≋ CHEST®

CrossMark

Sex-Specific Prediction Models for Sleep Apnea From the Hispanic Community Health Study/Study of Latinos

Neomi Shah, MD, MPH; David B. Hanna, PhD; Yanping Teng, MD, MSPH; Daniela Sotres-Alvarez, PhD; Martica Hall, PhD; Jose S. Loredo, MD, MPH; Phyllis Zee, MD, PhD; Mimi Kim, ScD; H. Klar Yaggi, MD; Susan Redline, MD, MPH; and Robert C. Kaplan, PhD

OBJECTIVE: We developed and validated the first-ever sleep apnea (SA) risk calculator in a large population-based cohort of Hispanic/Latino subjects.

METHODS: Cross-sectional data on adults from the Hispanic Community Health Study/Study of Latinos (2008-2011) were analyzed. Subjective and objective sleep measurements were obtained. Clinically significant SA was defined as an apnea-hypopnea index $\geq$ 15 events per hour. Using logistic regression, four prediction models were created: three sex-specific models (female-only, male-only, and a sex $\times$ covariate interaction model to allow differential predictor effects), and one overall model with sex included as a main effect only. Models underwent 10-fold cross-validation and were assessed by using the C statistic. SA and its predictive variables; a total of 17 variables were considered.

RESULTS: A total of 12,158 participants had complete sleep data available; 7,363 (61%) were women. The population-weighted prevalence of SA (apnea-hypopnea index $\geq$ 15 events per hour) was 6.1% in female subjects and 13.5% in male subjects. Male-only (C statistic, 0.808) and female-only (C statistic, 0.836) prediction models had the same predictor variables (ie, age, BMI, self-reported snoring). The sex-interaction model (C statistic, 0.836) contained sex, age, age $\times$ sex, BMI, BMI $\times$ sex, and self-reported snoring. The final overall model (C statistic, 0.832) contained age, BMI, snoring, and sex. We developed two websites for our SA risk calculator: one in English (https://www.montefiore.org/sleepapneariskcalc.html) and another in Spanish (http://www.montefiore.org/sleepapneariskcalc-es.html).

CONCLUSIONS: We created an internally validated, highly discriminating, well-calibrated, and parsimonious prediction model for SA. Contrary to the study hypothesis, the variables did not have different predictive magnitudes in male and female subjects.

CHEST 2016; 149(6):1409-1418

KEY WORDS: clinical decision-making; community health; epidemiology (pulmonary); sex-specific prediction; sleep apnea

---

Sleep apnea (SA) affects at least 25 million adults in the United States.[1,2] Historically, SA has been considered a predominately male condition. Clinic-based studies have estimated that the prevalence of SA among men compared with women approaches a ratio of 8:1.[3,4] As a result, SA-screening tools have been based on male-specific SA symptoms such as snoring and excessive daytime sleepiness. Women, however, tend to underreport these "classic" symptoms of SA.[3] Instead, women with SA frequently present with symptoms such as lack of energy, depression, and insomnia,[5,6] and these female-specific SA symptoms are not routinely captured in current SA prediction tools. Therefore, a need for sex-specific SA screening tools is obvious. Furthermore, sex-specific SA prediction tools are now even more necessary, with epidemiologic evidence revealing a higher than previously reported prevalence of SA in women. The ratio of SA in men compared with women is currently believed to be closer to 2 to 3:1 (vs 8:1).[3,7,8]

There is a widespread underdiagnosis of SA in female subjects, with > 90% of female subjects with SA going undiagnosed.[9] This scenario is partly due to the concerns highlighted earlier, which include sex-related differences in the clinical presentation of SA and lack of sex-specific prediction tools for SA. Interestingly, well-established prediction tools for other diseases (eg, coronary heart disease) have demonstrated the importance of incorporating sex-specific variables to improve the predictive ability.[10,11] The same may be true for SA, but further investigation is needed.

To create a comprehensive prediction tool for SA that captures sex-specific symptoms, we designed a study in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL), the largest known population-based study to date with objective SA data. In the present study, we developed and validated sex-specific prediction equations for SA by using symptom, demographic, and anthropometric data. We assessed whether incorporating sex differences improves SA prediction by using established measures of model discrimination. The study hypothesis was that sex-specific prediction models would perform better than a non-sex-specific prediction model for SA.

Materials and Methods

The Albert Einstein College of Medicine institutional review board approved the HCHS/SOL protocol prior to data collection. Their institutional review board number is 2007-432, and their reference number is 008048.

Study Population

HCHS/SOL is a community-based cohort study sponsored by the National Institutes of Health consisting of 16,415 self-identified Hispanic/Latino participants aged 18 to 74 years residing in four US sites (Bronx, New York; Chicago, Illinois; Miami, Florida; and San Diego, California). The details of the study design, sampling strategy,

CORRESPONDENCE TO: Neomi Shah, MD, MPH, Icahn School of Medicine at Mount Sinai, One Gustave L Levy Pl, Box 1232, New York, NY 10029; e-mail: Neomi.Shah@mssm.edu

and recruitment have been described elsewhere.[12,13] All participants underwent a baseline interview and an extensive clinic examination (2008-2011) that included but was not limited to anthropometry, ECG, blood pressures in both arms, phlebotomy, glucose tolerance testing, audiometry, lung function, and oral examination. Questionnaires included sociodemographic characteristics, health and medical history, smoking, alcohol, medication use, occupational history, diet, and physical activity.

Assessment of Sleep-Disordered Breathing

In HCHS/SOL, both subjective and objective sleep data were collected. The sleep questionnaire included SA symptoms, questions, and sleep patterns from the sleep habits questionnaire of the Sleep Heart Health Study[14] and the Epworth Sleepiness Scale (ESS).[15] Objective SA testing included a home sleep test using the ARES Unicorder 5.2 (B-Alert).[16] The ARES monitor is a self-applied device that measures blood oxygen saturation and pulse rate (reflectance pulse oximetry), airflow (by nasal cannula connected to a pressure transducer), snoring levels (calibrated acoustic microphone), and head movement and head position (accelerometers). The sleep data were scored by the HCHS/SOL Sleep Reading Center. Respiratory events were defined as a ≥ 50% reduction in airflow lasting at least 10 s. Apneas were not distinguished from hypopneas because thermistry testing was not available. Each respiratory event was manually identified and linked to its level of desaturation, and artifacts were manually edited on an epoch-by-epoch basis. The present analyses used the 3% desaturation variable in accordance with current recommended scoring criteria from the American Academy of Sleep Medicine.[17,18] The majority of the sleep studies (≥ 80%) included in the analyses had ≥ 4 h of total recording time. Sleep studies with < 30 min were defined as incomplete and were excluded. The apnea-hypopnea index (AHI) data obtained from the ARES has an excellent correlation with AHI data obtained from in-laboratory

polysomnography.[16,19] The primary definition of SA was established by using an AHI of ≥ 15 events per hour. An AHI ≥ 15 events per hour was chosen as our cut point to define SA and thus enable the capture of moderate and severe SA, which are typically associated with adverse cardiovascular outcomes rather than milder degrees of SA.[20,21] Mild SA was defined as an AHI ≥ 5 events per hour and < 15 events per hour; moderate SA as AHI ≥ 15 events per hour and < 30 events per hour; and severe SA as AHI ≥ 30 events per hour. Interscorer and intrascorer reliability estimates for the AHI, assessed over the course of the study, were excellent (intraclass correlation coefficients, > 0. 99).

Variables Considered for Prediction Models

Waist and hip circumference were measured to compute the waist-to-hip ratio. Insomnia, hypertension, diabetes, and heart disease were self-reported (yes/no). Depressive symptoms were assessed by using the Center for Epidemiologic Studies Depression (CESD) scale (10-item version).[22] A CESD score (range, 0-30) ≥ 10 was considered indicative of depression symptoms. The question "Have you felt downhearted and depressed?" was also used to assess symptoms of depression. Energy level of participants was assessed by using the question, "Did you have a lot of energy?" Reponses to these questions were collapsed into yes/no (yes = all or most of the time; no = some, a little, or none of the time). Snoring, witnessed apneas, naps, difficulty falling asleep, excessive daytime sleepiness (EDS), and nighttime awakenings were defined as present if participants reported each occurring ≥ 3 times per week based on standardized questions. An ESS score[23] (range, 0-24) ≥ 10 was used to define EDS. Restless sleep was present if participants reported "restless" or "very restless" sleep in the past 4 weeks. Finally, average self-reported sleep duration was obtained.

Of these variables, 17 candidate predictors were considered in risk prediction models after significance testing in bivariate comparisons. Figure 1 summarizes the model selection process for the three sex-specific models (female-only, n = 7,363; male-only, n = 4,795; sex × covariate interaction model, n = 12,158) as well as the non-sex-specific model (n = 12,158).

Statistical Analysis

Baseline demographic characteristics, comorbidities, and sleep characteristics were compared according to SA status and sex. Chi-

17 covariates considered and covariate × sex interactions	Female-only model n = 7,363			Male-only model n = 4,795			Sex × covariate Interaction model n = 12,158			Non-sex-specific model n = 12,158
	1[a]	2[b]	Final	1[a]	2[b]	Final	1[a]	2[b]	Final	Final
1. Sex							▓	▓	▓	▓
2. Age	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓
Age × sex							▓	▓	▓	
3. BMI	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓
BMI × sex							▓	▓	▓	
4. Waist circumference	▓			▓			▓			
Waist circumference × sex							▓			
5. Hypertension	▓	▓		▓	▓		▓			
Hypertension × sex							▓			
6. Diabetes	▓			▓			▓			
7. Heart disease	▓	▓		▓			▓	▓		
Heart disease × sex							▓			
8. Average sleep duration	▓						▓			
9. Trouble falling asleep	▓	▓		▓	▓		▓			
Trouble falling asleep × sex							▓			
10. ESS > 10	▓			▓			▓			
11. Snoring	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓
12. Witnessed apneas	▓	▓		▓	▓		▓			
13. Naps	▓			▓			▓			
14. Awakenings at night	▓			▓			▓			
Awakenings at night × sex							▓			
15. Energy level next day	▓			▓			▓			
16. Feeling down	▓						▓			
Feeling down × sex							▓			
17. Restless sleep	▓						▓			
Restless sleep × sex							▓			
Area under the curve	0.845	0.844	0.836	0.818	0.817	0.808	0.844	0.843	0.836	0.832

Figure 1 – The model selection process for the three sex-specific models and the non-sex-specific model. Models began in primary phase (1[a]) in which all predictor variables were considered. The secondary phase (2[b]) refined models by using backward elimination, the C statistic to evaluate discrimination, and the Hosmer-Lemeshow test to evaluate calibration, ending with a "Final" version of each model. Red indicates female-only model; blue indicates male-only model; gray indicates sex × covariate interaction model; orange indicates overall model in which sex is included as a main effect only. ESS = Epworth Sleepiness Scale.

square and Wilcoxon tests were used for significance testing as appropriate. To test our hypothesis (ie, sex-specific prediction models [including models that include sex interaction terms] would outperform a non-sex-specific model for predicting SA), we created and compared the three aforementioned models with a non-sex-specific model.

All continuous variables were examined graphically for linearity, with transformations and piecewise functions considered. Logistic regression was used with backward elimination to identify significant predictors ($P < .20$ to consider, $P < .05$ to stay) for each model. As noted earlier, sex × covariate interaction terms were considered for variables that differed significantly with respect to SA status according to sex: age, BMI, waist circumference, hypertension, heart disease, trouble falling asleep, waking up several times at night, feeling downhearted, and restless sleep. We then systematically left out predictors and evaluated changes in the C statistic to determine each covariate's contribution to model discrimination. The more parsimonious models were selected to develop a clinically useful prediction tool, and calibration was considered (based on the Hosmer-Lemeshow test), prioritizing models with better goodness-of-fit. After the final sets of predictors were selected, 10-fold cross-validation was used to internally validate prediction models.[24] Cross-validation is a model validation procedure to assess the generalizability of a prediction model to an independent dataset. The 10-fold cross-validation method divides the dataset into 10 sets of size n/10, using 9 datasets as "training" datasets and 1 as the "test" dataset; this process was repeated 10 times. Performance statistics were determined within each iteration, and these values were then combined across the 10 sets.

Discrimination was assessed by using the following performance measures: receiver-operating characteristic curves, category-free net reclassification improvement, and integrated discrimination improvement.[25] The specificity and sensitivity of each model were also determined, using the probability cutoff maximizing the sum of these two measures to define SA (AHI ≥ 15 events per hour).

Analyses were performed by using SAS version 9.3 (SAS Institute, Inc) and SUDAAN software Release 11 (RTI International). Cross-validation and performance measures were conducted by using R 2.15.2 (R Foundation for Statistical Computing) with the R package ROCR[26] and in SAS with the %add_predictive macro.[27]

Results

A total of 16,415 participants completed the baseline examination from 2008 to 2011. Of these, 15,277 participants had undergone home sleep testing; 12,531 of these subjects had acceptable sleep study quality (ie, > 30 min of recording). Complete data on all variables were available on 12,158 participants (373 had missing data).

The population-weighted prevalence of SA (AHI ≥ 15 events per hour) was 6.1% (95% CI, 5.3-7.0) in female subjects and 13.5% (95% CI, 12.4-14.8) in male subjects. Table 1 provides baseline demographic characteristics and comorbidities for HCHS/SOL participants with SA, grouped according to sex. Women with SA (AHI ≥ 15 events per hour) were older and had a higher BMI compared with men with SA. Among both male and female subjects, age, BMI, and waist circumference were significantly higher in the SA-positive group compared with the SA-negative group. Similarly, in both men and women, the prevalence of self-reported hypertension, diabetes, and heart disease was significantly higher in SA-positive participants compared with SA-negative participants (data for SA-negative subjects not shown in Table 1). Compared with SA-positive male subjects, SA-positive female subjects had a higher prevalence of symptoms of insomnia (nocturnal awakenings, 51% in women vs 35% in men [$P_{interaction} = .03$]; trouble falling asleep, 33% vs 13% [$P_{interaction} < .0001$]; and restless sleep, 25% vs 14% [$P_{interaction} = .06$]) and depressed mood (feeling downhearted or depressed, 18% in female subjects vs 8% in male subjects [$P_{interaction} = .03$]). Interestingly, the presence of snoring, EDS (ESS ≥ 10), and depression symptoms (CESD ≥ 10) among men and women with SA was not significantly different (68% and 67%, $P_{interaction} = .40$; 23.5% vs 24.4%, $P_{interaction} = .88$; and 20.6% vs 36.7%, $P_{interaction} = .19$, respectively).

Table 2 provides the results for predicting SA from the logistic regression analyses for each model. Female-only and male-only prediction models both consisted of age, BMI, and self-reported snoring. Older age was associated with SA more in women compared with men (OR, 1.09 per year vs 1.06 per year; $P_{interaction} < .0001$), and higher BMI was associated with SA more in men compared with women (OR, 1.19 per kg/m^2 vs 1.13 per kg/m^2; $P_{interaction} < .0001$).

Table 3 provides comparisons of performance indicators among the four prediction models after 10-fold cross-validation. Comparing C statistics (Fig 2), the female-only model (C statistic, 0.836) outperformed both the male-only model (0.808) and the non-sex-specific model (C statistic, 0.832) but performed similarly to the sex × covariate interaction model (C statistic, 0.836). The sex × covariate interaction model performed statistically better than the non-sex-specific model based on sensitivity at the probability cutoff that maximized these measures, area under the curve (C statistic, 0.836 vs 0.832; $P < .0001$), category-free net reclassification improvement (24.3% [95% CI, 18.7-29.9]; $P < .0001$), and integrated discrimination improvement (0.8% [95% CI, 0.5-1.0]; $P < .0001$).

TABLE 1] Baseline Demographic/Sleep Characteristics and Comorbidities for HCHS/SOL Participants With SA, Grouped According to Sex

Variable	Male Subjects (n = 4,795)			Female Subjects (n = 7,363)		
	SA-Positive (n = 807)	SA-Negative (n = 3,988)	P Value	SA-Positive (n = 561)	SA-Negative (n = 6,802)	P Value
Age, y	53 (46-61)	46 (33-55)	< .001	57 (52-64)	48 (37-56)	< .001
BMI, kg/m^2	32 (29-35)	28 (25-31)	< .001	34 (30-39)	29 (26-33)	< .001
Waist circumference, cm	106 (99-115)	96 (89-104)	< .001	107 (99-116)	96 (88-104)	< .001
ESS (range, 0-24)	6 (3-9)	5 (2-9)	< .001	6 (3-9)	5 (2-8)	< .001
ESS score ≥ 10						
Yes	190 (24)	728 (18)	< .001	137 (24)	1,269 (19)	< .001
No	617 (77)	3,260 (82)		424 (76)	5,533 (81)	
CESD-10 score ≥ 10						
Yes	166 (21)	840 (21)	.75	206 (37)	2,278 (34)	.12
No	641 (79)	3,148 (79)		355 (63)	4,524 (67)	
Hypertension[a]						
Yes	378 (47)	949 (24)	< .001	311 (55)	1,846 (27)	< .001
No	429 (53)	3,039 (76)		250 (45)	4,956 (73)	
Diabetes[a]						
Yes	181 (22)	440 (11)	< .001	171 (31)	1,047 (15)	< .001
No	626 (78)	3,548 (89)		390 (70)	5,755 (85)	
History of heart disease (excludes angina)[a]						
Yes	64 (8)	139 (4)	< .001	42 (8)	136 (2)	< .001
No	743 (92)	3,849 (97)		519 (93)	6,666 (98)	
Snoring[a]						
≥ 3 times a week	551 (68)	1,555 (39)	< .001	374 (67)	2,386 (35)	< .001
< 3 times a week	125 (16)	1,333 (33)		72 (13)	2,307 (34)	
Do not know	131 (16)	1,100 (28)		115 (21)	2,109 (31)	
Witnessed apneas[a]						
≥ 3 times a week	163 (20)	282 (7)	< .001	82 (15)	411 (6)	< .001
< 3 times a week	290 (36)	2,082 (52)		179 (32)	3,243 (48)	
Do not know	354 (44)	1,624 (41)		300 (54)	3,148 (46)	
Naps (≥ 5 min) per week[a]						
≥ 3 times a week	257 (32)	1,068 (27)	.003	153 (27)	1,365 (20)	< .001
< 3 times a week	550 (68)	2,920 (73)		408 (73)	5,437 (80)	
Wake up several times at night[a]						
≥ 3 times a week	280 (35)	1,132 (28)	< .001	284 (51)	2,487 (37)	< .001
< 3 times a week	527 (65)	2,856 (72)		277 (49)	4,315 (63)	
Restless sleep in past 4 weeks[a]						
Restless or very restless	116 (14)	591 (15)	.74	140 (25)	1,404 (21)	.02
Very sound, sound, or average	691 (86)	3,397 (85)		421 (75)	5,398 (79)	
Did you have a lot of energy?[a]						
Some, a little, or none of the time	252 (31)	1,008 (25)	< .001	298 (53)	3,033 (45)	< .001
All or most of the time	555 (69)	2,980 (75)		263 (47)	3,769 (55)	

(Continued)

Variable	Male Subjects (n = 4,795)			Female Subjects (n = 7,363)		
	SA-Positive (n = 807)	SA-Negative (n = 3,988)	P Value	SA-Positive (n = 561)	SA-Negative (n = 6,802)	P Value
Have you felt downhearted and depressed?[a]						
All or most of the time	61 (8)	302 (8)	.99	103 (18)	881 (13)	< .001
Some, a little, or none of the time	746 (92)	3,686 (92)		458 (82)	5,921 (87)	
Trouble falling asleep[a]						
≥ 3 times a week	106 (13)	703 (18)	.002	186 (33)	1,811 (27)	< .001
< 3 times a week	701 (87)	3,285 (82)		375 (67)	4,991 (73)	

Data are presented as median (interquartile range) or no. (%). CESD = Center for Epidemiologic Studies Depression Scale; ESS = Epworth Sleepiness Scale; SA = with sleep apnea.
[a]Self-reported variables.

However, because the C statistics were nearly identical in the sex × covariate interaction model and the non-sex-specific model, and because the non-sex-specific model is simpler, we report the non-sex-specific model as our final model, with the prediction equation as follows:

Predicted probability of SA $= 1/[1 + exp$
$[-[-10.2561 + 0.7006 \times (1 \text{ if male, } 0 \text{ if female})$
$+ 0.0655 \times (age) + 0.1391 \times (BMI) + 0.9481$
$\times (1 \text{ if snores } \geq 3 \text{ times a week, } 0 \text{ if otherwise})$
$+ 0.1012 \times (1 \text{ if does not know snoring status,}$
$0 \text{ if otherwise})]]]$

Following is an example of how our prediction model can be applied:

Predicted probability of SA for a 50-year-old man with of 32 kg/m^2 who does not know if he snores $= 1/[1 + \exp[-[-10.2561 + 0.7006 \times 1 + 0.0655 \times 50 + 0.1391 \times 32 + 0.9481 \times 0 + 0.1012 \times 1]]] = 0.151$

Therefore, the predicted probability of SA in this patient is 15.1% (95% CI, 13.2-17.1), with a sensitivity of 0.77 and a specificity of 0.75 using a probability cutoff of ≥12%.

Discussion

We developed and internally validated a simple yet highly discriminating, well-calibrated, and parsimonious prediction model for SA. Using this model, we created English (https://www.montefiore.org/sleepapneariskcalc. html) and Spanish (http://www.montefiore.org/sleepapneariskcalc-es.html) websites in which we display our SA risk calculator. The SA prediction model was created from a community-based sample of US Hispanic/Latino subjects, the largest population-based cohort study to date with subjective and objective SA measurements. Our model consisted of four commonly available variables: self-reported snoring, age, sex, and BMI. The model had a sensitivity of 0.77 and a specificity of 0.75. Contrary to our hypothesis, the same set of variables predicted SA in both sexes in our study sample. The risk prediction model we developed may be a more useful tool for predicting SA in the general population compared with previously published models due to its ease of use (ie, four easily measured variables) and high discrimination.

SA is common in the United States and, despite its high prevalence, it remains vastly underdiagnosed.[28] Furthermore, the majority of published data on SA screening and diagnosis originates from non-Hispanic/Latino study samples, which limits their generalizability. The prevalence of SA and its associated risk factors is common among US Hispanic/Latino subjects.[8] Development of prediction tools for SA in the US Hispanic/Latino population is therefore crucial for identifying high-risk individuals. In addition to the lack of representation of the Hispanic/Latino population in currently available prediction tool source cohorts, these other tools have several limitations: (1) most prediction tools for SA were developed in referral-based samples of symptomatic individuals, (2) most prediction tools require complex measurements (eg, neck, hip circumference) or completion of laborious questionnaires, and (3) sex-related SA symptoms are often not considered in these prediction tools.

TABLE 2] Results From Logistic Regression Analysis for Each Model Predicting Sleep Apnea

Variable	Beta Coefficient	Lower Limit of 95% CI	Upper Limit of 95% CI	P Value
Sex × covariate interaction model, n = 12,158				
Intercept	−11.24	−12.04	−10.43	< .0001
Male (vs female)	1.04	−0.06	2.14	.07
Age (per year)	0.08	0.07	0.09	< .0001
Age × male	−0.03	−0.04	−0.02	< .0001
BMI (per kg/m^2)	0.12	0.10	0.13	< .0001
BMI × male	0.06	0.04	0.08	< .0001
Snoring (vs < 3 nights a week)				
≥ 3 times a week	0.92	0.75	1.10	< .0001
Do not know	0.08	−0.13	0.29	.46
Male-only model, n = 4,795				
Intercept	−10.18	−10.95	−9.42	< .0001
Age (per year)	0.06	0.05	0.06	< .0001
BMI (per kg/m^2)	0.18	0.16	0.19	< .0001
Snoring (vs < 3 nights a week)				
≥ 3 times a week	0.89	0.67	1.11	< .0001
Do not know	0.06	−0.22	0.33	.67
Female-only model, n = 7,363				
Intercept	−11.26	−12.08	−10.45	< .0001
Age (per year)	0.08	0.07	0.09	< .0001
BMI (per kg/m^2)	0.12	0.10	0.13	< .0001
Snoring (vs < 3 nights a week)				
≥ 3 times a week	0.97	0.70	1.24	< .0001
Do not know	0.11	−0.20	0.43	.49
Non-sex-specific model, n = 12,158				
Intercept	−10.26	−10.79	−9.72	< .0001
Male sex (vs female)	0.70	0.63	0.77	< .0001
Age (per year)	0.07	0.06	0.07	< .0001
BMI (per kg/m^2)	0.14	0.13	0.15	< .0001
Snoring (vs < 3 nights a week)				
≥ 3 times a week	0.95	0.78	1.12	< .0001
Do not know	0.10	−0.11	0.31	.34

Unlike existing prediction models, our models systematically examined the impact of sex-related differences in SA prediction. A population-based study[29] conducted in 1990 examined the prevalence of SA symptoms, in particular snoring, in a Hispanic/Latino-American population. In this study of 1,222 individuals, regular loud snoring was more common in male subjects (27.8%) than in female subjects (15.3%). However, in our sample of community-dwelling US Hispanic/Latino subjects, we found no difference in self-reported snoring between the two sexes. We also found that the same set of variables predict SA in both sexes. This outcome was contrary to our hypothesis. Despite finding differences in the prevalence of insomnia and depression symptoms among men and women with SA, these symptoms did not result in improved SA prediction in either sex. This finding may be partly due to their expected high correlation with other predictor variables such as age and BMI. Another possible explanation is the increasing prevalence of obesity over the past few decades, which could be influencing sex-specific differences in SA. Finally, evidence that SA symptoms may differ according to sex originates from studies that consisted predominantly of non-Hispanic/Latino

TABLE 3] Comparisons of Performance Measures Among the Four Prediction Models After 10-Fold Cross-Validation

Model (Covariates Included)	Performance Measures				
	Training Set C Statistic	Test Set C Statistic (10-Fold Validation)	Probability at Which Sensitivity and Specificity Are Maximized, %	Sensitivity	Specificity
Female-only (age, BMI, snoring)	0.84	0.84	≥ 7.0	0.82	0.70
Male-only (age, BMI, snoring)	0.81	0.81	≥ 15.7	0.78	0.70
Sex × covariate interaction model (age, age × sex, BMI, BMI × sex, snoring)	0.84	0.84	≥ 11.3	0.79	0.73
Non-sex-specific model (age, BMI, snoring, sex)	0.83	0.83	≥ 12.0	0.77	0.75

individuals. Therefore, our findings could be partially explained by the possibility that sex-specific differences in reporting of SA symptoms are cultural.

Several questionnaire-based prediction tools exist in the current literature. For example, the Berlin Questionnaire captures symptoms of snoring, wake-time sleepiness or fatigue, and history of hypertension and obesity.[30] It was tested in 744 adults surveyed in five primary care sites in Cleveland, Ohio. Of the 744 surveys, only 100 had objective sleep data (from portable home sleep monitoring). Being in the high-risk group predicted presence of mild SA with a sensitivity of 0.86 and a specificity of 0.77. However, a high-risk score predicted moderate and severe SA, with a sensitivity of 0.54 and a specificity of 0.97. These findings suggest that the Berlin Questionnaire is an excellent screening tool for mild SA but is less sensitive for moderate to severe SA. Our prediction model, conversely, has a much higher sensitivity (0.77) for predicting moderate to severe SA without compromising specificity (0.75). We considered our model performance by using an AHI cutoff of 5 events per hour for mild SA and found that our model performs equally well (C statistic of 0.817 vs 0.833 using an AHI cut point of ≥ 15 events per hour).

The ESS[31] is another questionnaire-based prediction tool commonly used in clinical practice to assess for EDS. It has been examined for its ability to predict underlying SA and has been found to correlate well with the AHI.[23] However, the ESS did not improve the predictive ability of our models.

Limitations of the present study need to be considered. First, SA symptoms were captured by using a questionnaire and are therefore susceptible to

information bias. Self-reported snoring may be difficult to report because symptoms occur during sleep; nonetheless, this symptom is commonly used in clinical assessments and research. Second, the HCHS/SOL assessed SA with an in-home sleep study as opposed to in-laboratory, full-night, attended polysomnography. This approach may modestly underestimate SA severity and does not allow assessment of sleep stages, duration, or fragmentation. Furthermore, although we assessed the internal validity of our prediction model, we have not assessed the external validity of our prediction models. Although there are clear merits to examining sleep in the understudied US Hispanic/Latino population and,
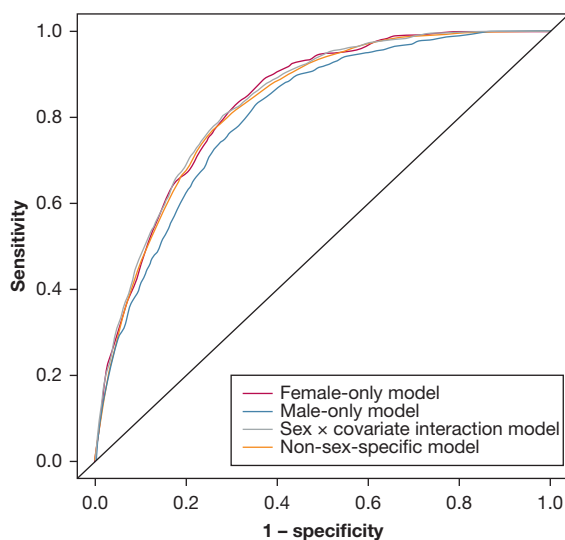

Figure 2 – Receiver-operating characteristic curves for sleep apnea prediction models displaying the trade-off between sensitivity and specificity for each model. Visual comparison of performance of the four models: (1) female-only, (2) male-only, (3) sex × covariate interaction model, and (4) non-sex-specific model.

accordingly, to developing prediction models for SA in this population (as we have), the generalizability of our prediction models to non-Hispanic/Latino subjects remains unclear and needs further investigation. We therefore plan to externally validate our prediction models in other population-based cohort studies. Finally, population-based prediction rules are useful in that they can identify SA among those less likely to be referred for sleep testing. However, such models may perform differently among patients already referred to specialty sleep clinics, at which the disease prevalence can vary significantly, resulting in underestimation or overestimation of the probability of disease. This scenario was illustrated by use of the Berlin Questionnaire, which performs relatively poorly in sleep clinic settings compared with primary care settings.[32]

Conclusions

We developed and internally validated a highly discriminating, well-calibrated, and parsimonious prediction model for SA in a community-based sample of US Hispanic/Latino subject. This model consisted of four commonly available variables (snoring, age, sex, and BMI).

References

1. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol.* 2013;177(9):1006-1014.

2. Morgenthaler TI, Croft JB, Dort LC, Loeding LD, Mullington JM, Thomas SM. Development of the National Healthy Sleep Awareness Project Sleep Health Surveillance questions. *J Clin Sleep Med.* 2015;11(9):1057-1062.

3. Redline S, Kump K, Tishler PV, Browner I, Ferrette V. Gender differences in sleep disordered breathing in a community-based sample. *Am J Respir Crit Care Med.* 1994;149(3 pt 1):722-726.

4. Redline S, Strohl KP. Recognition and consequences of obstructive sleep apnea hypopnea syndrome. *Otolaryngol Clin North Am.* 1999;32(2):303-331.

5. Chervin RD. Sleepiness, fatigue, tiredness, and lack of energy in obstructive sleep apnea. *Chest.* 2000;118(2):372-379.

6. Pillar G, Lavie P. Psychiatric symptoms in sleep apnea syndrome: effects of gender and respiratory disturbance index. *Chest.* 1998;114(3):697-703.

7. Bixler EO, Vgontzas AN, Lin HM, et al. Prevalence of sleep-disordered breathing in women: effects of gender. *Am J Respir Crit Care Med.* 2001;163(3 pt 1):608-613.

8. 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):1230-1235.

9. Young T, Evans L, Finn L, Palta M. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep.* 1997;20(9):705-706.

10. Anderson KM, Wilson PW, Odell PM, Kannel WB. An updated coronary risk profile. A statement for health professionals. *Circulation.* 1991;83(1):356-362.

11. Lloyd-Jones DM, Wilson PW, Larson MG, et al. Framingham risk score and prediction of lifetime risk for coronary heart disease. *Am J Cardiol.* 2004;94(1):20-24.

12. Lavange LM, Kalsbeek WD, Sorlie PD, et al. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol.* 2010;20(8):642-649.

13. Sorlie PD, Aviles-Santa LM, Wassertheil-Smoller S, et al. Design and implementation of the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol.* 2010;20(8):629-641.

14. Lind BK, Goodwin JL, Hill JG, Ali T, Redline S, Quan SF. Recruitment of healthy adults into a study of overnight sleep monitoring in the home: experience of the Sleep Heart Health Study. *Sleep Breath.* 2003;7(1):13-24.

15. Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. *Sleep.* 1991;14(6):540-545.

16. Ayappa I, Norman RG, Seelall V, Rapoport DM. Validation of a self-applied unattended monitor for sleep disordered breathing. *J Clin Sleep Med.* 2008;4(1):26-37.

17. Berry RB, Budhiraja R, Gottlieb DJ, et al. Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med.* 2012;8(5):597-619.

18. Grigg-Damberger MM. The AASM Scoring Manual four years later. *J Clin Sleep Med.* 2012;8(3):323-332.

19. Westbrook PR, Levendowski DJ, Cvetinovic M, et al. Description and validation of the apnea risk evaluation system: a novel method to diagnose sleep apnea-hypopnea in the home. *Chest.* 2005;128(4):2166-2175.

20. Shahar E, Whitney C, Redline S, et al. Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study. *Am J Respir Crit Care Med.* 2001;163(1):19-25.

21. Peppard P, Young T, Palta M, Skatrud J. Prospective study of the association between sleep-disordered breathing and hypertension. *N Engl J Med.* 2000;342(19):1378-1384.

22. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1:385-401.

23. Cao J, Chen B, Dong L, Guo M, Wang Y, Yu M. The primary diagnostic significance of the Epworth Sleepiness Scale in patients

with obstructive sleep apnea syndrome [in Chinese]. *Zhonghua Jie He He Hu Xi Za Zhi*. 2002;25(3):154-155.

24. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15(4):361-387.

25. Pencina MJ, Sr D'Agostino RB, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27(2):157-172.

26. Sing T, Sander O, Beerenwinkel N, Lengauer. T. ROCR: visualizing classifier performance in R. *Bioinformatics*. 2005; 21(20):3940-3941.

27. Kennedy KF, Pencina MJ. A SAS macro to compute added predictive ability of new markers predicting a dichotomous outcome. SESUG 2010: The Proceedings of the SouthEast SAS Users Group. Savannah, GA: 2010. http://analytics.ncsu.edu/sesug/2010/SDA07.Kennedy.pdf.

28. Kapur V, Strohl KP, Redline S, Iber C, O'Connor G, Nieto J. Underdiagnosis of sleep apnea syndrome in U.S. communities. *Sleep Breath*. 2002;6(2):49-54.

29. Schmidt-Nowara WW, Coultas DB, Wiggins C, Skipper BE, Samet JM. Snoring in a Hispanic-American population. Risk factors and association with hypertension and other morbidity. *Arch Intern Med*. 1990;150(3): 597-601.

30. 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):485-491.

31. Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. *Sleep*. 1992;15(4):376-381.

32. Karakoc O, Akcam T, Genc H, Yetkin S, Piskin B, Gerek M. Use of the Berlin Questionnaire to screen at-risk patients for obstructive sleep apnea. *B-ENT*. 2014;10(1):21-25.