Morphological Prediction of Continuous Positive Airway Pressure–associated Acute Respiratory Instability

Thijs-Enagnon Nassi[1,3], Eline Oppersma[3], Gonzalo Labarca[1,4], Dirk W. Donker[3,5], M. Brandon Westover[1], and Robert J. Thomas[2]

[1]Department of Neurology and [2]Division of Pulmonary and Sleep Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts; [3]Cardiovascular and Respiratory Physiology Group, TechMed Centre, University of Twente, Enschede, the Netherlands; [4]Department of Respiratory Diseases, School of Medicine, Pontifical Catholic University of Chile, Santiago, Chile; and [5]Department of Intensive Care, University Medical Center Utrecht, Utrecht, the Netherlands

ORCID IDs: 0000-0003-0269-3405 (T.-E.N.); 0000-0002-0150-306X (E.O.); 0000-0002-0069-3420 (G.L.); 0000-0001-6496-2768 (D.W.D.); 0000-0003-4803-312X (M.B.W.); 0000-0002-5575-3953 (R.J.T.).

Abstract

Rationale: Multiple mechanisms are involved in the pathogenesis of obstructive sleep apnea (OSA). Increased loop gain (LG) is a key target for precision OSA care and may be associated with treatment intolerance when the upper airway is the sole therapeutic target. Morphological or computational estimation of LG is not yet widely available or fully validated, and there is a need for improved phenotyping and/or endotyping of apnea to advance its therapy and prognosis.

Objectives: This study proposes a new algorithm to assess self-similarity (SS) as a signature of increased LG using respiratory effort signals and presents its use to predict the probability of acute failure (i.e., high residual event counts) of continuous positive airway pressure therapy.

Methods: Effort signals from 2,145 split-night polysomnography studies from the Massachusetts General Hospital were analyzed for SS and used to predict acute continuous positive airway pressure therapy effectiveness. Logistic regression models were trained and evaluated using fivefold cross-validation.

Results: Receiver operating characteristic and precision-recall curves with area under the curve values of 0.82 and 0.84, respectively, were obtained. SS combined with the central apnea index (CAI) and hypoxic burden outperformed CAI alone. Even in those with a low CAI by conventional scoring criteria or only mild desaturation, SS was related to poor therapy outcomes.

Conclusions: The proposed algorithm for assessing SS as a measure of expressed high LG is accurate and noninvasive and has the potential to improve phenotyping and/or endotyping of apnea, leading to more precise OSA treatment strategies.

Keywords: sleep apnea; loop gain; similarity

(Received in original form November 17, 2023; accepted in final form September 17, 2024)

Supported by U.S. National Institutes of Health grants RF1NS120947, R01AG073410, R01HL161253, R01NS126282, R01AG073598, R01NS131347, and R01NS130119 and National Science Foundation grant 2014431. M.B.M, is a co-founder, scientific advisor, and consultant to Beacon Biosignals and has a personal equity interest in the company.

Author Contributions: Study concept and design: T.-E.N., E.O., M.B.W., and R.J.T. Data acquisition: T.-E.N., G.L., and M.B.W. Data analysis: T.-E.N. and E.O. Data interpretation: T.-E.N., E.O., G.L., D.W.D., M.B.W., and R.J.T. Manuscript drafting and revising: T.-E.N., E.O., G.L., D.W.D., M.B.W., and R.J.T.

Correspondence and requests for reprints should be addressed to M. Brandon Westover, M.D., Ph.D., Department of Neurology, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215. E-mail: mwestove@bidmc.harvard.edu.

This article has a related editorial.

This article has a data supplement, which is accessible at the Supplements tab.

Ann Am Thorac Soc Vol 22, No 1, pp 138–149, Jan 2025

DOI: 10.1513/AnnalsATS.202311-979OC
Internet address: www.atsjournals.org

Obstructive sleep apnea (OSA) is a condition that can increase the risk of cardiometabolic and neurodegenerative diseases (1, 2). Several driver endotypes are now recognized to be important in the pathophysiology of OSA, resulting in an admixture of anatomical and nonanatomical mechanisms (3, 4). Of the latter, high loop gain (HLG) is considered a key treatment target for precision sleep apnea care. However, recognition of LG driving sleep apnea outside the classic central apnea of Hunter-Cheyne-Stokes respiration has proven difficult. Several methods have been proposed, including hypoxic-hypercapnic gas administration (5–9), proportional assist ventilation (10, 11), breath-holding estimates (12–14), mathematical modeling of underlying breathing mechanics (15–17),

cardiopulmonary coupling (18, 19), and mathematical estimates of self-similarity (SS) (20).

Although the non–rapid eye movement (REM) sleep CO_2 reserve can improve with sleep apnea treatment (21, 22), therapies targeting upper airway obstruction alone may result in adverse physiological and clinical outcomes when HLG is present (23–27). One third to half of patients receiving continuous positive airway pressure (CPAP) continue to experience residual events (28–30), and one mechnism is HLG. LG can be lowered by supplemental oxygen (31) or controlling the inspiratory CO_2 level (32) and by medications such as acetazolamide, zonisamide, or sulthiame (33–36). To help guide precision sleep apnea treatment strategies, there is a need for practically applicable and clinically predictive phenotyping and endotyping of apnea, which can allow risk stratification.

We previously described a method to assess SS using respiratory effort signals computed over multiple consecutive respiratory events; this analysis predicted acute failure (i.e., high residual events) during CPAP titration with modest accuracy (20). Our group has described the adverse impact on long-term CPAP use and residual respiratory events of this CPAP-related acute respiratory instability even when it is below the threshold of criteria for treatment-emergent central sleep apnea (CSA) (37). However, a key limitation of this previously published method is algorithmic rigidity: small asymmetries and/or variances in consecutive signal fluctuations, which are common in clinical practice and may reflect anatomical contributions to pathophysiology and sleep state and/or arousal variability, result in missed detection; moreover, brief clusters of events may remain undetected. The modified algorithm presented here focuses on the morphology of individual events and aims to offer a more accurate, clinically suitable increased LG surrogate assessment method. This method aids in risk-stratifying patients who are otherwise characterized as "obstructive" for possible adjunctive pharmacotherapy (38).

Methods

Dataset Description and Preparation

The dataset was selected from an archive containing clinical polysomnography (PSG) recorded at the Massachusetts General Hospital (MGH) Sleep Laboratory between 2008 and 2022. The MGH Institutional Review Board approved the retrospective analysis of clinically acquired PSG data. All sleep studies included abdominal respiratory inductance plethysmography (RIP) signals with a sampling frequency of 200 Hz or 512 Hz. A notch filter of 60 Hz (power line) and a low-pass filter of 10 Hz were applied to reduce nonphysiological noise. Subsequently, recordings were resampled to 10 Hz and normalized using the mean and standard deviation of the 5th to 95th percentile clipped signal.

The MGH Sleep Center is an American Academy of Sleep Medicine (AASM) accredited sleep center that scores obstructive, central, and mixed apneas; hypopneas (using the 4% desaturation rule); and respiratory effort–related arousals (39). To conform event scoring with the current AASM convention, hypopneas meeting the 3%/arousal criteria were added using an automated scoring algorithm while using the original manually scored arousals and sleep staging (40).

Estimation of Titration Success or Failure

To create clear success and failure categories, recordings with an apnea–hypopnea index (AHI) >10 before CPAP were categorized based on their AHI and central apnea index (CAI) during the titration phase of CPAP into a therapy success group (AHI <10 and CAI <5) and a therapy failure group (AHI ≥ 30 or CAI ≥ 10 despite CPAP). Recordings outside these two groups were excluded. For additional details on data preparation, *see* data supplement. To visualize the relationship between the number of respiratory events and CPAP effectiveness, histograms were created showing the fraction of patients in whom therapy failed across the range of CAI, AHI, and SS scores.

The SS score quantifies the similarity of consecutive breathing cycles in zones with detectable discrete respiratory events. We outline breathing cycles by marking the peaks of inhalation and exhalation (Figure 1). Patterns that repeat consistently are identified by comparing the shapes of consecutive breathing cycles (Figure 2). Each pattern is given a score from 0 to 1, with higher scores indicating more self-similar (i.e., clone-like) breathing patterns (Figure 3). This scoring helps differentiate between stable and unstable breathing, aiding in the indirect assessment of LG and the prediction of CPAP therapy outcomes. *See* data supplement for an in-depth technical description.

Predicting CPAP Effectiveness

SS scores during the diagnostic phase were summarized across 5-minute segments. Histograms were generated from these scores, and a logistic regression (LR) model was trained and evaluated using fivefold cross-validation. For each test patient, two mean absolute histogram distances were computed, first with reference to the average histogram of all training patients with a successful therapy outcome, and second with reference to the failure group (*see* data supplement for details). The difference between the histogram distances, summarized into a single error score, was used as input for the LR model. In addition, LR models were trained using the CAI and the hypoxic burden (41) as individual inputs and in combination with SS.

Models were evaluated for accuracy and calibration on the entire cohort and on a subpopulation of patients with a CAI <1. Receiver operating characteristic (ROC), precision-recall (PR) (including 95% confidence intervals using bootstrapping; $N = 10{,}000$), and calibration curves were computed. Heat maps were used to study the relationship between the level of SS and the hypoxic burden.

For additional validation, a second external dataset of home sleep apnea tests was used to compare SS with LG (*see* data supplement). For each recording, SS scores and LG estimates by phenotyping using PSG (PUP) (17) were obtained, and their relationship was studied by fitting a polynomial and computing a correlation coefficient (R^2).

Results

Patient Characteristics and Algorithm

Data from 2,145 unique split-night recordings were included and are summarized in Table 1. Examples comparing the new algorithm with the previously published method are shown in Figure 4. In Figure 5, histograms and example RIP segments from four patients with varying levels of SS are shown. These show that, as the level of SS increases, the number of automatically tagged HLG oscillations increases, and the distribution becomes increasingly skewed rightward. More than

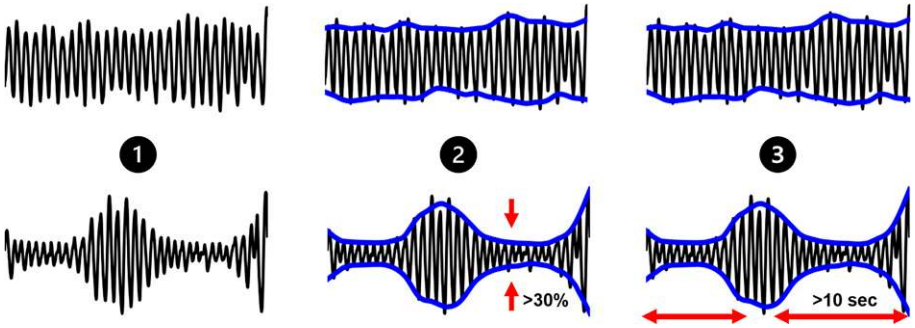

Figure 1. Two example effort traces, the upper one with regular breathing and the lower one with self-similar signal oscillations (1), in which inspiratory and expiratory peak envelopes were used to find amplitude reductions of ⩾30% (2) with a minimum duration of 10 seconds (3).

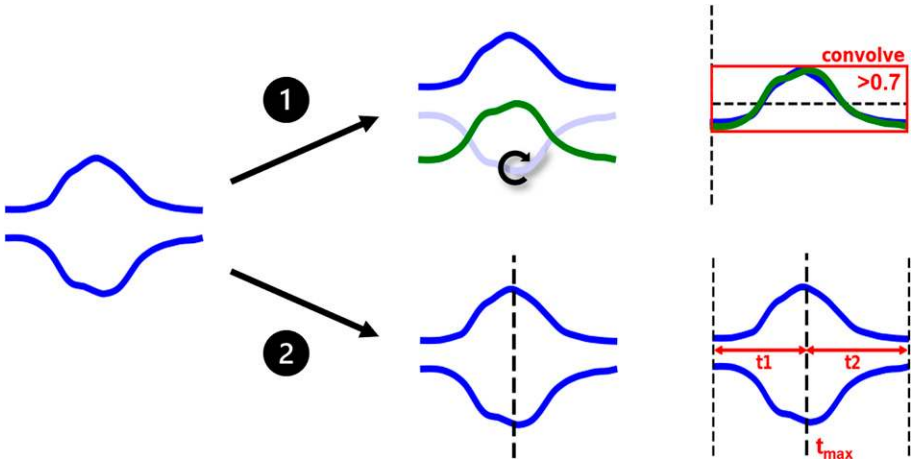

Figure 2. Example of how envelopes of signal oscillations were assessed by (1) computing a convolution between the positive and the rotated negative envelope and (2) assessing the location of the maximum distance between the positive and negative envelopes. t_{max} = maximum distance between positive and negative envelopes.

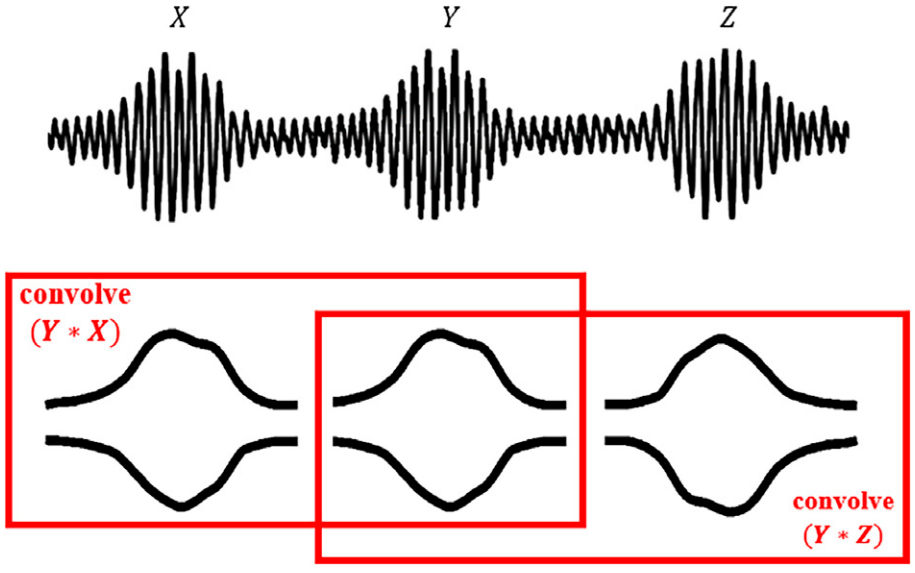

Figure 3. Example of how self-similarity in respiratory inductance plethysmography–derived effort is computed using the convolution scores across neighboring envelopes of oscillations.

Table 1. Cohort demographic characteristics ($N = 2,145$)

Characteristic	Incidence
Sex	
Female	642 (30.0%)
Male	1,503 (70.0%)
Age	
<20 yr	12 (0.6%)
20–40 yr	271 (12.6%)
40–60 yr	942 (43.9%)
60–80 yr	834 (38.9%)
>80 yr	86 (4.0%)
Race	
White	1,607 (74.9%)
Black	166 (7.7%)
Asian	51 (2.4%)
American Indian or Alaska Native	3 (0.1%)
Pacific Islander	1 (0.1%)
Other	201 (9.4%)
Unavailable	116 (5.4%)
CAI	
Diagnostic phase	5.6 ± 10.5
Titration phase	5.4 ± 9.4
AHI (3%/arousal)	
<15	47 (2.3%)
15–30	172 (8.1%)
>30	1,897 (89.7%)
AHI (4%)	
<15	119 (5.6%)
15–30	359 (16.7%)
>30	1,672 (77.8%)
Medications ($n = 1,539$)	
Hypertension	31 (2.0%)
Cardiac	21 (1.4%)
Sleep	91 (5.9%)
Antidepressants	28 (1.8%)
Antiseizure	8 (0.5%)
Opiates	29 (1.9%)
Benzodiazepines	46 (3.0%)
Diabetic	181 (11.8%)
Z-drugs	20 (1.3%)
Stimulants	71 (4.6%)
Neuroleptics	25 (1.6%)
RLS/PLMS	4 (0.3%)
Immune	53 (3.4%)
Neurodegenerative	3 (0.2%)
Other	1,079 (70.1%)

Definition of abbreviations: AHI = apnea–hypopnea index; CAI = central apnea index; PLMS = Periodic Limb Movement Syndrome; RLS = Restless Legs Syndrome.
Data presented as mean ± standard deviation where applicable.

95% of all the SS oscillations were found in non-REM sleep (Figure 6). After normalizing over the respective total sleep stage durations, the proportion of SS oscillations during REM sleep was 23.3%.

Self-Similarity and CPAP

A total of 1,089 of the 2,145 studies were categorized into CPAP success/failure groups (1,326 when using the 4% hypopnea criteria). Among these groups, the proportion of patients who continued to experience apnea and hypopnea during CPAP was positively correlated with higher baseline values of CAI, AHI, and SS (Figure 7). High SS was more prevalent in men; most patients in whom CPAP acutely failed were men, with the exception of individuals with an AHI >125, whereas the CPAP success groups show a more even distribution. When using the 4% hypopnea criteria to compute CPAP therapy effectiveness, a similar trend was observed, as shown in Figure E1 in the data supplement. However, the total number of patients considered to have failed CPAP by this criterion was lower, at 34.0% versus 49.6%. Figures E2 and E3 show two overnight recordings from patients exhibiting mild and severe SS, respectively. Each recording was cut into 10 consecutive RIP segments from left to right and top to bottom.

The predictability of CPAP failure using SS, CAI, and hypoxic burden as features to predict CPAP failure is shown in the ROC and PR curves in Figure 8. The area under the curve (AUC) values obtained by using SS, CAI, and hypoxic burden as individual features were 0.75 (95% confidence interval, 0.73–0.78), 0.73 (0.70–0.75), and 0.65 (0.63–0.68) and 0.76 (0.73–0.80), 0.77 (0.74–0.79), 0.69 (0.65–0.72) for the ROC and PR curves, respectively. The new SS algorithm showed significantly higher AUC values compared with the original algorithm. AUC values further increased when combining predictive features, up to an ROC AUC of 0.82 (0.80–0.84) and a PR AUC of 0.83 (0.81–0.86) when combining SS with the CAI and hypoxic burden as three input features for an LR model. All LR models showed excellent risk calibration (i.e., calibration curves are close to the diagonal). It is notable that, for the individual features, no calibration points below a probability of 0.3 were observed. When combining features, the range of prediction increased and calibration points between 0.2 and 0.9 were observed. Similar results using the 4% hypopnea criteria are shown in Figure E4.

Using the models trained on the entire cohort, we examined patients with a CAI <1 ($n = 455$). When using the CAI alone as a predictive feature for this patient subgroup, a clear decrease in performance was observed. Figure 9 shows that SS did provide ROC and PR AUC values of 0.73 and 0.57, respectively, outperforming the CAI and hypoxic burden. We observed that patients with treatment-emergent central apnea and high residual obstructive events, while being labeled with a CAI of 0 before titration, are better identified using SS.

In Figure 10, a clear cluster of patients with low SS and a low to moderate hypoxic burden is apparent within the CPAP success group, whereas, for patients from the CPAP failure group, clusters with low and high SS spanning the entire range of the hypoxic burden were observed. When using the 4% hypopnea criteria, less obvious clusters for the failure group were observed (Figure E5).

Self-Similarity versus PUP

In the external dataset ($n = 202$), we compared the amount of sleep characterized by a high SS score (similarity threshold >0.8) with PUP-computed LG, revealing a positive

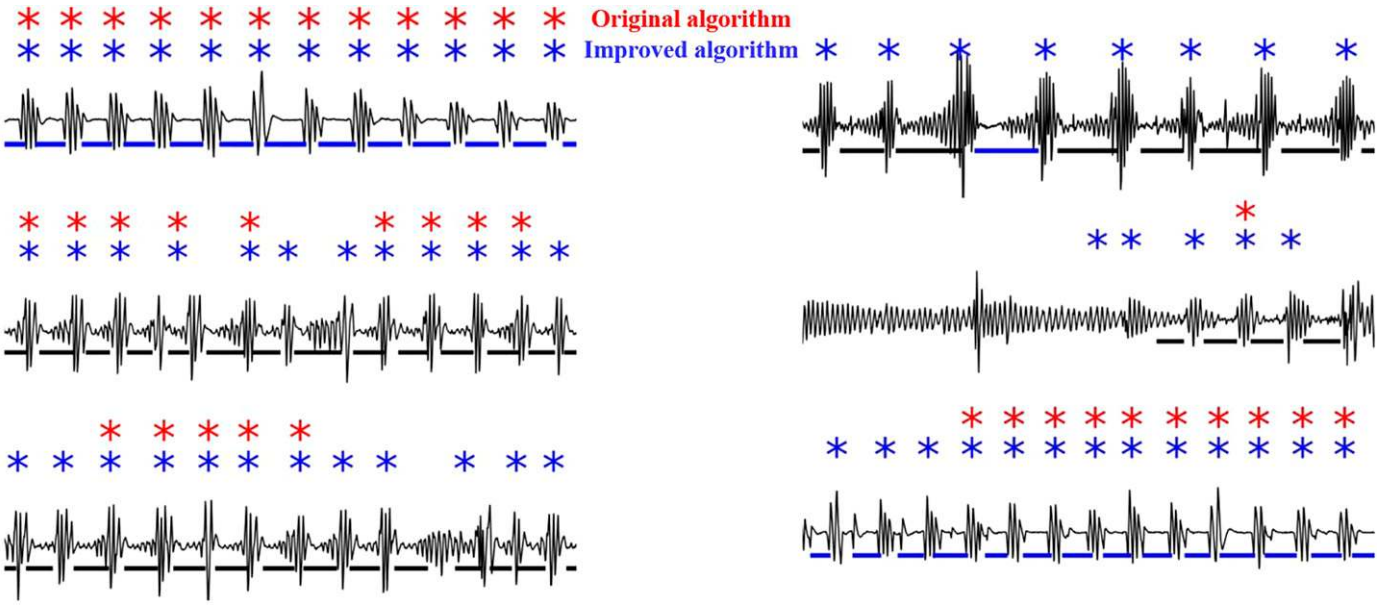

Figure 4. Example effort traces with technician-scored obstructive events (black lines) and central events (blue lines). The red and blue asterisks indicate signal oscillations marked as high loop gain by the old and improved self-similarity algorithms, respectively.

relationship. For lower SS values, the relationship follows a somewhat linear trajectory, which later grows exponentially. This characterization was captured by fitting a third-order polynomial, resulting in an R^2 value of 0.53, *see* Figure 11.

Discussion

This work demonstrates that the risk of acute CPAP failure is predictable from the oscillation morphology of RIP-derived effort signals. The described algorithm aims to detect expressed HLG using the horizontal and vertical morphological symmetry patterns shared in common across periodic breathing, CSA, and non-REM–dominant HLG-driven apnea (42). High SS scores indicate a pattern of breathing instability similar to that seen classic HLG patterns such as Hunter-Cheyne-Stokes respiration, supported by its evident relationship with LG estimates in 202 external sleep studies. High SS scores may be helpful for clinicians in assessing the risk of treatment-emergent CSA, particularly in patients who do not exhibit central apneas on a diagnostic study. Although our original algorithm showed reduced accuracy in scenarios with subtle asymmetry across consecutive oscillations, the morphology assessment in the present work improves sensitivity such that HLG-driven events are now better detected.

SS versus PUP

An immediate question that requires resolution is whether SS is a true reflection of LG. Although each method used to estimate LG gives slightly different results, the various measures covary positively such that a high value with one method generally aligns with a similarly high value of another. SS has a positive association with PUP-estimated LG. Up to approximately 20% SS, the relationship is relatively shallow and then increases. The highest SS values are when LG is 0.9–1. Factors that could impact the relationship of SS with PUP LG include sleep quality, coexisting upper-airway anatomically driven collapsibility, arousal threshold, and body position.

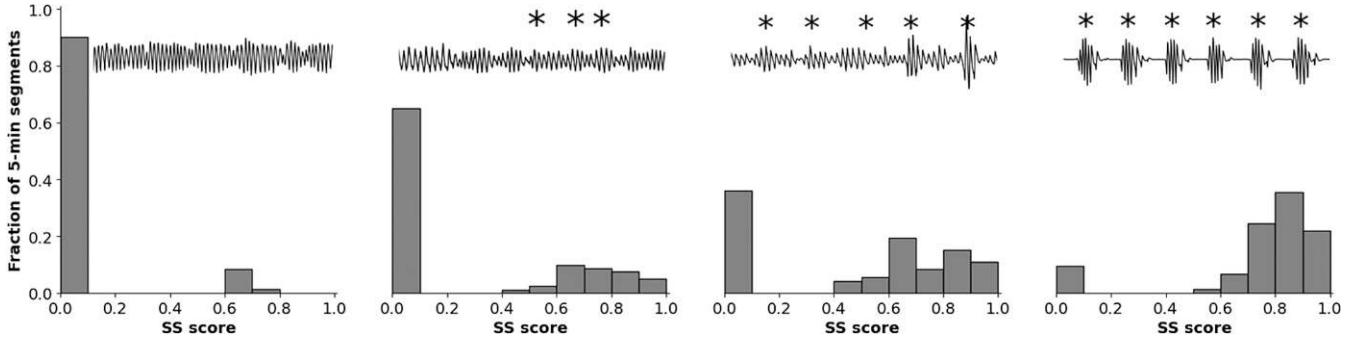

Figure 5. Respiratory effort example traces from four patients ranging from low to high self-similarity from left to right. The associated histograms show how self-similarity across all 5-minute segments was summarized for these patients. The automatically tagged high loop gain oscillations are marked with asterisks. SS = self-similarity.

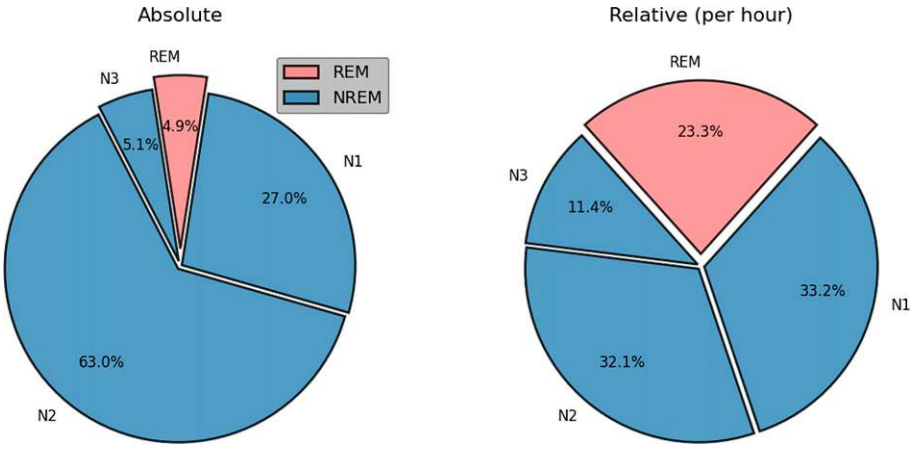

Figure 6. Pie chart of the proportion of automatically tagged high loop gain oscillations per sleep stage in absolute numbers and normalized by the hours of their respective sleep stage on the left and right, respectively. NREM = non–rapid eye movement; REM = rapid eye movement.

Central Apnea as Predictor of CPAP Effectiveness

The CAI does not accurately predict residual respiratory events during CPAP for all patients. Figure 7 shows that, for an increasing CAI, patients have a decreasing probability of sufficient benefit from CPAP, yet a significant subgroup of patients with few central apneas continue to experience respiratory events during titration. The results indicate that a patient's risk of persistent apneas and hypopneas cannot be fully captured by scoring central apneas alone, and accurate scoring of central hypopneas has remained elusive. Drive-dependent respiratory events dominate in obstructive sleep apnea (43), so it should be no surprise that central hypopneas are also associated with upper airway narrowing (44). Thus, manual/visual scoring of individual respiratory events cannot accurately determine the pathophysiological driver of hypopneas. The rigidity of the AASM scoring rules for central hypopneas exclude any obstructive feature, but it is well documented that respiratory events with coexisting central and obstructive

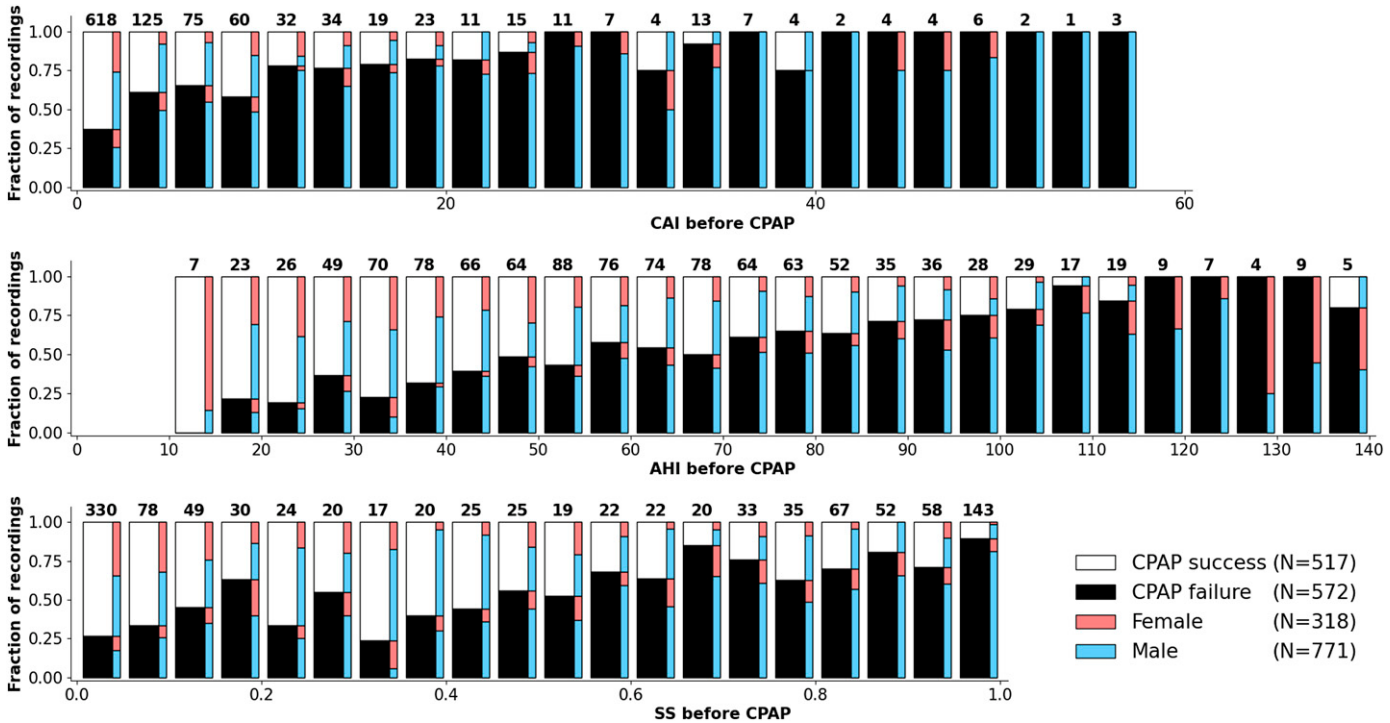

Figure 7. Histograms showing the fraction of patients who fail CPAP therapy across the range of the AHI, the CAI, and SS, based on the 3%/arousal hypopnea rule. Sex proportions are displayed in red and blue. The number above each bar indicates the sum of patients contributing to each bin. AHI = apnea–hypopnea index; CAI = central apnea index; CPAP = continuous positive airway pressure; SS = self-similarity.

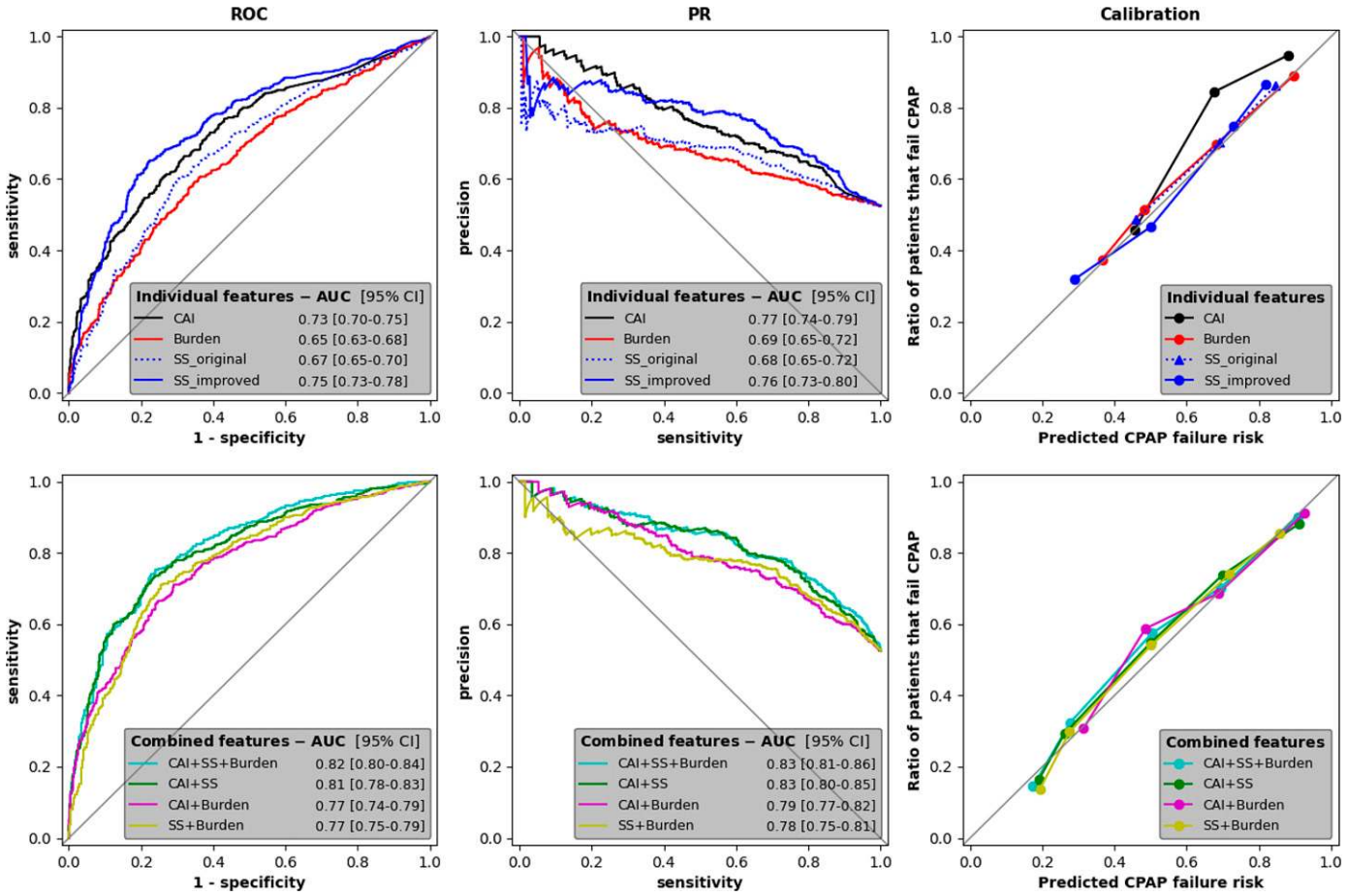

Figure 8. Receiver operating characteristic (ROC), precision-recall, and calibration curves of the CAI, hypoxic burden, and SS as individual predictive features for continuous positive airway pressure (CPAP) failure. The performance of each feature combination is shown below. CPAP treatment effectiveness is computed using the 3%/arousal hypopnea rule. AUC values and 95% confidence intervals for the ROC and precision-recall curves are shown in the respective legends. AUC = area under the curve; CAI = central apnea index; PR = precision recall; SS = self-similarity.

features are common, even at high altitude (45, 46).

Morphological CPAP Success/Failure Prediction

Improved prediction of CPAP effectiveness is obtained in our present approach by identification of expressed HLG in respiratory signals. The level of morphological SS across RIP-derived effort oscillations as an individual predictor achieves similar accuracy compared with CAI when this index is high, with ROC and PR AUC values of 0.75 and 0.76, respectively, including accurate calibration (Figure 8). However, the SS analysis is most useful when central apneas do not dominate in a given patient but hypopneas, events with some flow limitation, and physiologically mixed events are frequent. The struggle to differentiate central from obstructive hypopneas may be obviated by our method. Events detected by SS analysis could be called "HLG events" or "HLG OSA" to reflect driver pathophysiology, moving away from the scoring tension between obstructive and central hypopneas and expanding on traditional OSA versus CSA. This approach would be similar to what would be needed to clinically apply any more direct LG measurement. Although a high hypoxic burden may indicate that a patient's sleep apnea is more severe, SS also outperformed hypoxic burden for the prediction of residual events, as the latter showed AUC values ranging between 0.62 and 0.72. When combining SS with CAI and hypoxic burden, higher predictive performance was obtained, with AUC values of approximately 0.82–0.83, suggesting that the three metrics are complementary. Although high SS scores indicate HLG and potential challenges with CPAP therapy, they do not necessarily preclude its use. CPAP can still be effective, especially when combined with adjunctive treatments such as supplemental oxygen or medications like acetazolamide. A comprehensive, tailored approach ensures that patients with high SS scores receive the most appropriate and effective treatment. The SS concept has already impacted our own clinical practice. Recognizing SS during positive pressure titration allows the avoidance of excessive positive airway pressure and may drive decisions about the use of low-dose acetazolamide or zonisamide as adjunctive therapy for PAP, oral appliance, and hypoglossal nerve stimulation therapies.

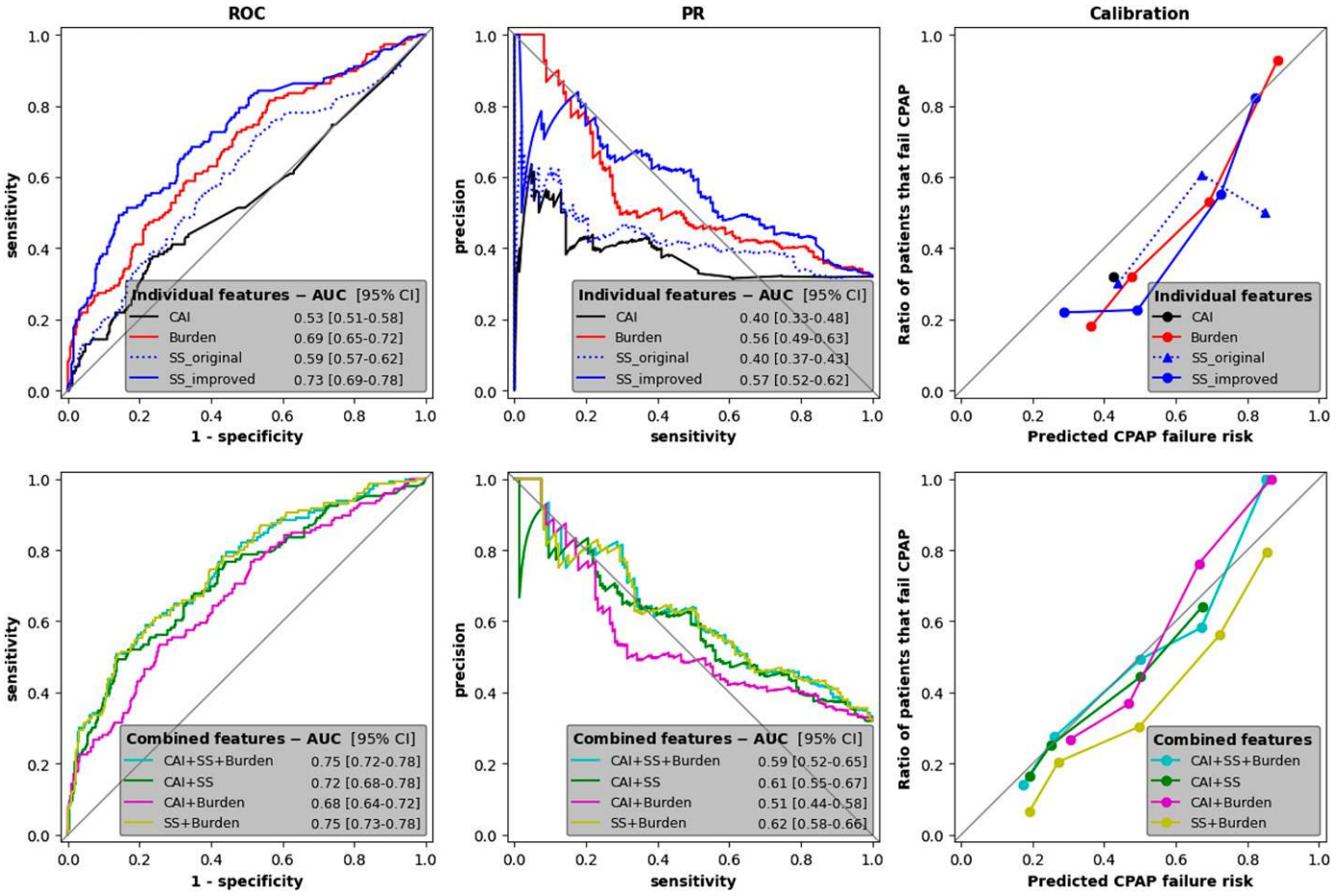

Figure 9. Receiver operating characteristic (ROC), precision-recall, and calibration curves of the central apnea index (CAI), hypoxic burden, and SS for all patients with a pretreatment CAI <1 ($n = 455$). The upper three graphs show the individual negative predictive features for continuous positive airway pressure (CPAP) therapy, with the performance of each feature combination shown below. CPAP treatment effectiveness is computed using the 3%/arousal hypopnea rule. AUC values and 95% confidence intervals for the ROC and precision-recall curves are shown in the respective legends. AUC = area under the curve; PR = precision recall; SS = self-similarity.

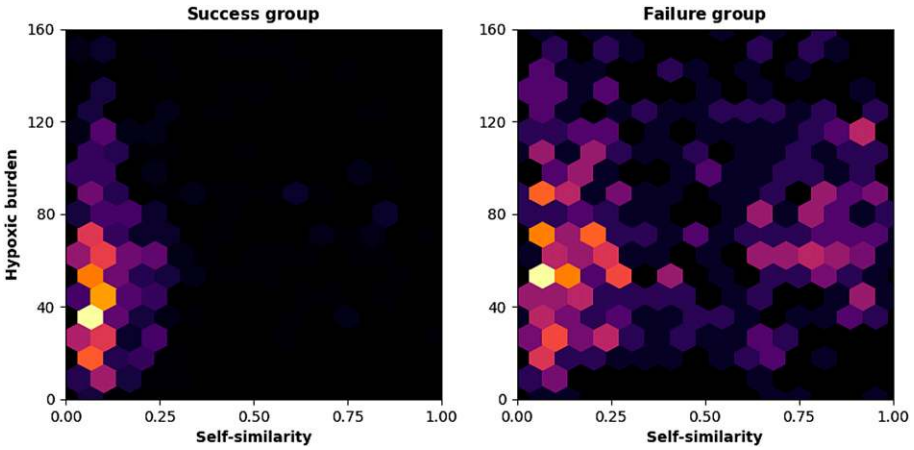

Figure 10. Heat maps showing the relationship between the level of self-similarity and the hypoxic burden within the continuous positive airway pressure success and failure groups (left and right, respectively) using the 3%/arousal hypopnea rule.

Hypoxic Burden and CPAP Failure

Intermittent hypoxia increases LG through carotid body sensitization (47), and CPAP reduces LG in patients with hypoxic sleep apnea (9). A low SS and low to moderate hypoxic burden may predict that CPAP will be effective, yet many patients with this phenotype still experienced unsuccessful treatment, as shown in Figure 10. Moreover, patients with a low or high hypoxic burden were likely to continue to experience respiratory events during CPAP if they displayed high SS in their RIP signals. These findings suggest that patients with HLG, regardless of oxygen desaturation, have a large positive likelihood of CPAP being acutely ineffective. In fact, desaturation is often mild in idiopathic CSA and heart failure–associated periodic breathing as a result of the vigorous arousals and hyperventilation inherent in these disorders (48, 49). Thus, events driven by HLG are related to poor therapy outcomes, even though such events do not necessarily cause significant desaturation. Future work will focus on exploring additional burdens, such as ventilatory and arousal factors, and their relationship to CPAP effectiveness.

Impact of Hypopnea Scoring Criteria

The predictive power of all features decreased when using the AASM 4% hypopnea criteria instead of 3%/arousal. Calibration points with greater distance to the diagonal were observed, and the heat maps suggest less clear clustering in SS among the CPAP failure group. These findings support the AASM's recent update of the hypopnea scoring criterion to 3%/arousal. These results indicate that the relationship between residual respiratory events and CPAP therapy effectiveness is better explained when including hypopneas meeting the 3%/arousal criterion. Typically, clinical scoring does not include the differentiation between obstructive and central hypopneas. Similarly, flow limitation not associated with sufficient desaturation or arousal are usually ignored when adhering to the AASM scoring criteria. Such mild events could be additional predictors for CPAP effectiveness but are not included in this work.

Comparison with the Published Literature

LG estimates are not yet part of standard clinical practice as a result of the lack of easily applied methodology. Various publications evaluate the ventilatory response and sensitivity to hypoxia and/or hypercapnia by administering a gas mixture containing adjusted O_2/CO_2 levels (5–9). These methods require monitoring of the gas concentrations in inhaled and exhaled air, whereas typically only flow, pressure, and effort signals are obtained sleep during conventional sleep studies. Controlled mechanical ventilation can also be used to assess LG. The response of the patient's respiratory system to changes in tidal volume and respiratory rate provided by the ventilator can be measured and used to calculate LG (10, 11, 50). Noninvasive LG assessment is possible by applying breath-holding maneuvers. However, this technique requires active participation from awake individuals when breathing control and breathing stability can be significantly different compared with sleep (12–14). The literature also describes techniques that use the response to spontaneous respiratory events to measure LG by leveraging mathematical parameter estimation of the underlying breathing mechanics (i.e., PUP) (15–17). Although this method can be used with standard PSG signals, the need for signal preparation and event labeling (manual or automated) currently limits this technique from wide clinical adoption. Clinical home sleep apnea test scoring is heavily weighted to hypoxia-linked events and may therefore underestimate the burden and severity of respiratory instabilities, resulting in errors of estimation if scoring respiratory events is required. However, at least one vendor of sleep testing equipment now integrates automatic endotype assessment from effort signals. Cardiopulmonary coupling can detect HLG by identifying the presence of narrow-band coupling, another expression of SS (18, 19). However, the technique requires 15–17 continuous minutes of self-similar cardiopulmonary coupled oscillations for detection, so brief bursts of periodic breathing are not detected. Table 2 shows an overview of existing LG measurement techniques, including their main advantages and disadvantages. Ultimately, multiple complementary methods applied to individual patient data may be best. In contrast to the manual computation of CAI, our method is fully automatic (with no laborious labeling required) and can be easily applied to signals from laboratory or home PSG; our goal is to have this available for common use as an aid to clinical risk stratification.

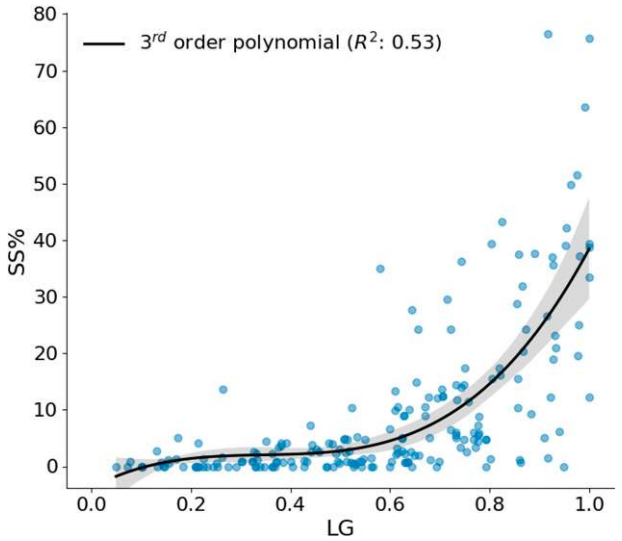

Figure 11. Scatter plot characterizing the relationship between the amount of sleep characterized by a SS score >0.8 and loop gain (LG) estimates. This positive relationship follows a somewhat initial linear trajectory that converges into exponential growth for higher LG estimates. A third-order polynomial was fitted including its 5th to 95th percentile confidence interval after bootstrapping (N = 1,000). SS = self-similarity.

Limitations

It is true that some patients with increased CAI may experience a resolution of central events over time, and the non-REM sleep CO_2 reserve improves over the initial several weeks of CPAP therapy. Respiratory instability short of classic CSA is more common during therapy, contributing to

Table 2. Overview of previously published HLG measurement techniques including their main advantages and disadvantages

Method	References	Advantage	Disadvantage
Hypoxic-hypercapnic gas	5–9	Most direct evaluation of a patient's chemosensitivity	Requires measurement of gas concentrations
Mechanical ventilation	10, 11, 49	Can be used to elicit the HLG threshold in any patient	Requires proportional assist ventilation
Breath-holding	12–14	Completely noninvasive	Requires active participation from awake patients
Spontaneous apneas	15–17	Works on typical airflow signals	Requires signal processing, manual labeling
Cardiopulmonary coupling	18, 19	Cost effectiveness with wearable applicability	Brief bursts of periodic breathing not detected

Definition of abbreviation: HLG = high loop gain.

high residual apnea and low adherence, and treatment-emergent CSA is a risk factor (37). The initial period of CPAP therapy is important for longer-term adherence. A high CAI in the short term can predict poor adherence and continued symptoms, which are detrimental to patient outcomes (51). The long-term effects of detected SS on CPAP outcomes could not be estimated in this study because the majority of patients were treated outside the MGH Sleep Center, limiting access to clinical or device-data outcomes. However, we will soon have access to data from another sleep center and be able to directly assess long-term predictive value. In this work, we exclusively used split-night studies to capture the risk of acute CPAP failure, which does not allow for the confounding effects of intermediate medication and/or treatment or changes in sleep hygiene.

We did not evaluate the whole range of possible combinations of AHI and CAI. One question is whether this method merely detects the severity of disease rather than specific morphological characteristics. HLG can increase the severity of apnea, but SS can be seen in those with predominantly hypopneas or mild hypoxia, suggesting that this metric is not just a severity estimator.

Our model provides a risk score that is monotonic in nature and systematically increases with the amount of SS oscillations found per hour of sleep. With any computational classifier, there is always a trade-off between sensitivity and false-positive rate. In comparison with our original SS method, the new approach is more sensitive, which might lead to a higher inclusion of noise, particularly at low SS scores. This could contribute to the higher failure rates observed in this range. However, comparison with the original model indicates that, regardless of the increased sensitivity, our improved method maintains a robust capability to discern true SS oscillations from noise and better predict acute CPAP failure. Nonetheless, we recognize that, in cases of low SS scores, the probability of noise inclusion is higher, potentially affecting the accuracy.

We acknowledge that the criteria of a 30% amplitude reduction and a 10-second duration for event detection, defined by the AASM, have limitations. Modifying these criteria could possibly improve the prediction of CPAP adherence but would significantly impact the outcomes and complicate the evaluation of the model. Our primary aim is to demonstrate that, in addition to routinely scored respiratory events, the level of SS adds important additional value for stability assessment in patients with apnea. Improving the respiratory event scoring criteria is a priority in our ongoing research. We are developing a novel completely artificial intelligence–based sleep report system that uses a more flexible respiratory event scoring approach, considering variations in the duration and amplitude criteria.

LG can vary significantly within a single overnight study (e.g., greater in the supine position) because, even in patients with severe HLG, some segments may present SS/periodic breathing, whereas other parts of the recording, e.g., during REM sleep, can display classic OSA with a normal LG. Our histogram approach does not capture the temporal trend of LG overnight because it merely summarizes the variation of SS across the entire recording. Nonetheless, our model shows that this summary metric by itself can identify the individuals at risk of suboptimal response to CPAP therapy.

Night-to-night variability of endotypes and stability over time requires further study (52), including this method. The present approach only identifies expressed HLG, making mildly increased LG undetectable in patients with apnea. We did not directly measure LG but used emergent central apneas and SS events as surrogate markers of HLG. SS has been part of the morphological description of periodic breathing/Hunter-Cheyne-Stokes respiration from the time of the first description, patterns with well documented HLG. Our method operationalizes objectively what is evident on visual morphological assessment and accommodates some "impurities" that are typical in clinical samples, that is, with minor deviations from the purest periodic breathing. The analysis is currently solely based on effort signals via RIP, whereas other signals such as oximetry morphology, the photoplethysmographic signal, and airflow may convey additional relevant information. In fact, SS is seen in a range of PSG signals in those with classic Hunter-Cheyne-Stokes respiration. However, the results from this study show that using RIP signals only can differentiate patients who will continue to experience significant respiratory events during CPAP from patients receiving effective therapy with sufficient clinical accuracy. A major advantage of using an effort belt is that it allows for noninvasive application such as home monitoring and monitoring in the intensive care setting.

Conclusions

A clinically applicable method to detect expressed HLG independent of manual scoring of obstructive versus central respiratory events is described. This estimate accurately predicts the acute failure of CPAP, which itself is associated with high residual apnea and CPAP nonadherence. SS may help risk-stratify and identify individuals who could benefit from adjunctive therapy targeting HLG, particularly when affected by centrally driven events that do not meet the criteria for CSA. ■

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

References

1 Lajoie AC, Lafontaine AL, Kimoff RJ, Kaminska M. Obstructive sleep apnea in neurodegenerative disorders: current evidence in support of benefit from sleep apnea treatment. *J Clin Med* 2020;9:297.
2 Zhao X, Li X, Xu H, Qian Y, Fang F, Yi H, *et al.* Relationships between cardiometabolic disorders and obstructive sleep apnea: implications for cardiovascular disease risk. *J Clin Hypertens (Greenwich)* 2019;21: 280–290.
3 Deacon NL, Catcheside PG. The role of high loop gain induced by intermittent hypoxia in the pathophysiology of obstructive sleep apnoea. *Sleep Med Rev* 2015;22:3–14.
4 Lee JJ, Sundar KM. Evaluation and management of adults with obstructive sleep apnea syndrome. *Lung* 2021;199:87–101.
5 McClean PA, Phillipson EA, Martinez D, Zamel N. Single breath of CO_2 as a clinical test of the peripheral chemoreflex. *J Appl Physiol (1985)* 1988;64:84–89.
6 Ghazanshahi SD, Khoo MC. Estimation of chemoreflex loop gain using pseudorandom binary CO_2 stimulation. *IEEE Trans Biomed Eng* 1997; 44:357–366.
7 Hudgel DW, Gordon EA, Thanakitcharu S, Bruce EN. Instability of ventilatory control in patients with obstructive sleep apnea. *Am J Respir Crit Care Med* 1998;158:1142–1149.
8 Younes M, Ostrowski M, Atkar R, Laprairie J, Siemens A, Hanly P. Mechanisms of breathing instability in patients with obstructive sleep apnea. *J Appl Physiol (1985)* 2007;103:1929–1941.
9 Loewen A, Ostrowski M, Laprairie J, Atkar R, Gnitecki J, Hanly P, *et al.* Determinants of ventilatory instability in obstructive sleep apnea: inherent or acquired? *Sleep* 2009;32:1355–1365.
10 Meza S, Younes M. Ventilatory stability during sleep studied with proportional assist ventilation (PAV). *Sleep* 1996;19:S164–S166.
11 Younes M, Ostrowski M, Thompson W, Leslie C, Shewchuk W. Chemical control stability in patients with obstructive sleep apnea. *Am J Respir Crit Care Med* 2001;163:1181–1190.
12 Stanley NN, Cunningham EL, Altose MD, Kelsen SG, Levinson RS, Cherniack NS. Evaluation of breath holding in hypercapnia as a simple clinical test of respiratory chemosensitivity. *Thorax* 1975;30:337–343.
13 Trembach N, Zabolotskikh I. Breath-holding test in evaluation of peripheral chemoreflex sensitivity in healthy subjects. *Respir Physiol Neurobiol* 2017;235:79–82.
14 Messineo L, Taranto-Montemurro L, Azarbarzin A, Oliveira Marques MD, Calianese N, White DP, *et al.* Breath-holding as a means to estimate the loop gain contribution to obstructive sleep apnoea. *J Physiol* 2018; 596:4043–4056.
15 Terrill PI, Edwards BA, Nemati S, Butler JP, Owens RL, Eckert DJ, *et al.* Quantifying the ventilatory control contribution to sleep apnoea using polysomnography. *Eur Respir J* 2015;45:408–418.
16 Nava-Guerra L, Edwards BA, Terrill PI, Sands SA, Amin RS, Kemp JS, *et al.* Quantifying ventilatory control stability from spontaneous sigh responses during sleep: a comparison of two approaches. *Physiol Meas* 2018;39:114005.
17 Finnsson E, Olafsdottir GH, Loftsdottir DL, Jonsson SAE, Helgadottir H, Agustsson JS, *et al.* A scalable method of determining physiological endotypes of sleep apnea from a polysomnographic sleep study. *Sleep* 2021;44:zsaa168.
18 Ramar K, Desrues B, Ramar P, Morgenthaler TI. Analysis of cardiopulmonary coupling to assess adaptive servo-ventilation success in complex sleep apnea management. *Sleep Breath* 2013;17:861–866.
19 Al Ashry HS, Ni Y, Thomas RJ. Cardiopulmonary sleep spectrograms open a novel window into sleep biology-implications for health and disease. *Front Neurosci* 2021;15:755464.
20 Oppersma E, Ganglberger W, Sun H, Thomas RJ, Westover MB. Algorithm for automatic detection of self-similarity and prediction of residual central respiratory events during continuous positive airway pressure. *Sleep* 2021;44:zsaa168.
21 Rao H, Thomas RJ. Complex sleep apnea. *Curr Treat Options Neurol* 2013;15:677–691.
22 Ginter G, Sankari A, Eshraghi M, Obiakor H, Yarandi H, Chowdhuri S, *et al.* Effect of acetazolamide on susceptibility to central sleep apnea in chronic spinal cord injury. *J Appl Physiol (1985)* 2020;128:960–966.
23 Sands SA, Edwards BA, Kee K, Turton A, Skuza EM, Roebuck T, *et al.* Loop gain as a means to predict a positive airway pressure suppression of Cheyne-Stokes respiration in patients with heart failure. *Am J Respir Crit Care Med* 2011;184:1067–1075.
24 Edwards BA, Andara C, Landry S, Sands SA, Joosten SA, Owens RL, *et al.* Upper-airway collapsibility and loop gain predict the response to oral appliance therapy in patients with obstructive sleep apnea. *Am J Respir Crit Care Med* 2016;194:1413–1422.
25 Deacon-Diaz NL, Sands SA, McEvoy RD, Catcheside PG. Daytime loop gain is elevated in obstructive sleep apnea but not reduced by CPAP treatment. *J Appl Physiol (1985)* 2018;125:1490–1497.
26 Op de Beeck S, Wellman A, Dieltjens M, Strohl KP, Willemen M, Van de Heyning PH, *et al.*; STAR Trial Investigators. Endotypic mechanisms of successful hypoglossal nerve stimulation for obstructive sleep apnea. *Am J Respir Crit Care Med* 2021;203:746–755.
27 Wong AM, Landry SA, Joosten SA, Thomson LDJ, Turton A, Stonehouse J, *et al.* Examining the impact of multilevel upper airway surgery on the obstructive sleep apnoea endotypes and their utility in predicting surgical outcomes. *Respirology* 2022;27:890–899.
28 Mulgrew AT, Lawati NA, Ayas NT, Fox N, Hamilton P, Cortes L, *et al.* Residual sleep apnea on polysomnography after 3 months of CPAP therapy: clinical implications, predictors and patterns. *Sleep Med* 2010; 11:119–125.
29 Denotti AL, Wong KK, Dungan GC, Gilholme JW, Marshall NS, Grunstein RR. Residual sleep-disordered breathing during autotitrating continuous positive airway pressure therapy. *Eur Respir J* 2012;39:1391–1397.
30 Reiter J, Zleik B, Bazalakova M, Mehta P, Thomas RJ. Residual events during use of CPAP: prevalence, predictors, and detection accuracy. *J Clin Sleep Med* 2016;12:1153–1158.
31 Sands SA, Edwards BA, Terrill PI, Butler JP, Owens RL, Taranto-Montemurro L, *et al.* Identifying obstructive sleep apnoea patients responsive to supplemental oxygen therapy. *Eur Respir J* 2018;52: 1800674.
32 Mulchrone A, Shokoueinejad M, Webster J. A review of preventing central sleep apnea by inspired CO_2. *Physiol Meas* 2016;37:R36–R45.
33 Edwards BA, Sands SA, Eckert DJ, White DP, Butler JP, Owens RL, *et al.* Acetazolamide improves loop gain but not the other physiological traits causing obstructive sleep apnoea. *J Physiol* 2012;590: 1199–1211.
34 Eskandari D, Zou D, Karimi M, Stenlof K, Grote L, Hedner J. Zonisamide reduces obstructive sleep apnoea: a randomised placebo-controlled study. *Eur Respir J* 2014;44:140–149.
35 Schmickl CN, Landry S, Orr JE, Nokes B, Edwards BA, Malhotra A, *et al.* Effects of acetazolamide on control of breathing in sleep apnea patients: mechanistic insights using meta-analyses and physiological model simulations. *Physiol Rep* 2021;9:e15071.
36 Hedner J, Stenlof K, Zou D, Hoff E, Hansen C, Kuhn K, *et al.* A randomized controlled clinical trial exploring safety and tolerability of sulthiame in sleep apnea. *Am J Respir Crit Care Med* 2022;205: 1461–1469.
37 Ni YN, Thomas RJ. Predictors and consequences of residual apnea during positive airway pressure therapy. *Sleep Med* 2023;106:42–51.

38 Ni YN, Holzer RC, Thomas RJ. Acute and long-term effects of acetazolamide in presumed high loop gain sleep apnea. *Sleep Med* 2023;107:137–148.
39 Troester MM, Quan SF, Berry RB, Plante DT, Abreu AR, Alzoubaidi M, *et al.* The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, version 3. Darien, IL: American Academy of Sleep Medicine; 2023.
40 Nassi TE. Algorithms for automated scoring of respiratory events in sleep [master's thesis]. Enschede, The Netherlands: University of Twente; 2021.
41 Azarbarzin A, Sands SA, Stone KL, Taranto-Montemurro L, Messineo L, Terrill PI, *et al.* The hypoxic burden of sleep apnoea predicts cardiovascular disease-related mortality: the Osteoporotic Fractures in Men Study and the Sleep Heart Health Study. *Eur Heart J* 2019;40: 1149–1157.
42 Thomas RJ, Terzano MG, Parrino L, Weiss JW. Obstructive sleep-disordered breathing with a dominant cyclic alternating pattern—a recognizable polysomnographic variant with practical clinical implications. *Sleep* 2004;27:229–234.
43 Gell LK, Vena D, Alex RM, Azarbarzin A, Calianese N, Hess LB, *et al.* Neural ventilatory drive decline as a predominant mechanism of obstructive sleep apnoea events. *Thorax* 2022;77:707–716.
44 Sankri-Tarbichi AG, Rowley JA, Badr MS. Expiratory pharyngeal narrowing during central hypocapnic hypopnea. *Am J Respir Crit Care Med* 2009;179:313–319.
45 Badr MS, Toiber F, Skatrud JB, Dempsey J. Pharyngeal narrowing/occlusion during central sleep apnea. *J Appl Physiol (1985)* 1995;78: 1806–1815.
46 Thomas RJ, Tamisier R, Boucher J, Kotlar Y, Vigneault K, Weiss JW, *et al.* Nocturnal hypoxia exposure with simulated altitude for 14 days does not significantly alter working memory or vigilance in humans. *Sleep* 2007;30:1195–1203.
47 Lindsey BG, Nuding SC, Segers LS, Morris KF. Carotid bodies and the integrated cardiorespiratory response to hypoxia. *Physiology (Bethesda)* 2018;33:281–297.
48 Eckert DJ, Jordan AS, Merchia P, Malhotra A. Central sleep apnea: pathophysiology and treatment. *Chest* 2007;131:595–607.
49 Dempsey JA. Central sleep apnea: misunderstood and mistreated!. *F1000Res* 2019;8:F1000.
50 Wellman A, Eckert DJ, Jordan AS, Edwards BA, Passaglia CL, Jackson AC, *et al.* A method for measuring and modeling the physiological traits causing obstructive sleep apnea. *J Appl Physiol (1985)* 2011;110: 1627–1637.
51 Pépin J-L, Woehrle H, Liu D, Shao S, Armitstead JP, Cistulli PA, *et al.* Adherence to positive airway therapy after switching from CPAP to ASV: a big data analysis. *J Clin Sleep Med* 2018;14:57–63.
52 Stoberl AS, Schwarz EI, Haile SR, Turnbull CD, Rossi VA, Stradling JR, *et al.* Night-to-night variability of obstructive sleep apnea. *J Sleep Res* 2017;26:782–788.