

ONLINE SUPPLEMENT

Respiratory effort during sleep and prevalent hypertension in obstructive sleep apnea

Jean-Benoit Martinot^{1,2}, Nhat-Nam Le-Dong³, Atul Malhotra⁴, Jean-Louis Pépin⁵

¹Sleep Laboratory, CHU Université Catholique de Louvain (UCL) Namur Site Sainte-Elisabeth,
Namur, Belgium

²Institute of Experimental and Clinical Research, UCL Bruxelles Woluwe, Brussels, Belgium

³Sunrise, Namur, Belgium

⁴Division of Pulmonary, Critical Care, and Sleep Medicine, Department of Medicine,
University of California San Diego, La Jolla, CA, USA

⁵HP2 Laboratory, Inserm U1300, University Grenoble Alpes, Grenoble, France

Supplemental methods

Gradient tree boosting is an approach where new tree-based classification rules are created to correct the errors made by existing rules until no further improvements can be made [1]. Then all rules are combined to make the final prediction and minimise the prediction error. Logarithmic loss was chosen as the evaluation metric (thus optimising the balanced accuracy between negative and positive labels). A hyperparameter tuning was performed with repeated cross-validation on training set. The best parameter configuration for training was as follow: learning rate of 0.01, maximum depth level of 5, subsample ratio of instances of 0.75. Lower learning rate allows prevention of overfitting; this will shrink the feature weights to make the boosting process more conservative. Due to the imbalanced proportion among the target labels, we balanced the training data using the synthetic minority oversampling technique (SMOTE) [2] before each training session.

Table S1. Evaluation of model performance by repeated 10-fold cross-validation and independent validation on test subset for a model that does not include the proportion of total sleep time with increased respiratory effort during sleep based on mandibular jaw movements measurement (REMOV).

Metrics	Repeated 10-fold cross-validation (n=901)			Independent validation on test subset (n=226)		
	Estimated	95% CI	p-value*	Estimated	95% CI	Significance**
F1 score	0.76	0.69–0.81	0.713	0.62	0.54–0.70	Yes
Balanced accuracy	0.75	0.69–0.80	0.009	0.73	0.68–0.79	Yes
Sensitivity	0.80	0.71–0.89	0.996	0.73	0.65–0.82	Yes
Specificity	0.70	0.62–0.76	<0.001	0.73	0.68–0.79	Yes
PPV	0.73	0.64–0.81	<0.001	0.54	0.51–0.59	Yes
NPV	0.87	0.67–0.87	<0.001	0.86	0.80–0.91	Yes
LR+	2.70	2.04–3.45	<0.001	2.81	2.18–3.61	Yes
LR-	0.29	0.16–0.43	0.931	0.36	0.24–0.49	Yes
ROC AUC	0.83	0.78–0.87	0.002	0.84	0.74–0.86	Yes

CI, confidence interval; LR, likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; ROC AUC, area under the receiver operator characteristic curve.

*p-value versus full model; paired samples t-test (Table 2).

**Statistically significant versus full model; permutation test (Table 2).

Figure S1. Global performance of a model that does not include the proportion of total sleep time with increased respiratory effort during sleep based on mandibular jaw movements measurement (REMOV), based on 10-fold cross-validation and independent data.

a) Average normalised confusion matrix obtained from 100 replications of extreme gradient boosting model in a 10-fold cross-validation; b) Normalised confusion matrix based on application of the final model on unseen data from 226 subjects; c) Receiver operating characteristic curve evaluating the global performance of the final model on an independent testing subset. AUC, area under the curve; Negative, hypertension absent; Positive, hypertension present.

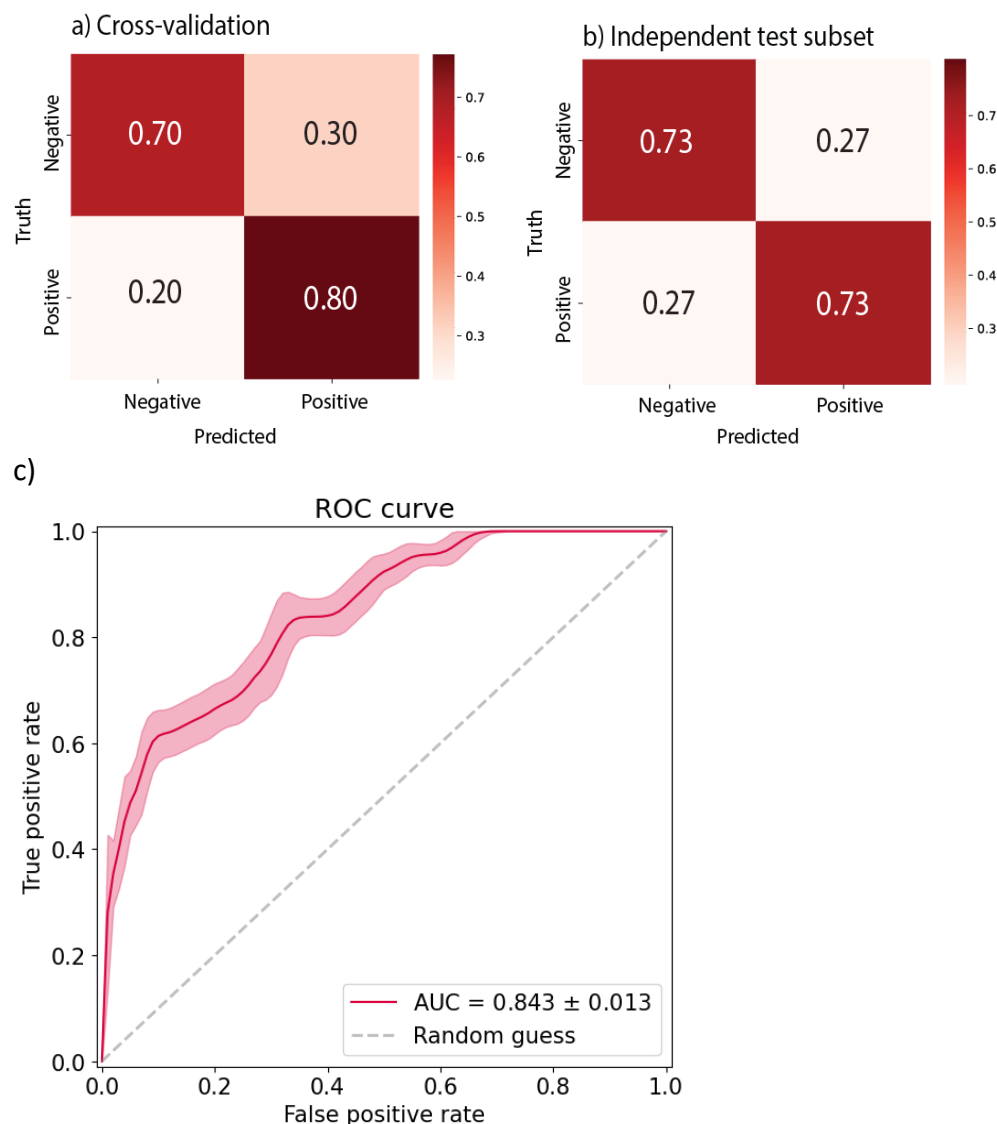


Figure S2. Ridge logistic regression analysis.

Bars show the average value of coefficients from a 10-fold cross-validation of the Ridge logistic regression model (this was a linear regression with binomial distribution and a L2 regularisation, providing more stable parameters). Higher model coefficient values indicate greater strength of association between the corresponding predictor and the risk of comorbid hypertension.

AHI, apnea-hypopnea index; ArI, arousal index; BMI, body mass index; circ., circumference; Desat_dt, time with oxygen saturation <90% or <95%; ESS, Epworth Sleepiness Scale; OAHl, obstructive apnea-hypopnea index; ODI, oxygen desaturation index; ORDI, obstructive respiratory disturbance index; PSG, derived from polysomnography; RDI, respiratory disturbance index; REMOV, increased respiratory effort during sleep (based on mandibular jaw movements measurement); Sr, derived from automatic analysis of mandibular jaw movements by the Sunrise system; TST, total sleep time.

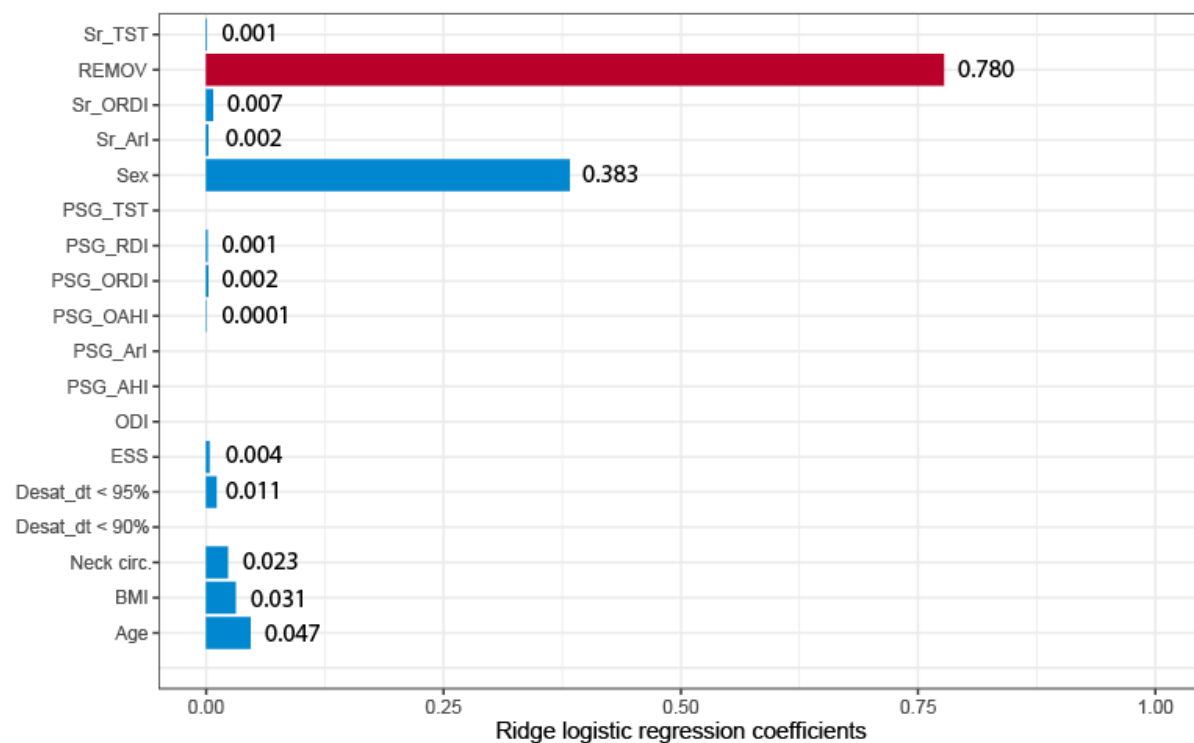
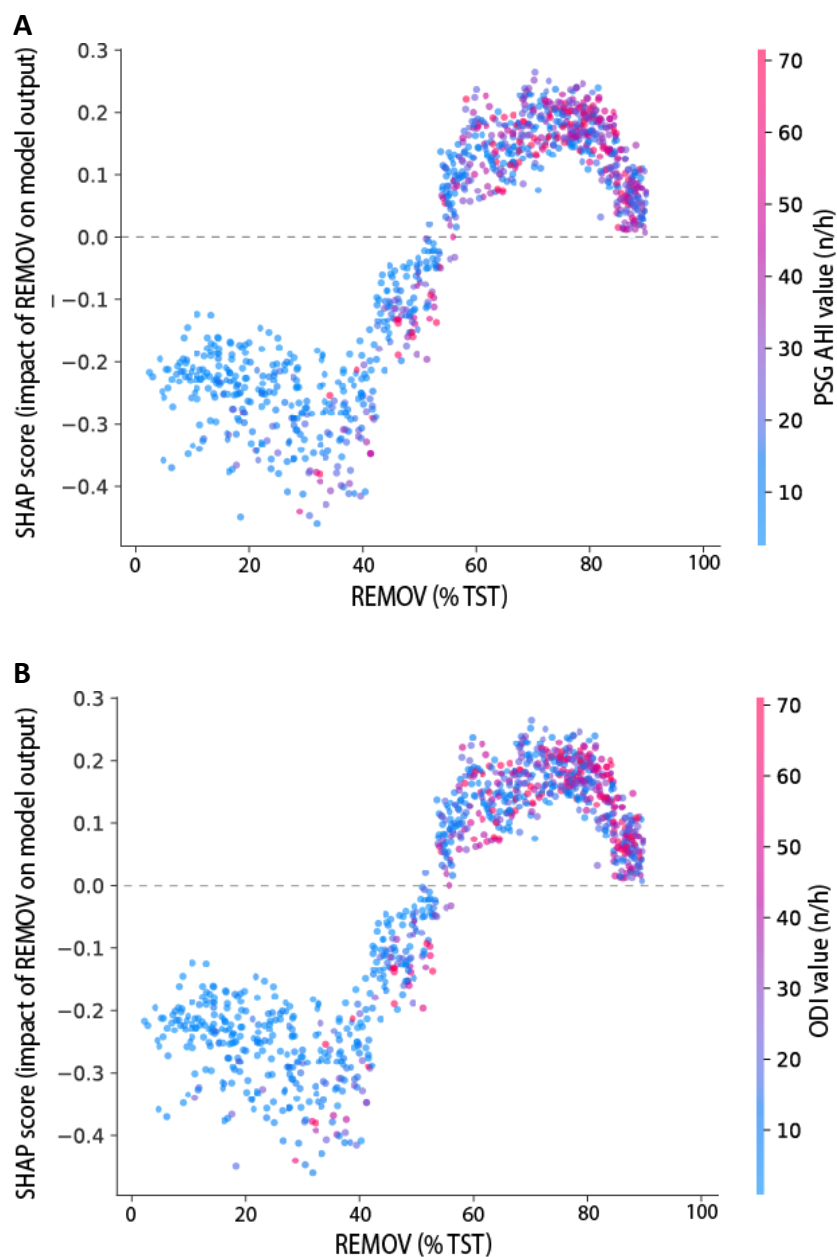


Figure S3. Shapley additive explanation method (SHAP) score for the contribution of increased respiratory effort duration in the absence of oxygen desaturation and when sleep fragmentation was limited. The red dots represent subjects with primary snoring and low values of the apnea-hypopnea index (**A**) and oxygen desaturation index (**B**) (<5 events/h). AHI, apnea-hypopnea index; ODI, oxygen desaturation index; REMOV, proportion of total sleep time with increased respiratory effort during sleep (based on mandibular jaw movements measurement).



References

1. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. August 2016. Pages 785–794.
2. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Art Intel Res.* 2002;16:321-357.