Published in final edited form as:

Sleep Med. 2020 November ; 75: 1–6. doi:10.1016/j.sleep.2020.05.004.

Prediction of sleep-disordered breathing after stroke

Devin L. Brown, MD¹, Kevin He, PhD², Sehee Kim, PhD², Chia-Wei Hsu, MS³, Erin Case, BA^{1,3}, Ronald D. Chervin, MD,MS⁴, Lynda D. Lisabeth, PhD^{1,3}

¹Stroke Program, University of Michigan

²Department of Biostatistics, School of Public Health, University of Michigan

³Department of Epidemiology, School of Public Health, University of Michigan

⁴Sleep Disorders Center and Department of Neurology, University of Michigan

Abstract

Objective/Background: Sleep-disordered breathing (SDB) is highly prevalent after stroke and is associated with poor outcomes. Currently, after stroke, objective testing must be used to differentiate patients with and without SDB. Within a large, population-based study, we evaluated the usefulness of a flexible statistical model based on baseline characteristics to predict post-stroke SDB.

Patients/Methods: Within a population-based study, participants (2010–2018) underwent SDB screening, shortly after ischemic stroke, with a home sleep apnea test. The respiratory event index

Correspondence: Devin L. Brown, MD, MS, 1500 East Medical Center Drive, Ann Arbor, Michigan, 48109-5855, Phone (734) 936-9075, Fax (734) 232-4447, devinb@umich.edu.

Devin L. Brown, MD, MS: designed the study, obtained funding, interpreted data, wrote the initial draft of the manuscript

Kevin He, PhD: designed analysis plan, interpreted data, performed statistical analyses, edited the manuscript for important intellectual content

Sehee Kim, PhD: designed analysis plan, interpreted data, performed statistical analyses, edited the manuscript for important intellectual content

Chia-Wei Hsu, MS: interpreted data, performed statistical analyses, edited the manuscript for important intellectual content

Erin Case, BA: obtained the data, edited the manuscript for important intellectual content

Ronald D. Chervin, MD, MD: interpreted data, edited the manuscript for important intellectual content

Lynda D. Lisabeth, PhD: designed the study, obtained funding, interpreted data, edited the manuscript for important intellectual content

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of interest/Disclosures:

Brown: received funding from R01HL126700, R01HL123379, R01 NS070941, R01 HL098065

He: None.

Kim: R01HL126700, R01NS091112, R01NS038916, R01DK070869, U01DK062456, HHSM-500-2016-RFP-0039

Hsu: None.

Case: has received funding from R01NS091112, R01 NS070941, R01NS038916, R01HL126700

Chervin: received research grant funding from the NIH (R01HL126700, R01HL123379, R01 NS070941, R01 HL098065, R01 HL105999, R43 HL117421, T32HL110952, R01HD082129, U01HL125295 and U01NS099043). He has consulted for Zansors; serves as an editor and author for UpToDate; edited a book published by Cambridge University Press; and has produced copyrighted material, patents, and patents pending, owned by the University of Michigan, focused on assessment or treatment of sleep disorders. He has served on the Boards of Directors for the American Academy of Sleep Medicine, Associated Professional Sleep Societies, International Pediatric Sleep Association, and the non-profit Sweet Dreamzzz.

Lisabeth: received research grant funding from the NIH (R01NS038916, R01NS107463, R01HL126700, R01HL123379, R21AG060277)

(REI) was calculated as the number of apneas and hypopneas per hour of recording; values ≥ 10 defined SDB. The distributed random forest classifier (a machine learning technique) was applied to predict SDB with the following as predictors: demographics, stroke risk factors, stroke severity (NIHSS), neck and waist circumference, palate position, and pre-stroke symptoms of snoring, apneas, and sleepiness.

Results: Within the total sample ($n=1,330$), median age was 65 years; 47% were women; 32% non-Hispanic white, 62% Mexican American, and 6% African American. SDB was found in 891 (67%). The area under the receiver operating characteristic curve, a measure of predictive ability, applied to the validation sample was 0.75 for the random forest model. Random forest correctly classified 72.5% of validation samples.

Conclusions: In this large, ethnically diverse, population-based sample of ischemic stroke patients, prediction models based on baseline characteristics and clinical measures showed fair rather than clinically reliable performance, even with use of advanced machine learning techniques. Results suggest that objective tests are still needed to differentiate ischemic stroke patients with and without SDB.

Keywords

Sleep-disordered breathing; stroke; sleep apnea

Introduction

Sleep-disordered breathing (SDB) is highly prevalent after stroke[1] and is associated with poor recovery and an increased risk of recurrent stroke and mortality.[2, 3] Interestingly, SDB characteristics differ between patients with stroke and those in the general population. For example, excessive daytime sleepiness and obesity are not highly characteristic of post-stroke SDB.[4] Thus, SDB screening questionnaires that perform well in the general population do not perform well after stroke.[5] A recent systematic review of this topic suggested that instruments often have good sensitivity but poor specificity, with overall performance typically in the poor or fair range.[6] The previous studies were limited by relatively small sample size, select study populations (from a clinical trial or academic medical center), sometimes retrospective design, exclusion of common measurements such as palate position, and absence of sophisticated statistical techniques.

Objective assessment of SDB in the acute stroke setting is expensive and a logistical challenge; therefore, a tool to aid in SDB identification could be useful. Within a large, population-based study, we tested the performance of a statistical model derived from machine learning approaches, based on easily-obtained baseline characteristics and simple measurements, in prediction of post-stroke SDB. The purpose of the study was to (1) create a model that performs well to identify post-stroke patients with SDB, (2) identify variables most influential in predictive models, and (3) compare model results based on traditional statistical methods and machine learning approaches.

Materials and Methods

The data were collected from the Brain Attack Surveillance in Corpus Christi (BASIC) Sleep Apnea Project. BASIC is an ongoing stroke surveillance study that uses active and passive surveillance to identify all strokes from the 7 acute care hospitals in Nueces County. Detailed methods have been published.[7, 8] To be eligible for BASIC, patients must have a stroke, live in Nueces County 6 months per year, and be age 45 years. To be eligible for the SDB testing, patients additionally must be identified within 30 days if through active surveillance and 45 days if through passive surveillance. Exclusions for SDB testing include current pregnancy, and current use of oxygen supplementation or positive airway pressure. All stroke cases are validated using source documentation by stroke-fellowship trained investigators. Only subjects with ischemic strokes (2010–2018), defined by clinical criteria, [9, 10] were included in the current analysis. Race/ethnicities other than non-Hispanic white, Mexican American, and African American were excluded due to small numbers. Institutional Review Board approval was obtained from the University of Michigan and the two Corpus Christi hospital systems. Written informed consent was obtained from subjects or their surrogates.

Participants in BASIC undergo a baseline interview as soon as possible after stroke. The interview includes administration of the Berlin questionnaire,[11] a screen for symptoms and signs of obstructive sleep apnea, in reference to the pre-stroke state. Baseline characteristics, including demographics, stroke risk factors, and clinical variables, are abstracted from the medical record.

Objective SDB evaluation is undertaken with the ApneaLink Plus portable home sleep apnea test (HSAT) either during the hospitalization or at home. This device is well validated[12–15] and HSATs have been validated in stroke patients.[16–19] The study includes measurements of nasal pressure, oxygen saturation, pulse, and respiratory effort. After input of raw data into the ApneaLink software program, a registered sleep technologist edits the start and stop times and artifacts. Detailed methods of this process and definitions of apneas and hypopneas have been published.[20] The respiratory event index (REI) is calculated from the sum of apneas and hypopneas per hour of recording. (This measure is similar to the apnea-hypopnea index (AHI) although the AHI is per hour of sleep.) In the current analyses, an REI 10 was considered to define SDB, but other commonly used cutoffs of 5, 15, and 30 were also applied. Friedman palate position,[21] neck circumference, and waist circumference are assessed by study coordinators at the time of the ApneaLink testing.

Statistical analysis:

Baseline characteristics were compared by SDB status using chi square or Wilcoxon Rank Sum tests as appropriate. Models to predict SDB (REI 10) included the following covariates: demographics (age, sex, race/ethnicity), hypertension, diabetes, atrial fibrillation, smoking, prior TIA/stroke, excessive alcohol use, congestive heart failure, coronary artery disease, high cholesterol, body mass index (BMI), initial stroke severity (National Institutes of Health Stroke scale (NIHSS)), neck and waist circumference, Friedman palate position, [21] and pre-stroke symptoms from the Berlin questionnaire[11] of snoring (yes vs no or

don't know), witnessed apneas (almost every day or 3–4 times per week vs all others), and sleepiness (almost every day or 3–4 times per week vs all others). For the primary analysis, machine learning techniques (Random Forests, Boosted Regression Models, XGBoost, Deep Learning and Stacked Ensembles) were applied using the R package h2o, where missing values were treated as a separate category by default. An attractive feature of the machine learning techniques, compared to the logistic regression, is that these approaches do not require (or are less reliant on) a specification of structures for the association between the outcome variable and predictors. To determine the machine learning method with the best performance, the area under the receiver operating characteristics curve (AUC) was calculated. The highest AUC (with REI 10 as the outcome variable) was obtained with the distributed random forest; therefore, individual variable importance was reported to understand which variables most contributed to the prediction decision. The individual importance accounts for both whether the variable was selected during the model building process and its relative influence. Each predictor's importance was scaled relative to the best performing variable. The variable with the highest influences on prediction was scored 1, and all other variables had lower scores ranging downwards toward zero.

To understand the utility of using a machine learning approach, we compared the predictive ability of the machine learning model with the highest AUC to the results from a logistic regression model including main effects only. Preliminary analyses to assess potential non-linear associations between continuous predictors and the log of the odds of having SDB were conducted using regression splines. As a result, linear splines were fitted for neck circumference and a log transformation was performed for NIHSS to account for its skewed distribution. To examine improvement led by consideration of interaction effects, we compared prediction performances resulting from two different logistic regression models: including marginal effects only and including both main and two-way interaction effects. Inclusion of the two-way interaction terms were determined by a p-value <0.1 using a stepwise selection procedure. In this procedure, after each added interaction, all interactions already selected in the model were checked and eliminated if $p > 0.1$. In the logistic models, the mean was substituted for missing covariate values (mean imputation).

The performance of each predictive model was evaluated based on 10-fold cross-validation. Machine learning models were also repeated for three alternative SDB severity cutoffs: REI 5, REI 15, and REI 30.

Results

Within the total analytic sample ($n=1,330$), median age was 65 years (inter-quartile range (IQR): 57, 74); 47% were women; 32% non-Hispanic white, 62% Mexican American, and 6% African American. The median time from stroke recognition to the ApneaLink Plus study was 14 days (IQR: 6, 21). SDB (defined as REI 10) was found in 67% of the sample. REI 5 was found in 88%; REI 15, in 51%; and REI 30, in 22%. Baseline characteristics by SDB status are found in Table 1. SDB differed across race/ethnic groups ($p<0.0001$): SDB was found in 591/819 (72%) of Mexican Americans, 258/432 (60%) of non-Hispanic whites, and 42/79 (53%) of African Americans. SDB was also more common in men (517/700, 74%) than women (358/360, 57%, $p<0.0001$). Waist and neck circumference, BMI, palate

position, hypertension, diabetes, coronary disease, prior history of stroke/TIA, and pre-stroke snoring and witnessed apneas were all associated with SDB (Table 1). Moreover, the correlation coefficient between waist and neck circumference was 0.55 ($p<0.0001$); between waist circumference and BMI, 0.73 ($p<0.0001$); and between neck circumference and BMI, 0.47 ($p<0.0001$).

Machine learning:

The highest AUC of the machine learning algorithms was for the Distributed Random Forest model. This approach resulted in an AUC of 0.75. The Distributed Random Forest correctly classified 72.5% of validation samples. The other machine learning techniques resulted in AUCs that ranged from 0.68–0.73. The next best performing machine learning techniques included the Random Forest, Stacked Ensemble Model, Deep Learning, and Generalized Boosted Regression Model. The most important features in the Distributed Random Forest model for predicting SDB were neck circumference, BMI, waist circumference, age, NIHSS, and pre-stroke daytime sleepiness (Figure). The AUCs for alternative thresholds of SDB demonstrated similar results. The machine learning algorithm with the highest AUCs was again the Distributed Random Forest model: for REI 5, the AUC was 0.74; for REI 15, the AUC was 0.75; and for REI 30, the AUC was 0.75.

Logistic regression:

The results of the main effects adjusted logistic regression model are shown in Table 2; the AUC was 0.70 and the model correctly classified 69.7% of validation samples. A significant non-linear association ($p=0.0048$ in the preliminary analysis using regression splines model) was found between neck circumference and the log odds of SDB. Specifically, a negative association was found for neck circumference less than 30 cm, while a positive association was shown for neck circumference greater than 30 cm. Therefore, a knot to allow different regression slopes was placed at the value of 30 cm. Age, Mexican American ethnicity, hypertension, stroke/TIA history, NIHSS, and neck circumference were significantly associated with SDB. In addition, the stepwise procedure selected 52 two-way interactions with p -values less than 0.1. The corresponding AUC for the model including these interactions was 0.71, which was not significantly different from the main effects model ($p=0.2194$). The logistic regression model with interactions correctly classified 71.4% of validation samples. AUCs from the logistic regression models (main effects and including interactions) were significantly lower than that from the Distributed Random Forest model ($p<0.0001$ for main effects and $p=0.0022$ for model with interactions).

Discussion

This large, population-based, multicenter, prospective study shows that despite application of machine learning techniques and inclusion of a wide array of baseline characteristics and measurements, model-based prediction of post-stroke SDB remains only fair. Therefore, to identify the highly prevalent condition of post-stroke SDB in clinical practice or for research purposes, objective testing is required. Questionnaires are often recommended for use in clinical practice to assess for SDB after stroke or implemented in research of SDB after stroke.[22–25] Models based on the variables we assessed do not seem likely to provide a

worthwhile screening process in post-stroke patients. Although physiologic testing is more time-consuming, requires more resources and more expense, it remains necessary for accurate SDB determination after stroke.

Other SDB inventories and prediction models such as the Berlin questionnaire, Epworth Sleepiness Scale, STOP-BANG and its variations, SLEEP INventory, and Sleep Apnea Clinical Score also have had poor to fair performance in stroke patients.[5, 6] Our prediction model, despite inclusion of demographic, clinical, anthropometric, and questionnaire data, and application of sophisticated statistical methods, did not perform substantially better. The best machine learning approach did out-perform logistic regression, but the difference in performance was small.

In the adjusted machine learning model, the most influential variables were neck circumference, BMI, waist circumference, age, and NIHSS. Sex, snoring, and witnessed apneas, classically associated with SDB in the general population, were less influential than the anthropometric measurements. Neurologic deficit measurements, such as the NIHSS, have not been shown to be associated with post-stroke SDB in prior studies, but due to sample size these studies were unable to accommodate covariate adjustment.[26, 27] This may have masked actual differences. Waist and neck circumference have not been well studied, including examination of non-linear relationships, as post-stroke SDB predictors. Interestingly, all three of the anthropometric variables remained influential despite adjustment for the other two, suggesting that each captures unique information, including an unexpected threshold effect related to neck circumference. Differences in predictors between the logistic regression and machine learning approaches may be due to the flexibility of machine learning approaches to incorporate non-linear associations and interactions. These findings highlight the importance of these flexible approaches.

The strengths of this study include the large sample size, population-based design, racial/ethnic diversity, well characterized study population, inclusion of anthropometrics and palate position often ignored post-stroke, internal model validation, and application of sophisticated statistical and machine learning techniques. Limitations include the use of an HSAT rather than the gold standard, in-laboratory polysomnography. Nocturnal periods of wakefulness, which may be more common during hospitalizations, could have contributed to an underestimation of REI. However, most research in post-stroke SDB has not used polysomnography,[3, 28–32] the ApneaLink Plus is commonly used and well-validated,[12–15] and HSATs have been validated in the post-stroke setting.[16–19] Ischemic stroke subtype and infarction location were not included. However, neither is strongly associated with post-stroke SDB and we did account for stroke severity.[33, 34]

Conclusions

This study demonstrated that despite the variety of variables included in a predictive model that used sophisticated statistical techniques within a large sample size, model performance was inadequate to identify SDB with high accuracy. These results support the need to use objective sleep apnea tests to identify stroke patients with SDB, and do not suggest that a simple screen is likely to limit the number of subjects who require testing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements and Sources of Funding:

The project was funded by the NIH (R01 HL098065, R01 NS070941, R01 HL123379, R01 HL126700). The funding source played no role in the decision to submit this analysis for consideration or in the interpretation of data. This study was performed in the Corpus Christi Medical Center and CHRISTUS Spohn hospitals, CHRISTUS Health system, in Corpus Christi, Texas.

References

- [1]. Seiler A, Camilo M, Korostovtseva L, Haynes AG, Brill AK, Horvath T, et al. Prevalence of sleep-disordered breathing after stroke and TIA: A meta-analysis. *Neurology*. 2019;92:e648–e654. [PubMed: 30635478]
- [2]. Brown DL, Shafie-Khorassani F, Kim S, Chervin RD, Case E, Morgenstern LB, et al. Sleep-Disordered Breathing Is Associated With Recurrent Ischemic Stroke. *Stroke*. 2019;50:571–576. [PubMed: 30744545]
- [3]. Turkington PM, Allgar V, Bamford J, Wanklyn P, Elliott MW. Effect of upper airway obstruction in acute stroke on functional outcome at 6 months. *Thorax*. 2004;59:367–371. [PubMed: 15115859]
- [4]. Arzt M, Young T, Peppard PE, Finn L, Ryan CM, Bayley M, et al. Dissociation of Obstructive Sleep Apnea From Hypersomnolence and Obesity in Patients With Stroke. *Stroke*. 2010;41:e129–e134. [PubMed: 20075361]
- [5]. Sico J, Yaggi H, Ofner S, Concato J, Austin C, Ferguson J, et al. Development, Validation, and Assessment of an Ischemic Stroke or Transient Ischemic Attack-Specific Prediction Tool for Obstructive Sleep Apnea. *J Stroke Cerebrovasc Dis*. 2017;26:1745–1754. [PubMed: 28416405]
- [6]. Takala M, Puustinen J, Rauhalu E, Holm A. Pre-screening of sleep-disordered breathing after stroke: A systematic review. *Brain Behav*. 2018;8:e01146–e01146. [PubMed: 30371010]
- [7]. Morgenstern LB, Smith MA, Sanchez BN, Brown DL, Zahuranec DB, Garcia N, et al. Persistent ischemic stroke disparities despite declining incidence in Mexican Americans. *Ann Neurol*. 2013;74:778–785. [PubMed: 23868398]
- [8]. Morgenstern LB, Smith MA, Lisabeth LD, Risser JM, Uchino K, Garcia N, et al. Excess stroke in Mexican Americans compared with non-Hispanic Whites: the Brain Attack Surveillance in Corpus Christi Project. *Am J Epidemiol*. 2004;160:376–383. [PubMed: 15286023]
- [9]. Asplund K, Tuomilehto J, Stegmayr B, Wester P, Tunstall-Pedoe H. Diagnostic criteria and quality control of the registration of stroke events in the MONICA project. *Acta Med Scand Suppl*. 1988;728:26–39. [PubMed: 3202029]
- [10]. Gillum R, Fortmann S, Prineas R, Kottke T. International diagnostic criteria for acute myocardial infarction and acute stroke. *Am Heart J*. 1984;108:150–158. [PubMed: 6731265]
- [11]. Netzer NC, Stoohs R, Netzer CM, Clark K, Strohl KP. Using the Berlin questionnaire to identify patients at risk for the sleep apnea syndrome. *Ann Intern Med*. 1999;131:485–491. [PubMed: 10507956]
- [12]. Chen H, Lowe A, Bai Y, Hamilton P, Fleetham J, Almeida F. Evaluation of a portable recording device (ApneaLink™) for case selection of obstructive sleep apnea. *Sleep Breath*. 2009;13:213–219. [PubMed: 19052790]
- [13]. Ng S, Chan T, To K, Ngai J, Tung A, Ko F, et al. Validation of a Portable Recording Device (ApneaLink) for identifying Patients with Suspected Obstructive Sleep Apnea Syndrome (OSAS). *Intern Med J*. 2009;39:757–762. [PubMed: 19220528]
- [14]. Nigro C, Serrano F, Aimaretti S, González S, Codinardo C, Rhodius E. Utility of ApneaLink for the diagnosis of sleep apnea-hypopnea syndrome. *Medicina*. 2010;70:53–59. [PubMed: 20228025]

- [15]. ResMed I ApneaLink™ Plus White Paper (D2231–127). https://www.resmedcom/au/dam/documents/articles/ApneaLink_Plus_White_Paperpdf2009.
- [16]. Chernyshev OY, McCarty DE, Moul DE, Liendo C, Caldito GC, Munjampalli SK, et al. A pilot study: portable out-of-center sleep testing as an early sleep apnea screening tool in acute ischemic stroke. *Nat Sci Sleep*. 2015;7:127–138. [PubMed: 26527904]
- [17]. Parra O, Abroix A, Bechich S, García-Eroles L, Montserrat J, López J, et al. Time course of sleep-related breathing disorders in first-ever stroke or transient ischemic attack. *Am J Respir Crit Care Med*. 2000;161:375–380. [PubMed: 10673174]
- [18]. Saletu M, Kotzian S, Schwarzsinger A, Haider S, Spatt J, Saletu B. Home Sleep Apnea Testing is a Feasible and Accurate Method to Diagnose Obstructive Sleep Apnea in Stroke Patients During In-Hospital Rehabilitation. *J Clin Sleep Med*. 2018;14:1495–1501. [PubMed: 30176970]
- [19]. Veasey S, Rosen I. Obstructive Sleep Apnea in Adults. *N Engl J Med*. 2019;380:1442–1449. [PubMed: 30970189]
- [20]. Brown D, McDermott M, Mowla A, De Lott L, Morgenstern L, Kerber K, et al. Brainstem infarction and sleep-disordered breathing in the BASIC Sleep Apnea Study. *Sleep Med*. 2014;15:887–891. [PubMed: 24916097]
- [21]. Friedman M, Ibrahim H, Bass L. Clinical staging for sleep-disordered breathing. *Otolaryngol Head Neck Surg*. 2002;127:13–21. [PubMed: 12161725]
- [22]. Mansukhani MP, Bellolio MF, Kolla BP, Enduri S, Somers VK, Stead LG. Worse outcome after stroke in patients with obstructive sleep apnea: an observational cohort study. *J Stroke Cerebrovasc Dis*. 2011;20:401–405. [PubMed: 20656506]
- [23]. Katzan IL, Thompson NR, Uchino K, Foldvary-Schaefer N. A screening tool for obstructive sleep apnea in cerebrovascular patients. *Sleep medicine*. 2016;21:70–76. [PubMed: 27448475]
- [24]. Camilo MR, Sander HH, Eckeli AL, Fernandes RMF, Dos Santos-Pontelli TEG, Leite JP, et al. SOS score: an optimized score to screen acute stroke patients for obstructive sleep apnea. *Sleep medicine*. 2014;15:1021–1024. [PubMed: 24934142]
- [25]. Swartz RH, Cayley ML, Lanctôt KL, Murray BJ, Cohen A, Thorpe KE, et al. The “DOC” screen: Feasible and valid screening for depression, Obstructive Sleep Apnea (OSA) and cognitive impairment in stroke prevention clinics. *PloS one*. 2017;12:e0174451–e0174451. [PubMed: 28376127]
- [26]. Bassetti C, Aldrich M. Sleep apnea in acute cerebrovascular diseases: final report on 128 patients. *Sleep*. 1999;22:217–223. [PubMed: 10201066]
- [27]. Lisabeth L, Sánchez B, Chervin R, Morgenstern L, Zahuranec D, Tower S, et al. High Prevalence of Post-stroke Sleep Disordered Breathing in Mexican Americans. *Sleep Med*. 2017;33:97–102. [PubMed: 28449915]
- [28]. Bassetti CL, Milanova M, Gugger M. Sleep-Disordered Breathing and Acute Ischemic Stroke: Diagnosis, Risk Factors, Treatment, Evolution, and Long-Term Clinical Outcome. *Stroke*. 2006;37:967–972. [PubMed: 16543515]
- [29]. Broadley SA, Jorgensen L, Cheek A, Salonikis S, Taylor J, Thompson PD, et al. Early investigation and treatment of obstructive sleep apnoea after acute stroke. *J Clin Neurosci*. 2007;14:328–333. [PubMed: 16790349]
- [30]. Harbison J, Ford GA, James OF, Gibson GJ. Sleep-disordered breathing following acute stroke. *QJM*. 2002;95:741–747. [PubMed: 12391386]
- [31]. Parra O, Arboix A, Montserrat JM, Quinto L, Bechich S, Garcia-Eroles L. Sleep-related breathing disorders: impact on mortality of cerebrovascular disease. *Eur Respir J*. 2004;24:267–272. [PubMed: 15332396]
- [32]. Ahn S, Kim J, Kim D, Choo I, Lee H, Kim H. Interaction between Sleep-Disordered Breathing and Acute Ischemic Stroke. *J Clin Neurol*. 2013;9:9–13. [PubMed: 23346154]
- [33]. Iranzo A, Santamaria J, Berenguer J. Prevalence and clinical importance of sleep apnea in the first night after cerebral infarction. *Neurology*. 2002;58:911–916. [PubMed: 11914407]
- [34]. Brown D, Mowla A, McDermott M, Morgenstern L, Hegeman G, Smith M, et al. Ischemic stroke subtype and presence of sleep-disordered breathing: the BASIC Sleep Apnea Study. *J Stroke Cerebrovasc Dis*. 2015;24:388–393. [PubMed: 25497720]

Highlights:

- Model-based prediction of post-stroke SDB based on baseline factors is only fair.
- Objective testing is needed to identify post-stroke SDB accurately.
- Machine learning out-performed logistic regression in predicting post-stroke SDB.

Figure.

Variable importance plot for the Distributed Random Forest (DRF) model. Scale 0–1, the relative importance of each variable considered, scaled relative to the best performing variable.

Table 1:Baseline characteristics by SDB status (REI ≥ 10).

	Median (IQR) or Number (column %)		p-value
Covariates	REI<10 (N = 439; 33.0%)	REI $\geq$ 10 (N = 891; 67.0%)	
Age (years)	65 (56, 73)	65 (58, 75)	0.0681
Sex			<.0001
Male	183 (26.1%)	517 (73.9%)	
Female	256 (41.7%)	358 (58.3%)	
Race/ethnicity			<.0001
Non-Hispanic white	174 (40.3%)	258 (59.7%)	
Mexican American	228 (27.8%)	591 (72.2%)	
African American/Black	37 (46.8%)	42 (53.2%)	
Hypertension			<.0001
No	107 (50.5%)	105 (49.5%)	
Yes	332 (29.7%)	784 (70.3%)	
Diabetes			<.0001
No	240 (38.7%)	380 (61.3%)	
Yes	199 (28.1%)	509 (71.9%)	
Atrial fibrillation			0.1933
No	399 (33.6%)	789 (66.4%)	
Yes	39 (28.3%)	99 (71.7%)	
Smoke			0.5549
Never	232 (30.7%)	523 (69.3%)	
Current	133 (40.9%)	192 (59.1%)	
Former	73 (29.3%)	176 (70.7%)	
Stroke History			0.0073
No	325 (35.2%)	597 (64.8%)	
Yes	113 (27.9%)	292 (72.1%)	
Excessive alcohol			0.2089
No	403 (32.7%)	831 (67.3%)	
Yes	35 (38.0%)	57 (62.0%)	
Congestive heart failure			0.296
No	408 (33.4%)	813 (66.6%)	
Yes	30 (28.6%)	75 (71.4%)	
Coronary artery disease			0.0086
No	332 (35.1%)	613 (64.9%)	
Yes	106 (27.8%)	275 (72.2%)	
High cholesterol			0.1405
No	228 (35.0%)	424 (65.0%)	
Yes	210 (31.2%)	464 (68.8%)	

	Median (IQR) or Number (column %)		p-value
Stroke severity (NIHSS)	3 (1,7)	3 (1,6)	0.523
Body mass index	27.12 (23.78, 31.47)	29.95 (26.00, 34.33)	<.0001
Neck circumference	38 (35.50, 41)	41 (38.00, 44.00)	<.0001
Waist circumference	103 (94, 112.5)	111 (102, 121)	<.0001
Friedman palate position			0.0117
1	40 (40.4%)	59 (59.6%)	
2	69 (38.1%)	112 (61.9%)	
3 or 4	259 (30.3%)	596 (69.7%)	
Do you snore *			<.0001
No	106 (47.7%)	116 (52.3%)	
Yes	280 (29.0%)	686 (71.0%)	
Noticed quit breathing *			<.0001
No	323 (36.7%)	558 (63.3%)	
Yes	76 (21.9%)	271 (78.1%)	
Tired after sleep *			0.9626
Nearly Every Day	141 (35.3%)	259 (64.7%)	
3–4 Times Per Week	37 (29.6%)	88 (70.4%)	
1–2 Times Per Week	64 (31.1%)	142 (68.9%)	
1–2 Times Per Month	23 (23.7%)	74 (76.3%)	
Never/Nearly Never	167 (35.6%)	302 (64.4%)	

* In reference to the pre-stroke state.

Table 2.

SDB predictors: a multivariable logistic regression model.

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	−4.0176	0.9080	19.5779	<.0001
Age	0.0246	0.00841	8.5782	0.0034
Women:men	−0.3337	0.2339	2.0360	0.1536
Mexican American: non-Hispanic white	0.4869	0.1355	12.9097	0.0003
African American/Black:non-Hispanic white	−0.3369	0.2177	2.3947	0.1217
Hypertension	0.5968	0.2293	6.7712	0.0093
Diabetes	−0.2724	0.1793	2.3078	0.1287
Atrial fibrillation	0.1389	0.2908	0.2282	0.6328
Smoking (current:never)	−0.1146	0.1350	0.7198	0.3962
Smoking (former:never)	0.1314	0.1492	0.7757	0.3784
Stroke/TIA History	0.4403	0.1888	5.4387	0.0197
Excessive alcohol	−0.0676	0.3235	0.0437	0.8345
Congestive heart failure	0.2025	0.3583	0.3196	0.5719
Coronary artery disease	0.0447	0.1981	0.0510	0.8213
High cholesterol	−0.1799	0.1729	1.0823	0.2982
Log of stroke severity	−0.4298	0.1949	4.8637	0.0274
Body mass index	0.0317	0.0209	2.2973	0.1296
Waist circumference (cm)	0.0072	0.0086	0.7020	0.4021
Neck circumference (cm) Less than 30	−0.2709	0.3709	0.5335	0.4651
Neck circumference (cm) 30	0.0941	0.0334	7.9512	0.0048
Friedman palate position 2 vs 1	0.0174	0.1572	0.0122	0.9121
Friedman palate position 3–4 vs 1	0.2148	0.1267	2.8731	0.0901
History of snoring [†]	0.3558	0.2031	3.0684	0.0798
History of witnessed apneas [†]	0.3253	0.1960	2.7563	0.0969
Tired after sleep (3–4 times per week) ^{*†}	−0.1717	0.2248	0.5835	0.4449
Tired after sleep (1–2 times per week) ^{*†}	0.0238	0.1815	0.0171	0.8959
Tired after sleep (1–2 times per month) ^{*†}	0.3254	0.2646	1.5131	0.2187
Tired after sleep (never/nearly never) ^{*†}	−0.0470	0.1456	0.1041	0.7469

* Reference: Nearly every day

[†] From the Berlin questionnaire in reference to the pre-stroke state.