

regular expression matches). We identified 14 different formats and 7 study types. We used these manual annotations as multiclass outcomes in a random forest classifier to evaluate prediction of sleep study format and type using document property features. Out-of-bag (OOB) error rates and multiclass area under the receiver operating curve (mAUC) were estimated to evaluate training and testing performance of each model.

Results: We successfully predicted sleep study format and type using random forest classifiers. Training OOB error rate was 5.6% for study format and 8.1% for study type. When evaluating these models in independent testing data, the mAUC for classification of study format was 0.85 and for study type was 1.00. When applied to the large universe of diagnostic sleep study reports, we successfully extracted hundreds of discrete fields in 38,252 reports representing 33,696 unique patients.

Conclusion: We accurately classified a sample of sleep study reports according to their format and type, using a random forest multiclass classification method. This informed the development and successful deployment of custom data extraction tools for sleep study reports. The ability to leverage these data can improve understanding of sleep disorders in the clinical setting and facilitate implementation of large-scale research studies within the EHR.

Support (if any): American Heart Association (20CDA35310360).

400

UTILITY OF THE STOP QUESTIONNAIRE IN PREDICTING SLEEP DISORDERED BREATHING IN OLDER WOMEN VETERANS

Monica Kelly,¹ Isabel Moghtaderi,² Sarah Kate McGowan,¹ Gwendolyn Carlson,¹ Karen Josephson,¹ Michael Mitchell,¹ Dominika Swistun,¹ Constance Fung,¹ Michelle Zeidler,¹ Safwan Badr,³ Jennifer Martin¹

¹VA Greater Los Angeles Healthcare System, ²David Geffen School of Medicine at UCLA, ³Wayne State University School of Medicine

Introduction: Sleep disordered breathing (SDB) is underdiagnosed in older women, despite a significant increase in SDB prevalence post-menopause. Few studies have assessed the diagnostic accuracy of SDB screening questionnaires in older women, particularly older Women Veterans (WV). WV have higher rates of SDB compared to non-Veteran women and are particularly vulnerable to sleep disorders in general. We examined the diagnostic accuracy of the STOP questionnaire compared to home sleep apnea testing (HSAT) that includes sleep time estimation (i.e., WatchPAT) in older WV.

Methods: Cross-sectional baseline data obtained from chart review were combined from two behavioral sleep intervention studies targeting WV with sleep difficulties (i.e., insomnia symptoms) or SDB risk factors (e.g., hypertension, obesity). A total of 136 older WV (50-91y; age=60.0±7.8y) completed the STOP questionnaire (yes/no: snoring, tiredness, observed breathing pauses, and high blood pressure [BP]) and had an apnea-hypopnea index (AHI) available from their baseline HSAT (WatchPAT). Sensitivity, specificity, and positive and negative likelihood ratios (+LR/-LR) were calculated to characterize the diagnostic accuracy of STOP≥2 for AHI≥5 (mild SDB) or AHI≥15 (moderate SDB).

Results: 70.6% (n=96) of participants endorsed a STOP≥2, 83.8% (n=114) demonstrated an AHI≥5 and 46.3% (n=63) demonstrated an AHI≥15. For AHI≥5, sensitivity was 73.7% (95% CI=64.6,81.5%), specificity was 45.5% (95% CI=24.4,67.8%), +LR was 1.35 (95% CI=0.91, 2.01), and -LR was 0.58 (95% CI=0.33,1.00). For AHI≥15, sensitivity was 76.2% (95% CI=63.8,86%), specificity was 34.2% (95% CI=23.5,46.3%), +LR was 1.16 (95% CI=0.93,1.44), and -LR was 0.70 (95% CI=0.30,1.20).

Conclusion: The likelihood ratios for STOP≥2 limited the utility of the STOP vs. an HSAT system with sleep scoring in determining AHI. While the STOP correctly identified 3/4 of older WV with SDB on WatchPAT, it correctly identified <50% of older WV without SDB. Screening measures that better capture predictors of moderate SDB in women at risk for SDB are needed, especially in older women who may not present clinically with the common SDB symptoms (i.e. snoring, tiredness, observed breathing pauses, and high BP). STOP compared to polysomnography studies are also needed.

Support (if any): VA HSR&D IIR-13-058, IIR 16-244 and RCS 20-191; NIH/NHLBI K24 HL143055, VAGLAHS GRECC, VA Office of Academic Affiliations, and AASM Foundation.

401

CLINICAL PHENOTYPES OF OBSTRUCTIVE SLEEP APNEA IN WORLD TRADE CENTER RESPONDERS

Ankit Parekh,¹ Bresne Castillo,¹ Do Hyung Kim,¹ Kathleen Black,² Rafael de la Hoz,¹ David Rapoport,¹ Jag Sunderram,² Indu Ayappa¹

¹Icahn School of Medicine at Mount Sinai, ²Robert Wood Johnson University Hospital

Introduction: The heterogeneity of symptoms in obstructive sleep apnea (OSA) patients has been recently formalized into 3 distinct clusters: Sleepy, Disturbed Sleep, and Minimally Symptomatic. Our previous data showed that OSA is highly prevalent (>75%) in World Trade Center (WTC) responders, and positive airway pressure (PAP) treatment adherence is very poor (<20%). To better understand the heterogeneity of OSA in the WTC cohort, here we sought to examine the distribution of these distinct clinical phenotypes.

Methods: 643 subjects with no history of OSA or reported loud and frequent snoring before 9/11/2001 from the WTC health program clinical centers at Rutgers RWJMS, New Jersey, NYU School of Medicine, and Icahn School Medicine at Mount Sinai, New York underwent 2 nights of home sleep testing using the ARES unicorder (SleepMed, Inc., West Palm Beach, FL, USA). Epworth Sleepiness Scale (ESS), sleep onset insomnia, and sleep maintenance insomnia were assessed with questionnaires. OSA was defined as (AHI4%≥5 or RDI≥15/hr). The three clusters were defined as 1) Sleepy (ESS>10 and/or sleep onset/maintenance insomnia); 2) Disturbed Sleep (not sleepy (ESS≤10) and sleep onset/maintenance insomnia); and 3) Minimally Symptomatic (not sleepy (ESS≤10) and no sleep onset/maintenance insomnia). Distribution of clusters in the WTC cohort was compared to published data from the Sleep Apnea Global Interdisciplinary Consortium (SAGIC) and the Hispanic Community Health Study/Study of Latinos (HCHS/SOL).

Results: Among the subjects diagnosed with OSA (N 440; AHI4%=13(15); RDI =28(19); median(iqr); 81% men; age, 33-87 years; BMI, 27.4±3.7 kg/m²), the distribution of clinical phenotypes was 31.4% sleepy, 48.9% disturbed sleep, and 19.7% minimally symptomatic, and did not differ between OSA severity groups. In comparison to SAGIC and HCHS/SOL, the WTC cohort exhibited significantly increased prevalence of the disturbed sleep phenotype (WTC vs SAGIC: 48.9% vs. 19.8%, $\chi^2=54.9$; $p<0.001$; WTC vs. HCHS/SOL: 48.9% vs. 38.1%, $\chi^2=26.1$, $p<0.001$).

Conclusion: The predominant clinical phenotype of OSA in the WTC cohort is disturbed sleep (insomnia) and its prevalence is significantly greater than what has been observed in other large OSA cohorts. These findings may help explain the poor adherence to PAP treatment observed in the WTC cohort.

Support (if any): NIOSH U01OH01415; AASM Foundation 233-BS-20.