crit Care Med* 1994;150:1279–1285.
- 11 Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, Khajehdehi A, Shapiro CM. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. *Anesthesiology* 2008;108:812–821.
- 12 Rowley JA, Aboussouan LS, Badr MS. The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. *Sleep* 2000;23:929–938.
- 13 Flemons WW. Clinical practice: obstructive sleep apnea. *N Engl J Med* 2002;347:498–504.
- 14 Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol* 1996;49:1225–1231.
- 15 Baxt WG. Use of an artificial neural network for the diagnosis of myocardial infarction. *Ann Intern Med* 1991;115:843–848.
- 16 Iber C, Ancoli-Israel S, Chesson A, and Quan SF; for the American Academy of Sleep Medicine. The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specifications. Westchester, IL: American Academy of Sleep Medicine; 2007.
- 17 Hoehfeld M, Fahlman SE. Learning with limited numerical precision using the cascade-correlation algorithm. *IEEE Trans Neural Netw* 1992;3:602–611.
- 18 Fahlman S, Lebiere C. The cascade-correlation learning architecture. In: Touretzky D, editor. *Advances in neural information processing systems*. San Mateo, CA: Morgan Kaufmann; 1990, pp. 524–532.
- 19 Kirby SD, Eng P, Danter W, George CF, Francovic T, Ruby RR, Ferguson KA. Neural network prediction of obstructive sleep apnea from clinical criteria. *Chest* 1999;116:409–415.

- 20 el-Solh AA, Mador MJ, Ten-Brock E, Shucard DW, Abul-Khoudoud M, Grant BJ. Validity of neural network in sleep apnea. *Sleep* 1999;22:105–111.
- 21 Robinson R. Costs and cost-minimisation analysis. *BMJ* 1993;307:726–728.
- 22 Higgins AM, Harris AH. Health economic methods: cost-minimization, cost-effectiveness, cost-utility, and cost-benefit evaluations. *Crit Care Clin* 2012;28:11–24, v.
- 23 Young TK, Martens PJ, Taback SP, Sellers EA, Dean HJ, Cheang M, Flett B. Type 2 diabetes mellitus in children: prenatal and early infancy risk factors among native Canadians. *Arch Pediatr Adolesc Med* 2002;156:651–655.
- 24 Kushida CA, Nichols DA, Simon RD, Young T, Grauke JH, Britzmann JB, Hyde PR, Dement WC. Symptom-based prevalence of sleep disorders in an adult primary care population. *Sleep Breath* 2000;4:9–14.
- 25 Heffner JE, Rozenfeld Y, Kai M, Stephens EA, Brown LK. Prevalence of diagnosed sleep apnea among patients with type 2 diabetes in primary care. *Chest* 2012;141:1414–1421.
- 26 Efron BTR. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Stat Sci* 1986;1:54–77.
- 27 Masa JF, Corral J, Pereira R, Duran-Cantolla J, Cabello M, Hernández-Blasco L, Monasterio C, Alonso A, Chiner E, Rubio M, *et al.* Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome. *Thorax* 2011;66:567–573.
- 28 Moskowitz CS, Pepe MS. Comparing the predictive values of diagnostic tests: sample size and analysis for paired study designs. *Clin Trials* 2006;3:272–279.
- 29 Nofuentes JAR, Luna del Castillo JdD. Comparison of the likelihood ratios of two binary diagnostic tests in paired designs. *Stat Med* 2007;26:4179–4201.
- 30 Fijorek K, Fijorek D, Wisniewska B, Polak S. BDTcomparator: a program for comparing binary classifiers. *Bioinformatics* 2011;27:3439–3440.
- 31 Jastrzebski M, Kukla P, Fijorek K, Sondej T, Czarnicka D. Electrocardiographic diagnosis of biventricular pacing in patients with nonapical right ventricular leads. *Pacing Clin Electrophysiol* 2012;35:1199–1208.
- 32 Arias-Loste MT, Bonilla G, Moraleja I, Mahler M, Mieses MA, Castro B, Rivero M, Crespo J, López-Hoyos M. Presence of anti-proteinase 3 antineutrophil cytoplasmic antibodies (anti-PR3 ANCA) as serologic markers in inflammatory bowel disease. *Clin Rev Allergy Immunol* 2013;45:109–116.
- 33 Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat* 1979;6:65–70.
- 34 Silva GE, Vana KD, Goodwin JL, Sherrill DL, Quan SF. 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;7:467–472.
- 35 Kunisaki KM, Brown KE, Fabbri AE, Wetherbee EE, Rector TS. STOP-BANG questionnaire performance in a Veterans Affairs unattended sleep study program. *Ann Am Thorac Soc* 2014;11:192–197.
- 36 McGee S. Simplifying likelihood ratios. *J Gen Intern Med* 2002;17:646–649.
- 37 Chung F, Subramanyam R, Liao P, Sasaki E, Shapiro C, Sun Y. High STOP-BANG score indicates a high probability of obstructive sleep apnoea. *Br J Anaesth* 2012;108:768–775.
- 38 Hinton GE. How neural networks learn from experience. *Sci Am* 1992;267:144–151.
- 39 Deeks JJ, Altman DG. Diagnostic tests. 4. Likelihood ratios. *BMJ* 2004;329:168–169.
- 40 Pack AI. Sleep medicine: strategies for change. *J Clin Sleep Med* 2011;7:577–579.
- 41 Peppard PE, Young T, Barnett JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol* 2013;177:1006–1014.
- 42 Kushida CA, Littner MR, Hirshkowitz M, Morgenthaler TI, Alessi CA, Bailey D, Boehlecke B, Brown TM, Coleman J Jr, Friedman L, *et al.*; American Academy of Sleep Medicine. Practice parameters for the use of continuous and bilevel positive airway pressure devices to treat adult patients with sleep-related breathing disorders. *Sleep* 2006;29:375–380.
- 43 Sun LM, Chiu H-W, Chuang CY, Liu L. A prediction model based on an artificial intelligence system for moderate to severe obstructive sleep apnea. *Sleep Breath* 2011;15:317–323.