

Original Article

The respiration–limb coactivation index: a continuous signal-based biomarker for stratifying cardiovascular disease risk in a cross-sectional sleep cohort

Shenghan Wen ¹, Raffaele Ferri ², Qichen Li¹, Wanlu Deng³, Michael Brandon Westover⁴, Robert J. Thomas ^{5,6,†} and Samaneh Nasiri ^{1,*,†}

¹Biomedical Informatics, Emory School of Medicine, Atlanta, GA, United States, ²Clinical Neurophysiology Research Unit, Department of Neurology and Sleep Research Centre, Oasi Research Institute-IRCCS, Troina, Italy, ³Department of Statistics and Data Science, Tsinghua University, Beijing, China, ⁴Department of Neurology and Neurological Sciences, Stanford University, Palo Alto, CA, United States, ⁵Division of Sleep Medicine, Harvard Medical School, Boston, MA, United States and ⁶Department of Medicine, Division of Pulmonary Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, MA, United States

*Corresponding author. Samaneh Nasiri, Biomedical Informatics, Emory School of Medicine, Atlanta, GA, United States. Email: snasiri@dbmi.emory.edu.

†Co-senior author.

Abstract

Study Objectives: Conventional respiratory-related limb movement (RRLM) metrics rely on event definitions and overlap rules that miss continuous respiratory-motor coupling intensity. We developed the respiration–limb coactivation index (RLCI), an algorithmic continuous biomarker, and evaluated its associations with cardiovascular disease (CVD).

Materials and Methods: In Human Sleep Project polysomnography, the RLCI was computed in S0001 ($N = 20737$) as the interaction between respiratory-flow recovery and the phasic leg-EMG envelope. The locked S0001 pipeline, without parameter tuning, was applied to two independent HSP centers (I0004, I0006). Prevalent coronary heart disease (CHD), cerebrovascular disease (CER), and peripheral vascular disease (PVD) associations were assessed by multivariable logistic regression with stage-restricted comparisons against LMI, PLMI, and RRLMI and signal-processing sensitivity analyses.

Results: After covariate adjustment, whole-night RRLMI was not significant for CHD or CER, whereas RLCI metrics, especially in N1 and first-half sleep, remained associated with CER and PVD. RLC N1 showed the strongest CER association (per-1-SD OR 1.11, 95% CI 1.06–1.15), and RLC FirstHalf was associated with PVD (OR 1.07, 95% CI 1.03–1.11). Stage-restricted RRLMI_N1 was significant for all four outcomes, supporting stage-matched benchmarking. Mutual adjustment suggested complementary RLCI information for CER and PVD, while conventional N1 respiratory-related movements better captured CHD. Findings were stable in sensitivity analyses and external validation.

Conclusions: RLCI associations with CER and PVD suggest respiratory-linked motor coactivation captures autonomic surges distinct from cumulative hypoxia or sleep fragmentation and may complement apnea-hypopnea index-based cardiovascular risk stratification.

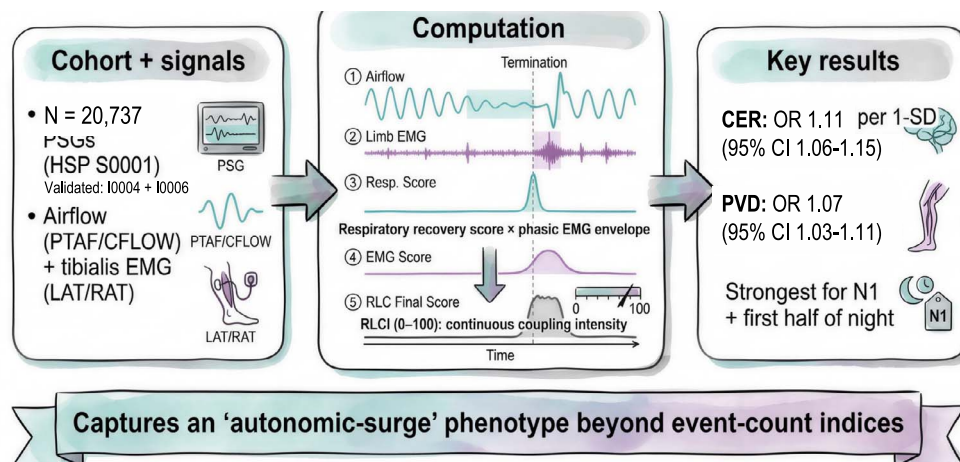
Key words: sleep apnea; respiratory-related leg movements (RRLM); biomarker; cardiovascular disease

Submitted: 1 February, 2026; Revised: 23 May, 2026; Accepted: 5 July, 2026

© The Author(s) 2026. Published by Oxford University Press on behalf of Sleep Research Society. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Graphical Abstract

The Respiration-Limb Coactivation Index (RLCI): A Continuous Signal-Based Biomarker for Stratifying Cardiovascular Disease Risk in a Cross-Sectional Sleep Cohort



Statement of Significance

Current clinical metrics for respiratory-related leg movements rely on discrete event counting and temporal overlap rules, which can be sensitive to event-boundary placement and may fail to capture the physiological magnitude of respiratory-motor coupling. This study introduces the Respiration-Limb Coactivation Index (RLCI), a fully algorithmic, continuous biomarker that quantifies the synchronization between respiratory recovery and limb activation. In a large cross-sectional cohort ($N > 20\,000$), the RLCI demonstrated associations with cerebrovascular and peripheral vascular diseases, independent of traditional sleep apnea severity and leg movement counts. These associations persisted under signal-processing sensitivity analyses and inter-center validation. These findings suggest that the RLCI captures a specific "autonomic surge" phenotype of sleep-disordered breathing, offering a scalable tool for stratifying cardiovascular risk beyond the apnea-hypopnea index.

Introduction

Vertical linking of multiple sleep-subsystems, such as respiration, electrocortical, motor, autonomic, and hemodynamic systems, is typical of the networked state of sleep [1]. Thus, pathological transients in one subsystem are often temporally associated with transients in others. In patients with sleep apnea, a common occurrence is limb movements concomitant with respiratory events, often in close approximation to recovery from apnea/hypopnea [2]. While periodic leg movements of sleep (PLMS) may be classically described as occurring independently of respiratory disturbances, this subset of PLMS temporally synchronized with respiratory event termination gives rise to respiratory-related limb movements (RRLMs) as part of a respiratory-modulated expression of the same motor phenomenon [2, 3]. RRLMs may be usefully viewed as leg movements embedded within an integrated respiratory-motor response. Although these movements are readily observed on polysomnography, they are infrequently quantified in clinical practice, and their long-term cardiovascular implications remain poorly defined, contributing to their frequent misclassification or conflation with conventional PLMS. Within the broader framework of network physiology, respiration, motor activity, cortical arousal, and autonomic regulation form a sleep-stage-dependent system of dynamically interacting subnetworks, and the strength of inter-system coupling, including its degradation under disease, is itself a quantifiable phenotype. Viewed in this light, RRLMs are not isolated motor events but a visible manifestation of respiratory-motor coupling, and any biomarker

that aspires to capture their pathophysiologic significance should ideally measure coupling intensity rather than count discrete events.

An important physiological relevance of PLMs lies in their association with heightened cortical arousal and pronounced sympathetic nervous system activation with nocturnal blood pressure surges [4-7]. The co-occurrence of motor activation can amplify this response [5, 8]. Parasympathetic and sympathetic co-activation may also occur [9]. Studies in large cohorts, such as the Osteoporotic Fractures in Men (MrOS) Sleep Study, have shown that obstructive respiratory events accompanied by RRLMs exhibit significantly larger heart rate surges and greater hemodynamic variability than respiratory events without motor activation [10, 11]. Such exaggerated autonomic responses, often occurring alongside significant oxygen desaturation, are biologically plausible mechanisms linking sleep-disordered breathing (SDB) to hypertension and cardiovascular disease (CVD). Emerging evidence confirms the clinical significance of this coupling: the conventional RRLM index (RRLMI) has been demonstrated to be an independent predictor of all-cause mortality in older men [11], even after accounting for apnea-hypopnea index), suggesting that RRLMs identify a high-risk phenotype that is independent of overall SDB burden.

A central barrier to progress is methodological. Current RRLM definitions (e.g. WASM/IRLSSG criteria [12]) require the precise temporal alignment of a scored respiratory event and a scored leg movement. This approach, which relies on discrete event

definitions and narrow temporal overlap rules, is highly susceptible to boundary-placement variability, and fragile event timing. Research has specifically shown that the resulting RRLM prevalence is highly sensitive to small alterations in the defined latency window—the time interval between event termination and leg movement onset [13]. A shift of only one to two seconds in marking an apnea or leg movement can alter whether an event is labeled as an RRLM or an isolated PLM, rendering the conventional RRLM index unstable and difficult to reproduce across scorers, laboratories, and cohorts.

This binary, event-based scoring approach fails to capture the continuous nature or magnitude of the underlying physiological coupling between the respiratory and motor systems. Reviewers of polysomnograms will notice that a minor respiratory event may have a vigorous motor component, while severe respiratory events can terminate without any motor activation. Thus, there seems to be an interplay between the firing threshold of the motor system and the severity of the respiratory events that cannot be captured by event counts and indices, whether respiratory or motor in isolation.

To overcome these limitations, we developed the Respiration-Limb Coactivation Index (RLCI), a continuous signal-based metric derived from algorithmically processed physiological signals. Rather than relying on discrete, unstable event overlap, the RLCI quantifies the synchronization between respiratory recovery and limb muscle activation, providing an objective and scalable biomarker of this co-activation. We hypothesized that the strength of this co-activation, as a surrogate of sympathetic/hemodynamic surges, would be associated with CVD. To test this hypothesis, we applied the RLCI to a large cross-sectional cohort ($N = 20\,737$ PSG sessions): (1) to establish the distribution of this novel metric across a large and diverse population; (2) to evaluate its association with prevalent CVD, independent of traditional metrics like apnea-hypopnea index (AHI); and (3) to compare its performance against the conventional, event-based metrics under matched analytic conditions.

Materials and methods

Dataset

The Human Sleep Project (HSP) [14] integrates clinical polysomnography (PSG) data from five U.S. institutions: Stanford University, Massachusetts General Brigham (MGB), Emory University, Beth Israel Deaconess Medical Center (BIDMC), and Boston Children's Hospital (BCH). Each study includes physiological signals and expert or automated annotations of sleep stages, arousals, and respiratory and limb-movement events. To preserve institutional anonymity, each site is assigned a unique identifier (S0001, I0002, I0003, I0004, I0006). All data are de-identified under the HIPAA Safe Harbor standard and made publicly accessible (website: <https://bdsp.io/content/hsp/2.0/>) under an approved IRB protocol (BIDMC IRB #2022P000417), which provided a waiver of informed consent.

In this study, we primarily analyzed data from the S0001 cohort, and we used studies from I0004 and I0006 for the purpose of inter-center validation. All polysomnography (PSG) recordings were acquired at 200 Hz with standard Electroencephalogram (EEG) derivations (F3–M2, F4–M1, C3–M2, C4–M1, O1–M2, O2–M1), a bipolar submental Electromyography (EMG) (CHIN1–CHIN2), a single-lead Electrocardiogram (ECG), respiratory flow (pneumotachograph airflow [PTAF] or nasal cannula flow [CFLOW]), thoraco-abdominal effort belts, oxygen saturation, and bilateral anterior tibialis EMG (LAT, RAT). All downstream processing was fully automated and implemented in Python and R. The detailed data

Table 1. Demographic and clinical characteristics of the study cohort from S0001

Characteristic	Overall
Total N	20 737
Female, n (%)	8875 (42.8%)
Male, n (%)	11 862 (57.2%)
Asian, n (%)	723 (3.5%)
Black, n (%)	1338 (6.5%)
Other, n (%)	1680 (8.1%)
White, n (%)	16 996 (82.0%)
CHD Present, n (%)	4607 (22.2%)
PVD Present, n (%)	3805 (18.3%)
CER Present, n (%)	2926 (14.1%)
All CVD Present, n (%)	7441 (35.9%)
Age (years)	53.8 (16.5)
BMI (kg/m²)	31.3 (6.9)
AHI (events/hr)	40.1 (41.2)
Ari (events/hr)	38.4 (17.5)
LMI (events/hr)	81.9 (100.1)
PLMI (events/hr)	44.3 (46.8)
RRLMI (events/hr)	6.5 (7.8)
RRLMI N1 (events/hr)	11.7 (13.5)
RRLMI N2 (events/hr)	7.0 (8.8)
RLC AllSleep	5.6 (5.0)
RLC FirstHalf	6.4 (5.9)
RLC SecondHalf	4.9 (5.6)
RLC REM	5.6 (6.0)
RLC N1	7.3 (5.7)
RLC N2	5.7 (5.1)
RLC N3	4.1 (6.0)

Data are presented as mean (standard deviation) for continuous variables and count (percentage) for categorical variables. Abbreviations: BMI: body mass index; AHI: apnea-hypopnea index; Ari: arousal index; LMI: leg movement index (all leg movements); PLMI: periodic leg movement index; RRLMI: respiratory-related leg movement index (WASM 05d2 criteria). RLC indices represent the mean Respiration-Limb Coactivation score calculated across the specified sleep stages or time periods.

processing pipeline, including subject exclusion criteria based on signal quality and demographic filters, is summarized in Figure 2, with the demographics of the remaining patients summarized in Table 1 for the main cohort S0001, and Table S1 for validation cohort I0004 and I0006.

Pre-processing

For spectral and visualization analyses, EEG and ECG signals were pre-processed using a standardized pipeline. For each available channel, the raw signal was first notch-filtered at 60 Hz to suppress power-line interference. EEG derivations were then band-pass filtered between 0.1 and 20 Hz, whereas ECG was high-pass filtered at 0.3 Hz (finite-impulse-response filters implemented in MNE-Python [15, 16]). When resampling to a common analysis sampling rate of 128 Hz, signals were resampled using polyphase filtering [17] (`scipy.signal.resample_poly` [18]) and truncated or zero-padded as needed to preserve the expected segment length.

Leg EMG signals were processed with a dedicated high-frequency pipeline. Raw EMG in volts was converted to microvolts, demeaned, and notch-filtered at 60 Hz. To attenuate low-frequency drift and movement artifacts while preserving phasic EMG bursts, we applied a high-pass filter with a 10 Hz cut-off. Finally, samples exceeding $\pm 18 \mu\text{V}$ were symmetrically clipped to reduce the impact of rare extreme artifacts (e.g. electrode pops) on subsequent envelope and coactivation calculations. Unless otherwise specified, all EMG processing was performed at the original 200 Hz sampling rate.

Respiratory flow was taken from the nasal pressure signal for diagnostic polysomnography (PTAF: pneumotachometer air flow)

and from positive airway pressure device flow for titration studies (CFLOW). These signals were first low-pass filtered at 10 Hz, then forwarded to the apnea-recovery scoring algorithm (see below) at the native sampling rate, without additional band-pass filtering, so as not to distort the shape of obstructive and hypopneic events.

Automated quality control of leg EMG

Because the RLCI is driven by phasic leg EMG activity, we implemented a fully automated quality-control (QC) procedure to exclude grossly corrupted EMG channels before RLCI computation. QC was applied independently to the pre-processed LAT and RAT signals (after notch + high-pass + clipping; `assess_emg_quality`) and relied on three complementary metrics:

Signal strength (root-mean-square amplitude)

For each EMG channel we computed the root-mean-square (RMS) amplitude over the entire recording,

$$RMS_{\mu V} = \sqrt{\frac{1}{N} \sum_{t=1}^N x_t^2}$$

Where x_t Denotes the Filtered EMG in Microvolts and N the Total Number of Samples. Channels with $RMS_{\mu V} < 1.0V$ Were Considered to Have Insufficient Signal Power (e.g. Detached Electrodes) and Were Rejected.

Flat-line percentage

To detect sustained loss of signal, we identified segments where the point-to-point difference of the EMG trace was numerically indistinguishable from zero over at least 5 s. Specifically, we formed a Boolean sequence, $d_t = |x_{t+1} - x_t| < 10^{-7}$, and grouped consecutive True values into runs. Any run with duration ≥ 5 s (i.e. $\geq 5 \times f_s$ samples, where f_s is the sampling rate) was labeled as flat-line. The flat-line percentage was defined as

$$Flat\% = 100 \times \frac{\text{number of flat - line samples}}{N}.$$

Channels with $Flat\% > 1\%$ were considered unreliable and excluded.

Envelope-based signal-to-noise ratio proxy

To quantify the contrast between phasic EMG bursts and baseline noise, we computed a smoothed EMG envelope using a standard rectification-smoothing procedure (smoothing window 0.1 s). Let e_t denote the resulting envelope. We then extracted the 50th and 90th percentiles of the envelope distribution (P_{50}, P_{90}), and defined “signal” and “noise” subsets as

$$S = \{t : e_t \geq P_{90}\}, \quad N = \{t : e_t < P_{50}\},$$

with corresponding mean envelope amplitudes $\mu_{sig} = \text{mean}(e_t, t \in S)$ and $\mu_{noise} = \text{mean}(e_t, t \in N)$. An SNR proxy was then given by

$$SNR_{proxy} = \frac{\mu_{sig}}{\mu_{noise} + 10^{-6}}.$$

Channels with $SNR_{proxy} < 2.0$ were classified as low-contrast recordings dominated by broadband noise.

Only channels with (1) $RMS_{\mu V} \geq 1.0V$, (2) $Flat\% \leq 1\%$, and (3) $SNR_{proxy} \geq 2.0$ were retained for RLCI computation. Thresholds are determined through distributions and passed over

visual examination (Figure S1). When both LAT and RAT passed QC, we constructed a combined leg EMG envelope as the sample-wise maximum of the two envelopes, thereby capturing unilateral movements without diluting the signal. When only one leg channel passed QC, only that envelope was used. Recordings in which neither LAT nor RAT satisfied QC criteria were excluded from all RLCI analyses. This procedure is fully automated and does not rely on any manual visual scoring or human adjudication.

Complete AI Sleep Report

To ensure consistent and standardized annotations across all eligible sessions, we applied the Complete AI Sleep Report (CAISR) [19] framework (<https://github.com/bdsp-core/CAISR-App>). Among the 19604 unique S0001 patients in the HSP metadata, 14941 patients (76%; 21004 PSG sessions) appeared in the CAISR auto-staging training table, whereas 4663 patients (24%; 5324 sessions) did not. CAISR was used here solely as a pretrained labeling pipeline to generate standardized sleep-stage, respiratory-event, arousal, and limb-movement annotations; the present study does not evaluate CAISR accuracy. Importantly, CAISR labels are used only as inputs to RLCI rather than as the prediction target; the cardiovascular-disease outcomes used in this study are derived from independent ICD-9 records and were not seen during CAISR training. To assess whether training-set overlap influenced the reported associations, we repeated the primary multivariable analyses in the subset of S0001 participants absent from the CAISR training table; these held-out results are summarized in the Limitations. CAISR operates on a predefined set of physiological channels (C3-M2, C4-M1, E1-M2, CHIN1-CHIN2, ABD, CHEST, ECG) and performs task-specific preprocessing tailored to sleep staging, arousal detection, respiratory event classification, and limb-movement detection. CAISR successfully processed 25649 of 26328 candidate recordings (97.4%) and generated four label streams: 30-s sleep stages, 1-s apnea/RERA events, 0.5-s arousals, and 1-s limb movements. The final analytic cohort ($N = 20737$) was established after excluding recordings with insufficient signal quality or missing demographic data, as detailed in Figure 2.

For sleep staging, the model used two EEG channels (e.g. C3-M2, C4-M1, or Cz-Oz), a single EOG channel, chin EMG, abdominal and thoracic effort signals, and ECG. All signals were notch filtered at 60 Hz and band-pass filtered using task-optimized ranges: 0.1–20 Hz for EEG, EOG, and chin EMG; 0–10 Hz for respiratory effort; and 0.3–40 Hz for ECG. These ranges were selected based on prior literature and preliminary analyses showing their suitability for sleep stage classification. Sleep stages were assigned according to AASM guidelines (W, N1, N2, N3, REM). Signals were normalized per subject and per channel using robust scaling (median=0, IQR=1), with values outside $\pm 20 \times \text{IQR}$ clipped to limit the influence of outliers. A graph convolutional network was then used to classify sleep stages [20].

For arousal detection, EEG, EOG, ECG, and chin EMG signals were notch filtered at 60 Hz, band-pass filtered between 0.3 and 35 Hz, resampled to 128 Hz, and clipped at $\pm 20 \times \text{IQR}$. Training and validation data were standardized using robust scaling to achieve zero mean and unit variance, while test data were normalized using a 5-minute rolling robust scaler calibrated on artifacts from a 5000-patient training cohort. Chin EMG underwent additional smoothing using a moving-average filter to attenuate movement artifacts. During REM sleep, arousal detection required a specialized phasic-EMG enhancement pipeline that involved notch filtering chin EMG at 60 Hz, band-pass filtering between 10 and 35 Hz, resampling to 128 Hz, and smoothing to further reduce

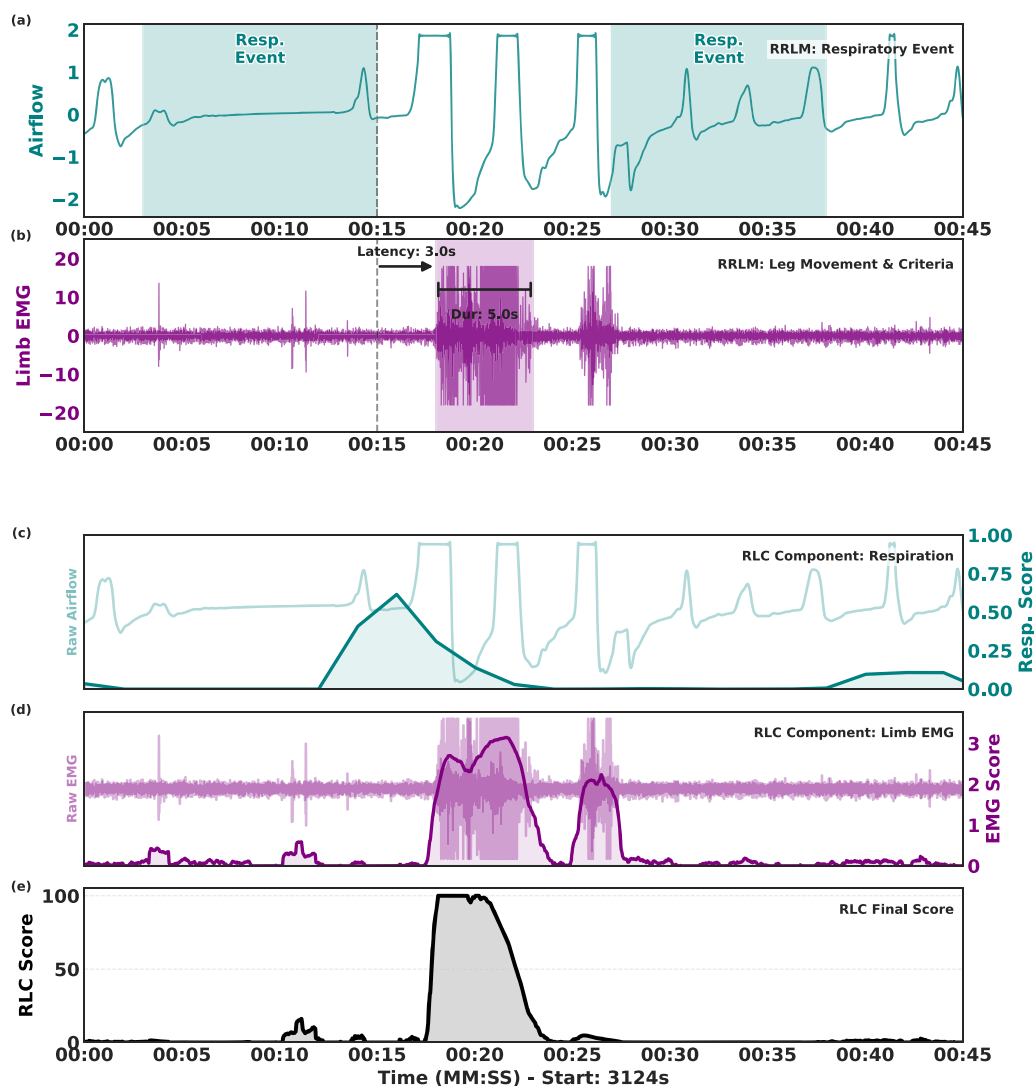


Figure 1. Illustration of RRLM detection versus RLCI computation. Visual comparison of the conventional RRLM scoring and the continuous RLCI within a 1-minute segment. (a, b) Discrete event detection based on WASM criteria (05D2). Panel (a) shows the airflow signal with detected respiratory events (shaded areas). Panel (b) shows the limb EMG signal (in μV), with bilateral signals LAT and RAT merged together; blocks indicate detected leg movements. An RRLM is scored (arrow) because the leg movement overlaps with the post-respiratory event window (0-5 s latency). (c-e) Construction of the continuous RLCI signals. Panel (c) displays the Respiratory Recovery Score, which captures the intensity of flow restoration. Panel (d) shows the Normalized Phasic Leg EMG Activity. Panel (e) presents the final RLC Score (shaded area), calculated as the smoothed multiplicative interaction of (c) and (d), scaled to a 0-100 range. Note how the RLCI provides a continuous measure of coupling intensity rather than a binary count.

noise. A U-Net model inspired by U-Sleep was then used to detect arousals [19, 21].

Respiratory event detection used oronasal airflow or pressure signals, respiratory inductance plethysmography (RIP), and oxygen saturation (SpO_2), all of which were extracted and temporally aligned with sleep stage and arousal annotations. Respiratory signals were notch filtered at 60 Hz, low-pass filtered at 10 Hz, resampled to 10 Hz, and normalized using the mean and standard deviation, with amplitudes clipped to the 5th–95th percentile. To reduce non-physiological artifacts, oronasal airflow segments exceeding $20\times$ the mean were removed, while RIP artifacts were suppressed using a $10\times$ -mean threshold. Segments with missing or uninterpretable respiratory waveforms lasting more than five minutes were automatically masked. Respiratory events were identified using a rule-based algorithm grounded in signal-processing heuristics and AASM criteria [19].

For limb movement detection, bilateral anterior tibialis EMG signals (LAT and RAT) were resampled to 200 Hz and high-pass filtered at 10 Hz using a polyphase implementation to suppress

low-frequency drift and enhance transient muscle activity. Limb movements and periodic limb movements were then detected using a rule-based algorithm derived from Ferri et al. work [22].

Respiratory-related leg movement index

We first quantified RRLMs using a conventional, event-counting index. RRLMs were defined following the World Association of Sleep Medicine (WASM) / IRLSSG framework, adopting the “05D2” variant used by Aritake et al. in the MROS Sleep Study [10]. Specifically, an RRLM was scored when a bilateral anterior tibialis leg movement (duration 0.5–10 s) occurred within a 5-second window before or after the termination of a respiratory event (apnea or hypopnea). The RRLMI was then computed as the number of RRLMs per hour of sleep, during all stages of sleep. An illustrative example of RRLM detection, showing the airflow signal, limb EMG, and the latency/duration window used to classify RRLMs, is provided in Figure 1a and b. In addition to the conventional whole-night index, we computed RRLMI restricted to each individual sleep stage (N1, N2, N3, REM) and the first/second half of the night,

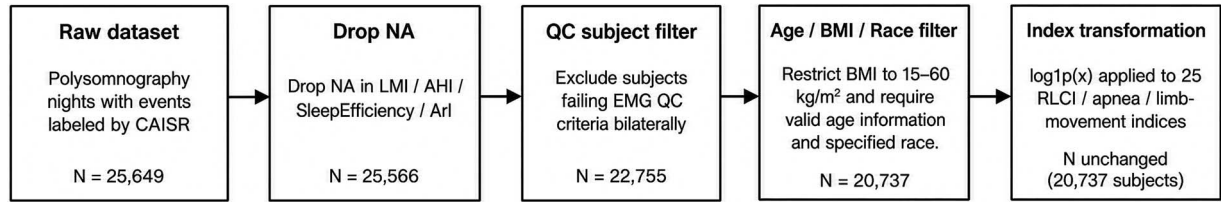


Figure 2. Data processing and subject selection flowchart. The diagram outlines the inclusion and exclusion steps applied to the S0001 dataset from the HSP. Starting with 25 649 recordings processed by the CAISR pipeline, subjects were sequentially excluded based on: NA in LMI, PLMI, AHI, Sleep Efficiency or Erousal Index (Ari), automated leg EMG quality control (QC) failing bilaterally; and BMI (< 15 or > 60 kg/m²), or missing age or BMI or unspecified race data. The final analytic cohort consisted of 20 737 subjects. Continuous indices were $\log_1p(x)$ transformed for statistical modeling.

with the corresponding stratum's hours as the denominator, so that it could be compared against the stage-specific RLCI on a matched footing.

Construction of the respiration–limb coactivation index

To obtain a more continuous and robust biomarker of respiratory–limb coupling, we developed the RLCI, a continuous signal-based measure quantifying the temporal coupling between respiratory recovery and phasic leg EMG activation. Conceptually, the index is derived from three components: (1) a respiratory recovery score, (2) a normalized phasic leg EMG activity signal, and (3) their multiplicative interaction followed by robust scaling.

The computation of the RLCI follows a multi-stage signal-processing pipeline designed to quantify synchronization between respiratory recovery events and limb movements (illustrated in Figure 1c–e). Briefly, in the respiratory stream, raw airflow (PTAF or CFLOW) is processed to detect abrupt increases in flow amplitude characteristic of apnea/hypopnea termination, and a sliding-window comparison yields a respiratory recovery score R_{raw} . In the limb EMG stream, bilateral anterior tibialis EMG is filtered and enveloped to capture phasic leg activation, then transformed into a normalized EMG activity score $E_{norm}(t)$ relative to the background baseline. In the coactivation integration step, these two components are combined such that high RLCI values are generated only when substantial limb activity occurs within the specific temporal window of respiratory recovery.

By operating on continuous signals rather than discrete onset/offset annotations, the RLCI reduces sensitivity to small timing differences in event definition, effectively smoothing over isolated leg movements or respiratory fluctuations and emphasizing arousal-related coactivation.

Respiratory recovery score. To capture the abrupt physiological recovery characterizing apnea termination, we computed a Respiratory Recovery Score (R_{raw}) based on the instantaneous change in flow amplitude from the previously selected respiratory flow channel (PTAF or CFLOW). The raw airflow signal $x(t)$ was first rectified to obtain the flow magnitude $x_{proc}(t) = |x(t)|$. We applied a discrete sliding window analysis with a stride of $\Delta_{stride} = 2.0$ seconds. At each time step t_c , we identified the maximum flow amplitude within a preceding baseline window ($t_c - T_{win}$ to t_c) and a subsequent recovery window (t_c to $t_c + T_{win}$), with window length $T_{win} = 4$ s, denoted as M_{prev} and M_{latter} respectively.

The recovery intensity was derived from the logarithmic ratio of these peaks. This approach linearizes the multiplicative nature of flow recovery (e.g. gasps). The final score was normalized to the interval $[0, 1]$ using a saturation threshold $L_{max} = 5.0$ (corresponding to an approximate 148-fold increase in

amplitude):

$$R_{raw}(t) = \left(1, \frac{\ln \left(\left(1.0, \frac{M_{latter}}{M_{prev} + \epsilon} \right) \right)}{L_{max}} \right)$$

where $\epsilon(10^{-9})$ is a small constant to ensure numerical stability. This metric yields high values specifically when a period of suppressed breathing is immediately followed by a significant surge in airflow.

To incorporate temporal tolerance around respiratory events and reduce sensitivity to boundary definitions, $R_{raw}(t)$ was smoothed using a 10-s rectangular kernel:

$$R(t) = \frac{1}{T_w} \int_{t-T_w/2}^{t+T_w/2} R_{raw}(u) du,$$

with $T_w = 10$ s. In discrete time, this corresponds to a moving-average convolution over $10 \cdot f_s$ samples using the `convolve(..., mode="same")` function for convolution from `scipy.signal` [18]. The resulting signal $R(t)$, highlights sustained respiratory disturbance.

We also explored a respiratory-envelope formulation based on the Hilbert amplitude of airflow during method development [23, 24], but found it less specific for post-event respiratory recovery than the current recovery-score formulation and therefore, we did not use it as the primary respiratory component, which could be shown by the visual comparison of a segment (Figure S2).

Normalized phasic leg EMG activity. For each subject, only EMG channels passing automated QC criteria—(1) $RMS_{\mu V} \geq 1.0V$, (2) $Flat\% \leq 1\%$, and (3) $SNR_{proxy} \geq 2.0$ —were used. If both LAT and RAT passed QC, the combined envelope was taken as their pointwise maximum; if only one passed, that channel was used exclusively. The phasic EMG envelope was extracted as follows:

- (1) Hilbert transform envelope:

$$E_{env}(t) = |\text{Hilbert}(x(t))|,$$

where $x(t)$ is the filtered EMG.

- (2) Smoothing with a moving-average window of duration 1.0 s:

$$E_{smooth}(t) = E_{env}(t) * \text{boxcar}(s).$$

- (3) Percentile normalization and log compression to suppress tonic background activity:

$$E_{norm}(t) = \log \log (1 + (E_{smooth}(t) - P_{25}, 0)),$$

where P_{25} is the 25th percentile of the envelope distribution for that subject.

This transformation strongly emphasizes phasic EMG bursts while minimizing low-level tonic activity.

Respiration–limb coactivation signal. The unscaled coactivation signal was defined as the pointwise product of respiratory disturbance and normalized phasic EMG:

$$C_{base}(t) = R(t) \cdot E_{norm}(t).$$

Thus, $C_{base}(t)$ is high when respiratory disturbance and leg movement co-occur, and near zero when either occurs in isolation.

Robust scaling and final RLCI. To ensure comparability across subjects and recordings, the coactivation signal was scaled using robust percentiles. After median-IQR normalization (via `sklearn.preprocessing.RobustScaler` [25]), the 1st and 99th percentiles of the transformed signal excluding Wake periods, P_1 and P_{99} , were used to map values to a standardized 0–100 range:

$$RLCI(t) = \begin{cases} 0, & Z(t) \leq P_1 \\ 100, & Z(t) \geq P_{99} \\ 100 \times \frac{Z(t) - P_1}{P_{99} - P_1}, & P_1 < Z(t) < P_{99} \end{cases}$$

Values outside this range were clipped.

The final RLCI for each subject was defined as the time-average of $RLCI(t)$ across selected epochs (across all sleep, first half of sleep, second half of sleep, REM, N1, N2, or N3). Crucially, the full pipeline, from preprocessing to QC to RLCI computation, is deterministic and requires no human event scoring or manual tuning.

Clinical outcomes definition

CVD outcomes were identified based on International Classification of Diseases, Ninth Revision (ICD-9) [26] diagnostic codes present in the participants' medical records. We categorized outcomes into three distinct subtypes:

- (1) **Coronary Heart Disease (CHD):** defined by codes for acute myocardial infarction (410.x), other acute/subacute ischemic heart disease (411.x), old myocardial infarction (412.x), angina pectoris (413.x), other chronic ischemic heart disease (414.x), and unspecified CVD (429.2).
- (2) **Peripheral Vascular Disease (PVD):** defined by codes for atherosclerosis (440.x), aortic aneurysm/dissection (441.x), other aneurysm (442.x), other PVD (443.x), arterial embolism/thrombosis (444.x), and other disorders of arteries/arterioles (447.x).
- (3) **Cerebrovascular Disease (CER):** defined by codes for subarachnoid hemorrhage (430.x), intracerebral hemorrhage (431.x), other/unspecified intracranial hemorrhage (432.x), occlusion/stenosis of precerebral (433.x) or cerebral (434.x) arteries, transient cerebral ischemia (435.x), and other acute or ill-defined CER (436.x, 437.x, 438.x).

A composite outcome, **All CVD**, was defined as the presence of any diagnosis across the CHD, PVD, or CER categories.

We summarized all CVD occurrences, physiological indices based on CAISR per AASM, RLCI scores in Table 1.

Statistical analysis: Association with CVD outcomes

To investigate the associations between the novel RLCI metrics and cardiovascular outcomes, we employed a statistical framework designed to adjust for prescribed potential confounders, evaluate whether observed associations were driven by high-burden or extreme values, and assess the incremental information contained in RLCIs beyond traditional limb movement indices, including limb movement index (LMI), periodic LMI (PLMI), and RRLMI. These conventional indices were defined as the number of corresponding events during sleep divided by the overall total sleep time in hour. To address the asymmetry between whole-night conventional indices and the stage-specific RLCI, we additionally recomputed stage-specific LMIs, PLMIs, and RRLMIs, as the average number of corresponding events restricted to the same sleep stages (N1, N2, N3, REM) or night halves used for the RLCI, expressed per hour spent in the corresponding sleep stage. We also evaluated the consistency of the main associations across centers to assess inter-center robustness.

FDR-controlled association via multivariable-adjusted logistic regression.

Given the high correlation between sleep fragmentation, respiratory events, and limb movements, we first performed a comprehensive screening analysis to identify RLCI metrics associated with CVD (CHD, PVD, CER, and All-CVD) independent of standard confounders. To ensure comparability across metrics with varying physiological ranges (e.g. LMI vs. RLC), all indices were first log1p-transformed ($\log_{10}(x+1)$) [27] to approximate normality and then standardized to Z-scores ($z = (x - \mu) / \sigma$). This scaling allows Odds Ratios (OR) to be interpreted as the change in disease risk associated with a one-standard-deviation (1-SD) increase in the log1p-scale index.

For each index and disease outcome, we fitted multivariable logistic regression models adjusted for Age, Sex, Race, BMI, Sleep Efficiency, and log1p-transformed AHI and arousal index (Arl) $\log_{10}[\text{AHI} + 1]$ and $\log_{10}[\text{Arl} + 1]$ to account for sleep apnea severity and fragmentation. To address multiple hypothesis testing, p -values were adjusted using the Benjamini-Hochberg procedure to control the False Discovery Rate (FDR) [28]. Results were visualized using volcano plots to simultaneously display effect size, $\beta = \log(\text{OR})$, and statistical significance $-\log_{10}\text{FDR}$.

High-end effect via adjusted Q4 vs. Q1–Q3 comparison.

To evaluate whether the per-1-SD associations identified in the primary screening were driven specifically by participants at the high end of each metric, we performed a sensitivity analysis comparing participants in the highest quartile (Q4) of each index against the remainder of the cohort (Q1–Q3), a conventional grouped-predictor approach in logistic-regression model building [29]. For each of the four cardiovascular outcomes, we fitted a covariate-adjusted logistic regression model in which the binary Q4 indicator was entered jointly with the same baseline covariates used in the primary analysis: Age, Sex, Race, BMI, Sleep Efficiency, log1p-AHI, and log1p-Arl. Adjusted odds ratios for Q4 vs Q1–Q3 with 95% confidence intervals were derived from the exposure coefficient and its model-based standard error.

Independence analysis via mutual adjustment. To determine whether the novel RLCI metrics provide independent risk-stratification information beyond established markers, we performed pairwise mutual-adjustment analyses using nested multivariable logistic regression. For each cardiovascular outcome (CHD, PVD, CER, and All CVD) and each (RLCI, conventional

index) pair, we fitted two logistic regression models on the same analytic subset, sharing the locked covariate set: a baseline containing the candidate predictor alone and an augmented model with the rival predictor added on top. The five RLCI candidates (RLC N1, RLC N2, RLC N3, RLC FirstHalf, RLC AllSleep) were each tested against the five traditional rivals (whole-night LMI, PLMI, RRLMI, plus the newly added stage-restricted RRLMI N1 and RRLMI N2), and vice versa, yielding 168 augmented comparisons across the four outcomes. Predictors were log1p transformed and z-scored (per 1-SD) to enable direct comparison of effect sizes across metrics. For every augmented model we report four metrics: (1) the per-1-SD adjusted OR with 95% CI and Wald *p*-value for the candidate predictor; (2) a likelihood-ratio test (LRT) *p*-value comparing the augmented model to its baseline, evaluating whether the rival significantly improves model fit; (3) $\Delta\text{AIC} = \text{AIC}(\text{augmented}) - \text{AIC}(\text{baseline})$, with negative values favoring inclusion of the rival; and (4) the variance inflation factor (VIF) of both predictors inside the augmented model as a collinearity diagnostic, with $\text{VIF} < 5$ considered acceptable. We assessed (1) whether RLCI associations persisted after adjustment for each traditional index and whether the traditional rival significantly improved model fit on top of an RLCI-only baseline, and (2) conversely, whether traditional indices retained significance after adjustment for each RLCI metric and whether the RLCI rival significantly improved model fit on top of a traditional-only baseline. Scenario where one direction yielding a significant LRT while the other did not was interpreted as evidence that one index family carries information not recoverable from the other. Symmetric LRT significance, by contrast, was read as cumulative and non-redundant contribution.

All logistic regression models were fitted using the base stats package in R (4.4.2) [30], with the Benjamini–Hochberg procedure applied via *p.adjust* per disease to control the FDR across indices. Data manipulation and visualizations were performed using the *dplyr*, *tidyr*, *broom*, *ggplot2*, *ggrepel*, and *patchwork* packages [31].

Results

Physiological characteristics of the RLCI

Visual inspection of the derived RLCI demonstrated its ability to capture the temporal overlap between respiratory effort and limb movements. As illustrated in Figure 1, the RLC score remains low during isolated respiratory events or isolated limb movements but exhibits a distinct surge when limb muscle activation coincides with the post-apneic respiratory recovery, indicating strong respiratory-related limb movements.

Whole-night analysis (Figure 3) revealed that the density of high-RLCI events was not uniform across the night. Consistent with the known physiology of sleep-disordered breathing, RLCI burden was frequently clustered in N2 sleep and often showed higher density during the first half of the sleep period. The continuous RLCI showed strong alignment with CAISR-labeled limb movements based on AASM criteria, specifically when they occurred in the context of respiratory instability. The 2-minute segments (Figure 4) further showed that high RLCI peaks arise specifically when phasic leg EMG bursts co-occur with the post-event respiratory recovery pattern, whereas isolated limb movements or respiratory fluctuations produce low or modest RLCI values. Furthermore, the largest peaks of RLCI correspond with RRLM events labeled by CAISR; in contrast, for limb movements lagging behind flow recovery that do not qualify as RRLMs, the RLCI peaks are significantly lower. This highlights

that the RLCI captures contextualized respiratory–motor coupling intensity rather than simple event counts.

FDR-controlled RLCI markers association with CVD

To identify which RLCIs were most strongly linked to cardiovascular pathology independent of standard confounders, we conducted a screening analysis using multivariable-adjusted logistic regression (Stage 1; see Methods). For each RLCI and conventional limb-movement metric (PLMI, LMI, RRLMI), odds ratios (ORs) were estimated per 1-standard deviation (1-SD) increase in the log1p-transformed metric at both the total sleep and stage-specific levels. We applied the Benjamini–Hochberg procedure to control the FDR at 5% within each disease category, jointly across all whole-night and stage-restricted indices (Figure 5a, Figure S3, and Table S2). E-values and lower-confidence-bound E-values (E_{CI}) were additionally reported as sensitivity metrics, representing the minimum strength of association that an unmeasured confounder would need to have with both the index and the outcome to explain away the observed association or its confidence-limit evidence.

At the whole-night level, conventional limb-movement indices showed strong associations across outcomes, especially CHD, PVD, and All CVD. LMI was significantly associated with all four outcomes with $\text{FDR} < 0.05$, with standardized ORs ranging from 1.06 to 1.09. Similarly, PLMI showed consistent positive associations across the four phenotypes. These conventional whole-night associations had modest E-values, generally around 1.20–1.26, indicating that relatively modest unmeasured confounding could attenuate the point estimates, although several lower-bound E-values remained above the null. RRLMI was disease-specific: significant for PVD (OR 1.10, $\text{FDR} = 0.002$) and All CVD (OR 1.05, $\text{FDR} = 0.030$), but not for CHD ($\text{FDR} = 0.12$) or CER ($\text{FDR} = 0.12$). Under stage-restriction, LMI and PLMI behaved as essentially temporally homogeneous markers: their ORs varied within a narrow range across N1, N2, N3, REM, and night-half strata (e.g. for All CVD, LMI ORs ranged 1.04 to 1.07 and PLMI ORs ranged 1.04 to 1.08), indicating that LMI and PLMI primarily index a generic motor-burden signal that is relatively insensitive to temporal stratification. By contrast, RRLMI was more stage-dependent. RRLMI N1 retained associations across all four outcomes, with significant effects for CHD (OR 1.09, $\text{FDR} = 8.5 \times 10^{-4}$), PVD (OR 1.08, $\text{FDR} = 1.5 \times 10^{-3}$), CER (OR 1.06, $\text{FDR} = 0.045$), and All CVD (OR 1.08, $\text{FDR} = 6 \times 10^{-4}$), with E-values ranging from 1.20 to 1.27. In contrast, RRLMI N2 lost significance for CHD, CER, and All CVD, with only PVD remaining significant (OR 1.09, $\text{FDR} = 2.8 \times 10^{-3}$), while RRLMI metrics in N3 and REM were null across all four outcomes.

These conventional-index results provided a stage-matched reference for evaluating whether the RLCI captured cardiovascular-relevant information beyond generic motor burden, represented by LMI/PLMI, and beyond discrete respiratory-related limb-movement counts, represented by RRLMI. In CER, where the stage-restricted differences were most apparent, RLC N1 showed a larger point estimate than stage-matched LMI N1 and PLMI N1, although the confidence intervals overlapped: RLC N1 was associated with CER with an OR of 1.11 (95% CI 1.06–1.15, $\text{FDR} = 4.0 \times 10^{-5}$), compared with LMI N1 (OR 1.07, 95% CI 1.02–1.11, $\text{FDR} = 0.017$) and PLMI N1 (OR 1.07, 95% CI 1.03–1.13, $\text{FDR} = 0.012$). The corresponding E-value was also larger for RLC N1 in CER (E-value 1.29, E_{CI} 1.21) than for LMI N1 (1.22, 1.11) or PLMI N1 (1.23, 1.13), supporting a somewhat stronger sensitivity profile for the RLCI signal in this stratum. A clearer divergence was observed in N3, where stage-matched LMI N3 and PLMI N3

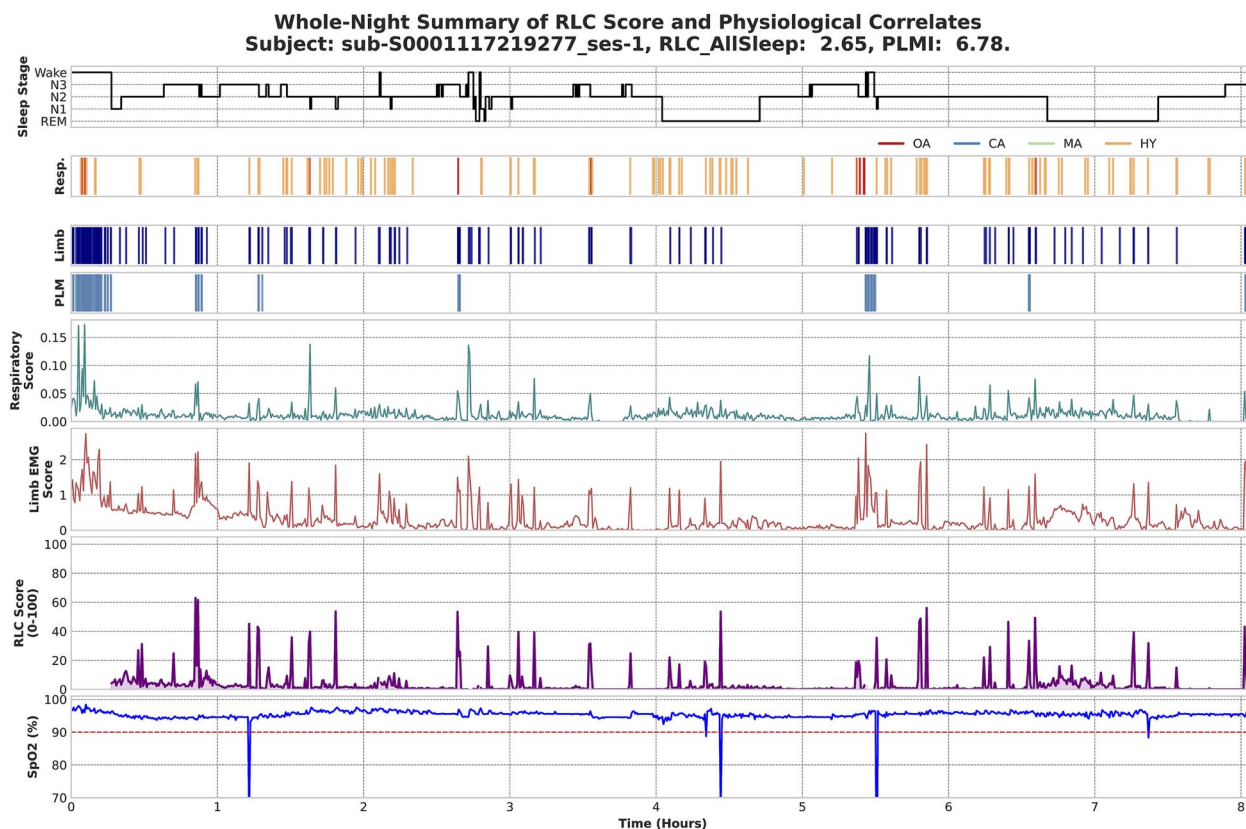


Figure 3. Whole-night visualization of physiological signals and RLC burden. A representative whole-night recording illustrating the distribution of sleep stages and biomarkers. From top to bottom: (1) Hypnogram showing sleep architecture (Wake, N3, N2, N1, REM). (2) Raster plot of discrete respiratory events (OA = Obstructive Apnea, CA = Central Apnea, MA = Mixed Apnea, HY = Hypopnea). (3) Raster plot of discrete limb movements (Limb) and (4) periodic limb movements (PLM). (5) Continuous Respiratory Score. (6) Continuous Limb EMG Score. (7) The derived RLC Score excluding wakefulness. (8) Oxygen Saturation (SpO₂, blue trace, red dashed line indicating 90%). Periods of high Respiration-Limb Coactivation, which typically cluster during N2 sleep and align with repetitive desaturations and respiratory instability, contrasting with periods of stable breathing or isolated limb movements.

were not significant for CER (LMI N3: OR 1.03, FDR=0.27; PLMI N3: OR 1.03, FDR=0.19), whereas RLC N3 remained significant (OR 1.08, 95% CI 1.03–1.12, FDR=0.004; E-value 1.23, E_CI 1.14). RLC N3 also retained significant associations with PVD (OR 1.07, FDR=0.002; E-value 1.21, E_CI 1.13), whereas RRLMI N3 was null for both CER and PVD (CER: OR 1.01, FDR=0.56; PVD: OR 1.02, FDR=0.22). Together, these stage-matched comparisons suggest that the RLCI preserves respiratory-motor coupling information in sleep-stage strata where conventional motor-burden indices are nonspecific and event-count-based RRLMI becomes sparse or unstable.

Among the novel RLCIs, distinct disease-specific patterns emerged. RLC N1 was significantly associated with all four outcomes (ORs 1.05–1.11, FDR < 0.05), and RLC FirstHalf with PVD, CER, and All CVD (ORs 1.04–1.09, FDR < 0.05). RLC N2 displayed particularly significant associations with CER (OR 1.08, FDR=0.004) and PVD (OR 1.06, FDR=0.004). RLC N3 was significantly associated with CER, PVD and All CVD (ORs 1.04–1.08, FDR < 0.05). Across FDR-significant RLCI associations, point-estimate E-values ranged from approximately 1.15 to 1.29, with the largest sensitivity values observed for RLC N1 in CER (1.29), RLC FirstHalf in CER (1.26), and RLC AllSleep in CER (1.24). The lower-confidence-bound E-values were smaller, ranging from approximately 1.04 to 1.21, indicating modest but non-negligible robustness to residual confounding. Notably, the RLCI averaged during REM sleep (RLC REM) did not pass the FDR threshold for any outcome, consistent with the distinct physiology of REM

atonia. Based on these findings, RLC N1, RLC N2, RLC N3, RLC FirstHalf, and RLC AllSleep were selected as primary markers for subsequent robustness testing. Accordingly, Figure 5b and c display the fitted probability curves (with 95% CIs) and ORs for these five selected RLC indices alongside baseline metrics.

High-end effect analysis via adjusted Q4 vs. Q1–Q3 comparison

To test whether the observed associations were driven by individuals at the high end of each metric, we performed an analysis comparing participants in the highest quartile (Q4) of each index against the remainder of the cohort (Q1–Q3) under the same covariate adjustment used in the primary screening (see Methods; Figure 6, Table S3). E-values were again used to quantify the sensitivity of the observed Q4 associations to potential unmeasured confounding.

The principal patterns from the primary analysis were preserved under this dichotomized framing. LMI and PLMI remained robustly associated with all four cardiovascular outcomes (LMI CHD: OR 1.16, 95% CI 1.07–1.26; LMI All CVD: OR 1.17, 95% CI 1.09–1.26; PLMI all four $p < .001$). The E-values for LMI and PLMI ranged from 1.32 to 1.43, with lower-bound E-values from 1.14 to 1.31, suggesting modest-to-moderate robustness of these high-burden motor signals. The whole-night RRLMI retained significance for three of the four outcomes (PVD: OR 1.17, 95% CI 1.06–1.28, $p = .001$; CER: OR 1.15, 95% CI 1.03–1.27, $p = .010$; All CVD:

Subject: sub-S0001113744485_ses-1 | Window: 3084-3204s

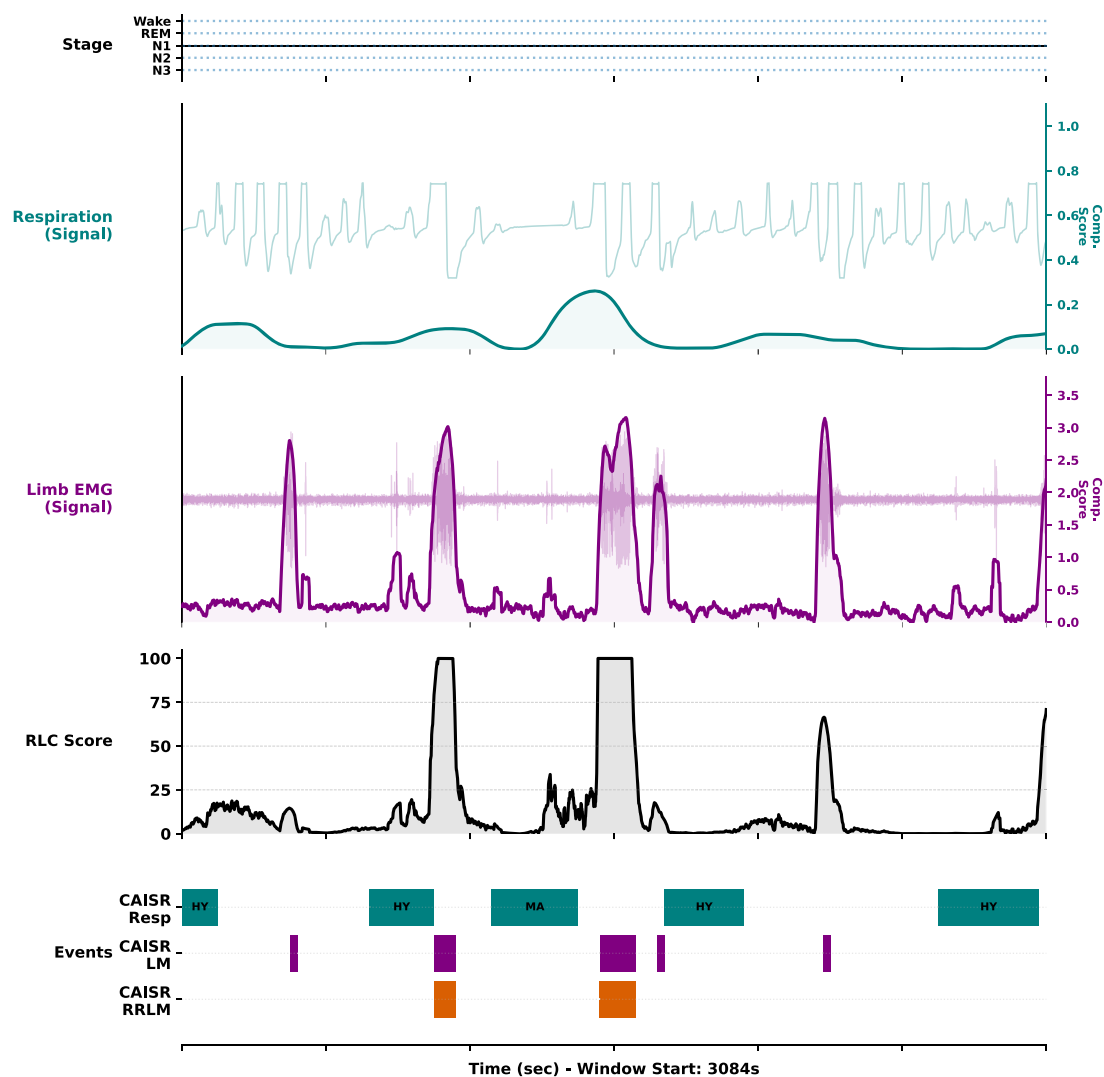


Figure 4. Comparison of continuous RLC score with CAISR event labels in a representative 2-minute segment. The airflow-derived respiratory recovery score (top) and normalized leg EMG activity (middle) jointly produce the continuous RLC score (0-100; excluding wake). Event rasters at the bottom show CAISR-labeled respiratory events, limb movements, and RRLM events based on WASM 05D2 criteria. Peaks in RLC occur when phasic leg EMG bursts align temporally with respiratory recovery, while isolated limb movements or respiratory events without concurrent leg activation yield low RLCI, illustrating how RLCI quantifies coupling intensity rather than binary RRLM counts. Each grid indicates 20 s.

OR 1.12, 95% CI 1.03–1.22, $p = .006$), with only CHD attenuating to non-significance (OR 1.08, 95% CI 0.98–1.18, $p = .11$). The stage-stratified RRLMI N1 produced the strongest profile of all traditional indices, surpassing the whole-night RRLMI on every outcome (CHD: OR 1.16, 95% CI 1.07–1.27, $p < .001$; PVD: OR 1.19, 95% CI 1.09–1.30, $p < .001$; CER: OR 1.18, 95% CI 1.07–1.30, $p < .001$; All CVD: OR 1.20, 95% CI 1.12–1.30, $p < .001$), with E-values ranging from 1.37 to 1.42 and lower-bound E-values from 1.22 to 1.30. This reinforces the inference that the RRLMI is concentrated in the N1 stratum. RRLMI N2 had significance for CER (OR 1.15, 95% CI 1.04–1.28, $p = .008$), but was not significant for CHD, PVD, or All CVD.

The novel RLC metrics demonstrated broad, stratum-specific consistency under the same adjustment. RLC N3 reached significance across all four outcomes (CHD: OR 1.11, 95% CI 1.02–1.21, $p = .012$; PVD: OR 1.14, 95% CI 1.05–1.24, $p = .002$; CER: OR 1.16, 95% CI 1.05–1.27, $p = .002$; All CVD: OR 1.09, 95% CI 1.01–1.17, $p = .019$), with E-values of 1.26–1.36 and lower-bound E-values of

1.09–1.19, supporting a robust high-end association for deep-sleep respiratory–motor coactivation. RLC FirstHalf retained strong significance for CER (OR 1.17, 95% CI 1.07–1.29, $p < .001$; E-value 1.39, E_CI 1.23) and reached significance for PVD (OR 1.10, 95% CI 1.01–1.19, $p = .027$; E-value 1.27, E_CI 1.08), with All CVD approaching significance (OR 1.06, $p = .099$). RLC N1 reached significance for CHD (OR 1.09, 95% CI 1.00–1.18, $p = .043$; E-value 1.25, E_CI 1.04) and PVD (OR 1.09, 95% CI 1.01–1.19, $p = .033$; E-value 1.27, E_CI 1.06), and RLC N2 reached significance for PVD (OR 1.09, 95% CI 1.00–1.18, $p = .047$; E-value 1.25, E_CI 1.02) and CER (OR 1.10, 95% CI 1.01–1.21, $p = .036$; E-value 1.28, E_CI 1.06). The whole-night RLC AllSleep showed more modest but directionally consistent associations with PVD (OR 1.09, 95% CI 1.00–1.18, $p = .040$; E-value 1.26, E_CI 1.05) and CER (OR 1.11, 95% CI 1.02–1.22, $p = .019$; E-value 1.30, E_CI 1.10). Together, these findings indicate that the leg-movement and respiratory-effort link captured by the RLCI shows a complementary high-end risk pattern not fully mirrored by traditional movement counts, with the broadest signal arising

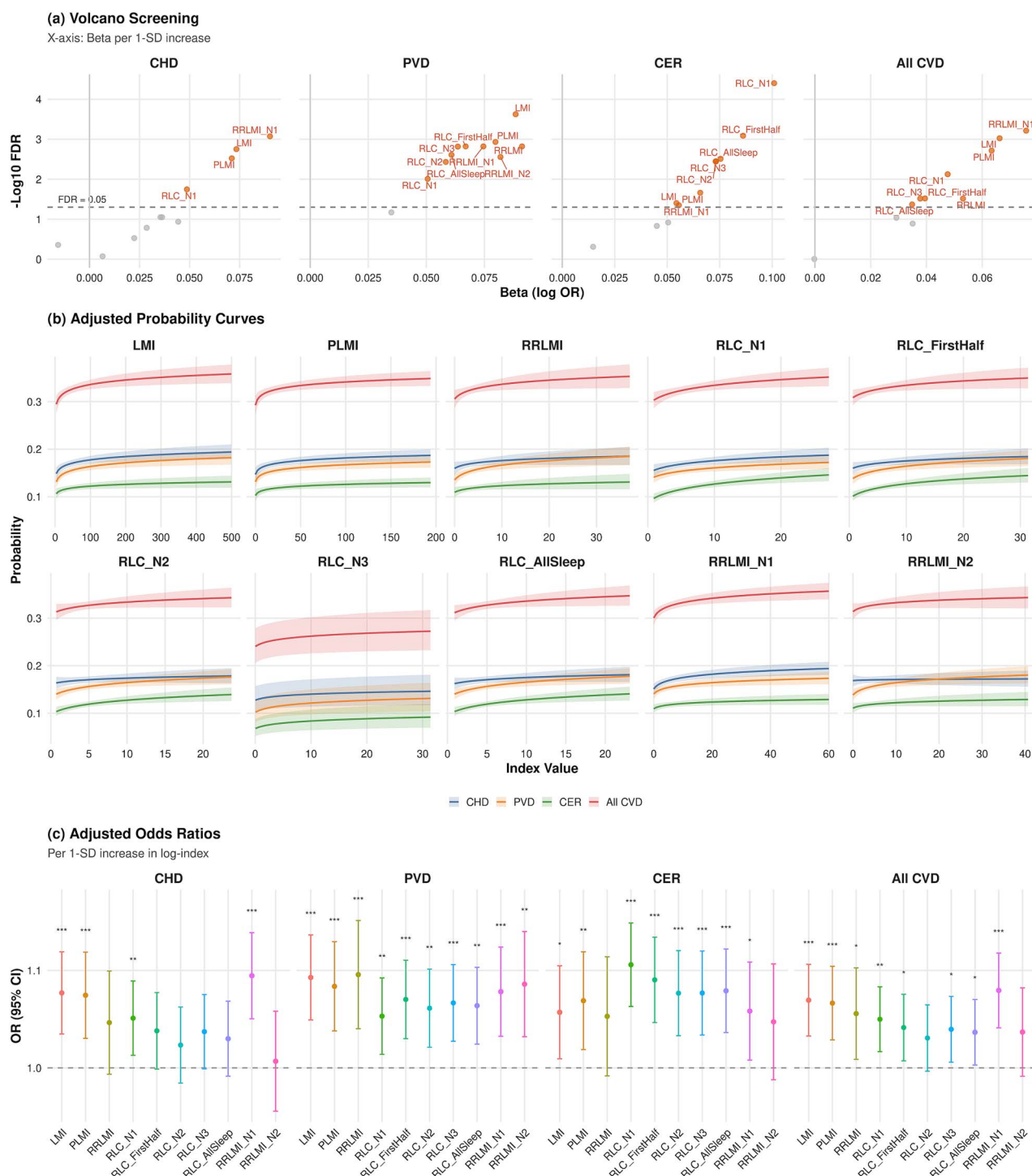


Figure 5. Multivariable screening and characterization of RLCI associations with cardiovascular outcomes. (a) Volcano Screening: Results from the Stage 1 screening analysis. Each point represents a leg movement metric, including LMI, PLMI, whole-night RRLMI, the stage-restricted RRLMI N1 and RRLMI N2, and the RLCIs; those above the dashed line ($FDR < 0.05$) are statistically significant. RLC metrics show distinct significance patterns, particularly for PVD and CER. (b) Adjusted Probability Curves: Predicted probability of disease outcomes across the physiological range of selected indices, adjusted for Age, Sex, BMI, Sleep Efficiency, and log₁₀transformed AHI and ArI. Shaded regions indicate 95% confidence intervals. (c) Standardized Odds Ratios: Forest plot displaying the adjusted Odds Ratio (OR) per 1-standard deviation (1-SD) increase in the log₁₀-transformed index. Significance levels: * $p < .05$, ** $p < .01$, *** $p < .001$.

from deep-sleep coactivation (RLC N3) and the strongest single-outcome signal arising from RLC FirstHalf for CER. At the same time, the E_CI values remained modest, so these results should be interpreted as supportive rather than definitive evidence against residual confounding.

Independence from traditional leg movement indices (LMI/PLMI/RRLMI /RRLMI N1/RRLMI N2)

To determine whether RLC metrics provide prognostic information beyond traditional limb movement indices, including the newly added stage-matched RRLMI N1/RRLMI N2, we fit

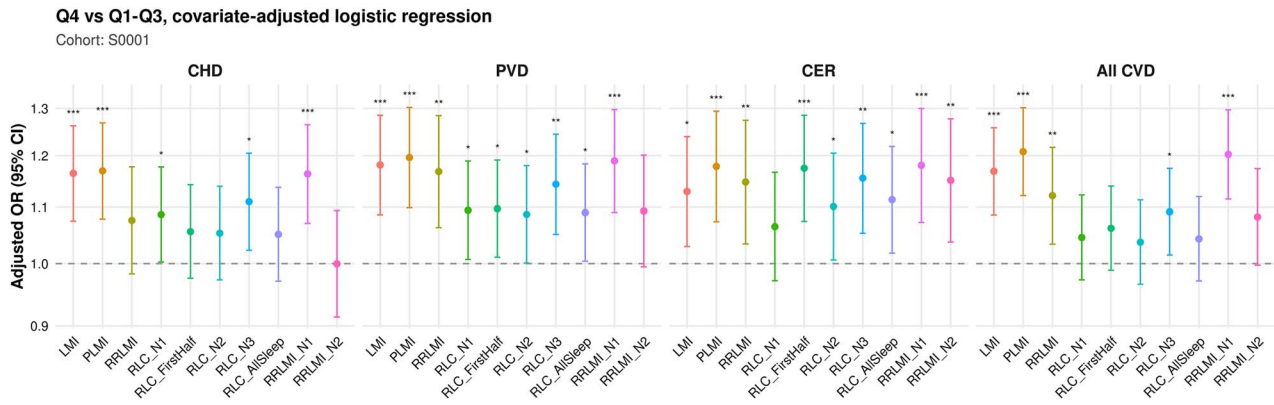


Figure 6. Robustness of associations using adjusted Q4 vs Q1-Q3 logistic regression. Sensitivity analysis comparing the risk of cardiovascular outcomes in participants with high leg movement burden (Top Quartile, Q4) versus the remainder of the cohort (Q1-Q3). For each outcome, a covariate-adjusted logistic regression model was fitted with the binary Q4 indicator entered jointly with the same baseline covariates used in the primary screening (Age, Sex, Race, BMI, Sleep Efficiency, log1p-AHI, log1p-ArI); 95% confidence intervals were derived from model-based standard errors. Notably, while the whole-night RRLMI attenuates to non-significance for CHD under this adjustment, the stage-stratified RRLMI N1 surpasses the whole-night metric across all four outcomes, and novel RLCs, particularly RLC N3 across all four cardiovascular endpoints and RLC FirstHalf for CER, retain statistical significance, supporting the independence of the respiratory-motor coupling signal from overall movement burden and from sleep apnea severity. Significance levels: * $p < .05$, ** $p < .01$, *** $p < 0.001$.

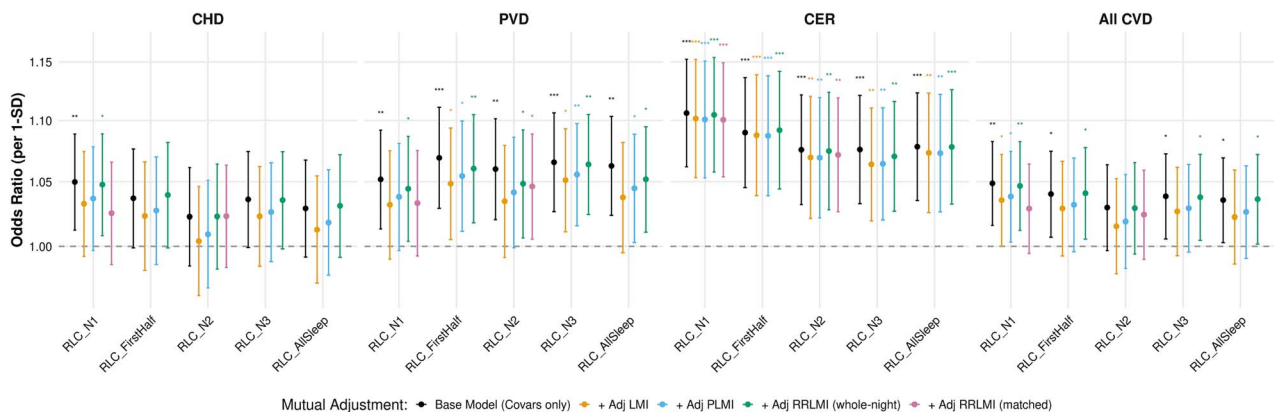


Figure 7. Stability of RLCI predictive power under mutual adjustment (Competition Models). Assessment of the independence of RLCI scores from traditional leg movement markers. Each cluster displays the Odds Ratio (per 1-SD) for the specified RLC metric across up to five model specifications, presented from left to right: (1) Base Model: Adjusted for demographic and sleep covariates only. (2.5) Mutual Adjustments: Additionally adjusted for (2) LMI, (3) PLMI, (4) the whole-night RRLMI, or (5) the corresponding stage-restricted RRLMI (available only for the RLC N1 / RRLMI N1 and RLC N2 / RRLMI N2 pairings). Key patterns of disease specificity are observed: For CER, RLC associations remain stable and highly significant regardless of adjustment, indicating RLC captures risk independent of movement frequency. For PVD, RLC metrics retain independent significance. Conversely, for CHD, RLC associations are attenuated by adjustment for PLMI, suggesting frequency is the dominant driver for coronary outcomes. Significance levels: * $p < .05$, ** $p < .01$, *** $p < .001$.

head-to-head competition models. For each (candidate, rival) index pair, we estimated two nested logistic regressions sharing the same covariate set—one containing the candidate alone, and one adding the rival on top—and tested whether the rival meaningfully improved the fit (Figure 7; Table S4). Each head-to-head comparison was characterized by four metrics: the per-1-SD adjusted Wald OR, a likelihood-ratio test (LRT) p -value evaluating whether the rival predictor improves model fit, Δ AIC against the baseline that omits the rival, and the variance inflation factor (VIF) for both predictors in the augmented model. Across all 168 augmented comparisons the maximum VIF was 2.24 (median 1.48), well below the conventional 5/10 concern thresholds, so the head-to-head ORs are not distorted by multicollinearity.

CER: RLC metrics carried the dominant predictive signal. The association between RLC FirstHalf and CER remained highly significant (OR 1.08–1.09 per 1-SD, $p < .001$) even after adjusting for PLMI, LMI, RRLMI, or RRLMI N1, and none of these traditional rivals improved model fit when added on top of RLC FirstHalf +

covariates (LRT $p = .288, 0.122, 0.453$ against LMI, PLMI, RRLMI; Δ AIC = +0.87, −0.39, +1.44). Conversely, in reverse analyses, the association between traditional PLMI and CER lost statistical significance ($p = .123$) when adjusted for RLC FirstHalf, and adding RLC FirstHalf delivered a strongly significant improvement in fit against every traditional rival (LRT $p < .001$; Δ AIC = −9.3 to −12.5 across LMI / PLMI / RRLMI / RRLMI N1 / RRLMI N2). The same one-sided pattern held for RLC N1, which retained OR ≈ 1.10 ($p < .001$) under all five traditional adjusters, while none of those adjusters added measurable fit (rival LRT $p = .192$ –.488, Δ AIC $\geq +0.30$). This directional asymmetry suggests that for cerebrovascular risk, the continuous interaction of limb movement and recovery of flow carries information that event-rate-based indices cannot recover.

PVD: RLC metrics provided independent and bidirectional value. RLC FirstHalf ($p = .015$) and RLC N3 ($p = .008$) retained independent significance after adjusting for PLMI, and the LRT confirmed mutual non-redundancy: adding PLMI on top of RLC N3 + covariates yielded LRT $p = .003$ (Δ AIC −6.69), while adding

RLC N3 on top of PLMI + covariates yielded LRT $p = .008$ ($\Delta AIC = -5.11$); the same bidirectional improvement held for RLC FirstHalf vs PLMI (LRT $p = .007$ and $.015$) and for RLC AllSleep vs PLMI (LRT $p = .008$ and $.033$). VIFs across these models stayed between 1.10 and 1.45. This indicates that the RLCI, the continuous coactivation, and conventional movement frequency contribute cumulatively rather than redundantly to peripheral vascular risk.

CHD: Conventional metrics were dominant. The association between RLC N1 and CHD was attenuated when adjusted for LMI or PLMI ($p = .126$ when adjusted for LMI, $.083$ when adjusted for PLMI, compared with $.009$ unadjusted) and was fully absorbed by the stage-matched RRLMI N1 (OR 1.03, $p = .207$). LMI and PLMI retained high significance ($p < .01$) in all models. RRLMI N1 was the strongest CHD comparator (OR 1.08–1.09, $p < .001$ across all RLC adjusters). The LRT made the asymmetry explicit: no RLC variant improved fit on top of LMI (LRT $p = .126$ –.899, $\Delta AIC = -0.34$ to $+1.98$) or on top of RRLMI N1 (LRT $p = .077$ –.571, $\Delta AIC \geq -1.14$), whereas LMI (LRT $p = .003$, $\Delta AIC = -6.95$), PLMI (LRT $p = .007$, $\Delta AIC = -5.39$), and especially RRLMI N1 (LRT $p < .001$, $\Delta AIC = -10.40$) each significantly improved fit when added on top of RLC N1 + covariates. This suggests that CHD risk is more tightly linked to the hemodynamic load of frequent N1 movements—captured most efficiently by RRLMI N—than to the averaged link to flow recovery.

All CVD: Traditional indices retained the primary signal, with the RLCIs contributing an incremental rather than independent value. LMI, PLMI, and RRLMI N1 retained tight significance under every RLC adjuster ($p \leq .007$ throughout; OR 1.05–1.08 per 1-SD), and each of these three traditional indices significantly improved fit when added on top of RLC N1 + covariates (LRT $p = .003$, $.007$, $.001$ for LMI / PLMI / RRLMI N1; $\Delta AIC = -6.53$, -5.20 , -8.56). In the reverse direction, no RLC variant improved fit on top of LMI (LRT $p \geq .136$, $\Delta AIC \geq -1.32$) and only RLC N1 marginally improved fit on top of PLMI (LRT $p = .045$, $\Delta AIC = -2.00$). RLCI did, however, retain incremental value against the whole-night RRLMI (RLC N1 / RLC N3: LRT $p = .013$ –.041, $\Delta AIC = -4.2$ to -2.2) and against RRLMI N2 (RLC N1 / RLC N3: LRT $p = .007$ –.025, $\Delta AIC = -5.3$ to -3.0). Consistent with All CVD being a composite outcome that aggregates hemodynamic load from frequent N1 movements across vascular territories, the continuous-coactivation signal captured by RLCI supplements rather than replaces traditional event-rate measures for global cardiovascular risk.

Robustness to signal-processing hyperparameters

To assess whether the observed RLCI-CVD associations were dependent on specific signal-processing parameter choices, we re-computed the RLCI under 19 configurations spanning all major hyperparameters of the pipeline; apnea-recovery stride, pre/post window, and saturation; respiratory-blur window; EMG envelope smoothing, baseline percentile, and log-compression; and dynamic-range clipping percentiles, including three joint perturbations of the respiratory, EMG, and dynamic-range subsystems (Methods; Table S5). Adjusted ORs were recomputed for each configuration in a randomly selected representative subset of $N = 4846$ PSG recordings. Across all 19 configurations, the standardized OR for every (stratum $\times$ outcome) cell varied by less than 3% from the baseline value (coefficient of variation $< 1.3\%$ in all cells), and the direction of association was preserved in every stratum that was significant in the main analysis. For example, the association between RLC FirstHalf and PVD ranged from an OR of 1.072 to 1.101 (baseline 1.092), and the association between RLC N1 and CHD ranged from an OR of 1.062 to 1.095 (baseline 1.088). These results indicate that the reported RLCI associations are not artifacts of any single parameter choice and that the

index is robust to plausible perturbations of its underlying signal-processing pipeline.

Inter-center validation (I0004, I0006)

To assess external validity, we re-ran the locked S0001 pipeline on two additional HSP centers chosen for their demographic and measurement contrast with S0001. All RLCI signal-processing parameters and analytic-cohort exclusion criteria were locked on S0001 prior to applying the pipeline to I0004 and I0006; no parameter tuning or model selection was performed on the validation cohorts. All detailed results are in Table S6 (S6.1 for per-SD primary regression; S6.2 for Q4-vs-Q1–Q3 sensitivity; S6.3 for mutual adjustment of RLCIs and conventional limb indices).

I0004 ($N = 10\,503$ sessions; 58.4% male; 68.6% White, 14.9% Asian; age 48.4 ± 21.0 y; BMI 28.1 ± 6.8) provided a younger, leaner, more Asian-enriched sample carrying the full outcome set (CHD 19.0%, PVD 15.7%, CER 14.4%, CVD 31.5%); limb-events scored by CAISR at this center were comparable to S0001 (LMI 135.3 ± 151.2 events/hr; PLMI 47.8 ± 44.7 ; RRLMI 5.9 ± 8.4). I0006 ($N = 5427$ sessions; 44.4% male; 49.6% White, 47.9% Black; age 55.4 ± 14.6 y; BMI 34.1 ± 8.3) is a smaller and demographically distinct cohort with CER the only available ICD endpoint (CER 13.0%); its RLC burden was systematically lower than at S0001 (RLC AllSleep: I0006 4.6 ± 3.1 vs S0001 5.6 ± 5.0), consistent with site-level scoring/recording differences. Full demographics and index distributions are in Table S1.

In I0004 the primary S0001 signals replicated, with effect sizes broadly comparable to S0001 on the per-SD log-index scale. RLC N1 was the most consistent external signal (CHD OR 1.07 [1.01–1.13], $p = .019$; PVD 1.07 [1.01–1.14], $p = .018$; CER 1.09 [1.02–1.15], $p = .008$; CVD 1.07 [1.02–1.13], $p = .003$). RLC N2 (PVD 1.08, CER 1.10, CVD 1.07; all $p \leq .007$), RLC N3 (PVD 1.06, CVD 1.06; both $p \leq .049$), RLC FirstHalf (PVD 1.07, CVD 1.06; both $p \leq .019$), and RLC AllSleep (PVD 1.07, CER 1.07, CVD 1.06; all $p \leq .025$) were also significant, with confidence intervals overlapping S0001. A pre-specified high-end driven Q4-vs-Q1–Q3 analysis at I0004 was directionally concordant and, by construction, yielded larger contrasts (e.g. RLC N2 $\rightarrow$ CER 1.25 [1.10–1.42]; RLC AllSleep $\rightarrow$ CER 1.22 [1.07–1.39]; RLC FirstHalf $\rightarrow$ PVD 1.24 [1.10–1.41]; RLC N1 $\rightarrow$ CVD 1.19 [1.06–1.34]). In I0006 (CER only), RLC FirstHalf $\rightarrow$ CER replicated in direction (per-SD OR 1.10 [1.01–1.20], $p = .030$; Q4-vs-Q1–Q3 1.28 [1.07–1.54], $p = .008$); other strata were non-significant but directionally concordant, consistent with the smaller sample and lower RLC scale at I0006.

RLC–limb independence held only partially in the validation cohorts. In I0004, where limb-event scoring is now comparable to S0001, conventional indices were themselves strong predictors of every endpoint (LMI $\rightarrow$ CHD 1.16, PVD 1.21, CER 1.23, CVD 1.20; all $p < .001$) and partially absorbed the RLC effects (RLC N1 $\rightarrow$ CVD 1.07 $\rightarrow$ 1.04 after LMI, $p = .14$; RLC AllSleep $\rightarrow$ CER 1.07 $\rightarrow$ 1.02 after LMI, $p = .59$); RLC N1 retained incremental value over PLMI/RRLMI/RRLMI N1 for the composite CVD endpoint (LRT $p \leq .04$, $\Delta AIC = -2.1$ to -3.2). In I0006, RLC FirstHalf $\rightarrow$ CER strengthened after LMI adjustment (per-SD 1.10 $\rightarrow$ 1.14, $p = .004$), while LMI $\rightarrow$ CER, non-significant unadjusted (0.93, $p = .10$), became modestly protective only with RLC FirstHalf in the model (0.88, $p = .011$).

Discussion

In this analysis, we present a method to assess the integration of respiratory and motor pathology during sleep, via measurement

of coactivation strength. Associations with vascular disease are presented as an initial use-case of this measurement.

Automating the capture of pathological coupling

The RLCI is a continuous, signal-based biomarker designed to capture the temporal synchronization between respiratory instability and limb activation. Our findings demonstrate that quantifying this coactivation provides significant information regarding cardiovascular status, particularly for cerebrovascular and PVDs. The primary methodological distinction of the RLCI is not automation versus manual scoring, but continuous signal-based quantification versus event-based quantification that depends on discrete boundaries and predefined temporal overlap rules. Traditional RRLM scoring is inherently fragile; it relies on the “perfect storm” of a scored respiratory event and a scored leg movement overlapping within a narrow time window. Even when such indices are derived computationally, small differences in event onset, offset, or duration can alter whether a movement is classified as respiratory-related. By shifting from discrete event counting to continuous signal integration, the RLCI reduces dependence on boundary placement and latency-window definitions, offering a scalable tool for population-level sleep analysis.

Comparison with standard indices

Our results indicate that the RLCI has health outcome associations comparable to, and in some contexts stronger than, established metrics like the PLMI and LMI. While the conventional whole-night RRLMI (derived from CAISR’s discrete event labels and normalized by total sleep time) was significantly associated with PVD and All CVD in the multivariable screening (Figure 5; Table S2), its associations with CHD and CER did not meet FDR control, underscoring the vulnerability of event-based coupling metrics to boundary/latency definitions and residual confounding. The stage-matched variants RRLMI N1 and RRLMI N2 partially recovered this signal, with RRLMI N1, in particular, reached FDR-significance for all four outcomes and remained robust in the Q4 vs. Q1–Q3 sensitivity analysis across all four, suggesting that the whole-night denominator dilutes a respiratory–motor signal that is concentrated in N1 sleep. The continuous RLCIs, specifically RLC FirstHalf and RLC N1, showed significant associations with CVD outcomes. This suggests that the intensity and continuity of the coupling, which the RLCI captures, may be more clinically relevant than a simple count of overlapping events. Notably, in the context of CER, the RLCI provided effect sizes marginally superior to the traditional PLMI, which may reflect amplified hemodynamic surges during this coactivation, although confirmation via direct beat-to-beat blood pressure monitoring is required.

These findings should not be interpreted as positioning the RLCI as a replacement for established limb-movement metrics such as the PLMI or LMI. Rather, the RLCI captures a distinct physiological construct: limb motor activity occurring in the specific context of respiratory recovery and associated autonomic activation. The persistence of RLCI associations, especially for CER, under the Q4 vs Q1–Q3 sensitivity analysis and mutual adjustment (Figures 6 and 7; Tables S3 and S4), with RLC N3 emerging as the most resilient stratum (significant across all four outcomes) and RLC FirstHalf retaining the strongest CER effect, supports the interpretation that respiratory-linked motor coactivation encodes risk beyond overall movement frequency, consistent with an autonomic/hemodynamic surge phenotype. Whereas the PLMI and LMI quantify overall motor burden during sleep, the RLCI emphasizes contextualized motor responses

embedded within periods of respiratory instability and heightened sympathetic drive. From this perspective, conventional limb-movement indices and the RLCI should be viewed as complementary rather than competing measures, indexing different aspects of sleep-related motor and autonomic dysregulation that may differentially contribute to cardiovascular risk. Importantly, these conclusions did not depend on a single analytic framing. The main RLCI associations persisted when conventional indices were recomputed within matched sleep stages and night segments, remained stable across broad perturbations of the signal-processing pipeline, and were directionally replicated in two independent HSP cohorts. Together, these findings support the interpretation that the observed cardiovascular associations reflect a stable respiratory–motor coupling signal rather than an artifact of whole-night asymmetry, parameter choice, or site-specific characteristics.

This conceptual distinction may also explain why the conventional whole-night RRLMI showed weaker and less consistent associations with cardiovascular outcomes in the present analyses—for CHD and CER, where it was outperformed by both standard limb-movement indices (PLMI, LMI) and selected RLC-derived metrics—though the attenuation was largely rescued when the RRLMI denominator was restricted to N1 sleep. Unlike the PLMI and LMI, which robustly capture overall motor burden, the RRLMI depends on the discrete temporal overlap between scored respiratory events and leg movements within a narrow latency window. As such, it remains vulnerable to small variations in event boundaries and scoring precision. In contrast, the continuous nature of the RLCI allows it to integrate the intensity and temporal persistence of respiratory–motor coupling, thereby reducing sensitivity to small timing differences in event definition while preserving the physiological context in which arousal-related limb activation occurs.

An additional consideration is the stage- and timing-specific pattern of associations observed for the RLCI. The strongest and most consistent relationships with cardiovascular outcomes emerged for the RLC measured during N1 sleep and during the first half of the night. This pattern must be interpreted in the context of the well-established nocturnal distribution of leg movements, which are themselves more frequent and pronounced in the early part of sleep [32, 33]. Importantly, however, the RLCI does not simply reflect increased motor activity, but rather the co-occurrence of limb activation with respiratory recovery. Early-night sleep and lighter non-REM stages are characterized by greater ventilatory instability, lower arousal thresholds, and more pronounced sympathetic responses to respiratory events [34, 35]. Under these conditions, respiratory recovery is more likely to coincide with phasic motor activation, producing amplified autonomic and cardiovascular surges. In line with prior Network Physiology work showing sleep-stage stratification of respiratory–motor–cardiac coupling [36–38] and its disruption under disease [39], our N1- and first-half-of-night findings are consistent with light sleep and early-night sleep being states in which respiratory–motor coupling is intrinsically more pronounced, more labile, and more sensitive to disease-related perturbation than deep or REM sleep. Within this framework, the temporal specificity of the RLCI reflects the convergence of heightened motor propensity and respiratory-autonomic instability, supporting its interpretation as a state-dependent coupling marker rather than a generic measure of motor burden.

A central piece of physiological information not directly captured here is the hemodynamic response. That is, whether high-RLCI episodes are accompanied by larger blood-pressure or

sympathetic surges than respiratory recovery or motor activation in isolation. Heart-rate variability provides only a partial proxy; a rigorous demonstration would require simultaneous beat-to-beat blood pressure recordings. If RRLMs are to be considered a therapeutic target, the RLCI may offer a quantitative readout to support clinical decision-making, given that available treatment options (e.g. intravenous iron) are limited and carry their own risks. The persistence of high coactivation after adequate sleep-apnea treatment may signal the need to target more subtle residual respiratory events, and assessing how treatment of coexisting restless legs syndrome modifies the RLCI is a natural next step.

Limitations

While the RLCI demonstrates promise, the current implementation relies on a set of physiologically motivated signal-processing parameters, including the stride and window lengths used to quantify respiratory recovery, the smoothing applied to the EMG envelope, and the percentile-based normalization and scaling rules. Although our sensitivity analyses showed that the main associations were robust across plausible perturbations of these parameters, the current configuration was not derived through formal mathematical optimization. While the RRLMI captures the frequency of discrete RRLM events, it is inherently sensitive to small discrepancies in the onsets and offsets of both respiratory events and EMG bursts: shifts of 1–2 s can move an event in or out of the 0–5 s latency window. This dependence on precise event boundaries may limit the robustness and reproducibility of the RRLMI across scoring algorithms and datasets.

A second consideration concerns the use of CAISR labels as inputs to the RLCI pipeline. In the S0001 metadata, 19 604 unique patients contributed 26 328 PSG sessions. Of these, 14 941 patients (76%; 21 004 sessions) appeared in the CAISR auto-staging training table, whereas 4663 patients (24%; 5324 sessions) did not. CAISR labels were therefore not generated under fully out-of-sample conditions for much of the S0001 discovery cohort. Because CAISR labels are used as input features—including sleep stage, respiratory events, arousals, and limb movements—rather than as cardiovascular prediction targets, and because the cardiovascular outcomes were derived independently from ICD-9 records, the principal concern is inflated label quality rather than direct outcome leakage. To assess this possibility, we repeated the primary multivariable analyses after applying the same analytic filters to sessions from patients absent from the CAISR training table. This non-training analytic subset comprised 3764 complete-label, QC-passing S0001 sessions. As detailed in Table S7, the RLCI-cerebrovascular-disease signal replicated across all five strata, with per-1-SD log-transformed adjusted ORs of 1.11–1.13, closely matching the primary cohort. Across the full 124-cell screen, standard errors increased as expected from the reduced sample size, and 80 of 84 primary BH-FDR-significant cells retained their direction. Thus, the smaller number of FDR-significant associations in the non-training subset is consistent with reduced power rather than systematic effect attenuation.

The inter-center analyses in I0004 and I0006, neither of which contributed to CAISR training, and the held-out S0001 subset analysis described in the Methods together mitigate this concern, but readers should interpret single-site label-quality metrics with this overlap in mind.

We also considered a simpler Hilbert-envelope formulation for the respiratory signal. Visual comparison (Figure S2) suggested that the Hilbert airflow envelope was less specific to post-event recovery than the recovery-score formulation, because it also captured normal breathing amplitude fluctuations and low-flow

periods within respiratory events. A head-to-head quantitative comparison on the same $N = 4846$ subset used for the signal-processing hyperparameter sensitivity analysis (after QC, BMI, race, and age filters), under the identical adjusted logistic regression, supported this choice: the two recovery-targeted formulations were strongly correlated on the log scale ($r \approx 0.66$ for AllSleep) and produced broadly similar adjusted ORs, while the during-apnea Hilbert variant was essentially orthogonal to both and to the AHI / ArI / LMI ($r \leq 0.10$), describing a mechanistically distinct phenomenon that we did not pursue (Table S8).

Future work should involve larger-scale ablation and optimization studies to determine the sensitivity of the index to these hyperparameters and to identify the optimal configuration for maximizing clinical correlation.

The current study is cross-sectional, associating the RLCI with prevalent CVD diagnoses. While this establishes a strong concurrent link, it cannot definitively prove causality. The longitudinal prognostic utility of the RLCI needs to be validated in longitudinal cohorts, testing its ability to predict the future onset of cardiovascular events in patients who are currently disease-free.

The comparatively weaker association between RLCI metrics and CHD, relative to cerebrovascular outcomes, warrants comment. Coronary disease risk may be more strongly driven by cumulative metabolic and inflammatory burden, which is well captured by overall limb-movement frequency and sleep fragmentation. In contrast, CER may be particularly sensitive to transient surges in blood pressure, heart rate, and sympathetic tone occurring at arousal and respiratory recovery. The stronger relationship between the RLCI and cerebrovascular outcomes is therefore consistent with a model in which acute autonomic stressors, rather than global motor burden alone, play a disproportionate role in cerebrovascular vulnerability.

Conclusions

The RLCI provides a novel, algorithmic framework for quantifying the coupling between respiratory events/instability and limb motor activation during sleep. Our findings suggest that this coupling, particularly when occurring during early-night sleep and lighter non-REM stages, is meaningfully associated with CVD, with the strongest signals observed for cerebrovascular outcomes. These associations persisted under matched comparisons with conventional indices, remained stable across signal-processing perturbations, and were directionally replicated across independent HSP cohorts. Future studies should extend this work in several directions, including longitudinal validation to establish prognostic utility, integration with direct measures of autonomic and hemodynamic reactivity, and evaluation of RLCI dynamics in response to therapeutic interventions such as positive airway pressure or pharmacologic modulation of arousal and motor activity. A particular near-term opportunity is to use the RLCI as a treatment-response readout: the persistence of high coactivation after adequate AHI control would identify patients in whom respiratory-motor coupling, rather than apnea burden alone, sustains residual cardiovascular risk, and who may benefit from adjunctive intervention. More broadly, the RLCI illustrates how continuous, multimodal signal integration can enhance sleep phenotyping and may help identify high-risk physiological patterns that remain invisible to conventional, event-based scoring approaches.

Supplementary material

Supplementary material is available at SLEEP online.

Funding

M.B.W.'s laboratory is supported by grants from the NIH (R01AG073410, R01HL161253, R01NS126282, R01AG073598, R01NS131347, and R01NS130119).

Disclosure statement

Financial disclosure: R.J.T. is a co-inventor of a cardiopulmonary sleep spectrogram that is licensed by the Beth Israel Deaconess Medical Center to MyCardio, LLC. He also engages in general sleep medicine consulting for GLG Councils, Guidepoint, Beacon Biosignals, and Jazz Pharmaceuticals. M.B.W. has a personal equity interest in Beacon Biosignals.

Non-financial disclosure: R.J.T. is a co-inventor of a cardiopulmonary sleep spectrogram used to assess sleep stability/quality and sleep apnea. He also holds a patent for an Enhanced Expiratory Rebreathing Space to treat high-loop-gain sleep apnea, and a patent for estimating respiratory self-similarity for the detection of high-loop-gain sleep apnea. M.B.W. is a co-founder and serves as a scientific advisor and consultant to Beacon Biosignals.

Data availability

The HSP dataset is hosted on the Brain Data Science Platform (BDSP), a secure, NIH-supported repository facilitating the sharing of de-identified, multi-center sleep and brain-related clinical data. BDSP provides standardized access to HSP data in BIDS and HDF5 formats, supporting reproducible analyses and machine learning workflows. HSP is distributed through BDSP under a credentialed-access model. Approved users can download the data locally for offline analysis. Access requires: (1) registration of an account on BDSP, (2) submission of proof of current HIPAA training, and (3) signing a brief Data Use Agreement in which the user commits to use the data for research purposes only, not to attempt re-identification of any participant, and not to redistribute the data to third parties. The Data Use Agreement follows the same framework adopted for other clinical datasets hosted on the BDSP. Once these requirements are met, access is typically granted within a few business days, and users can download the standardized HDF5 files, per-cohort metadata CSVs, and accompanying documentation directly from the platform. For more information, visit bdsp.io.

The code used for the RLC framework, including all data preprocessing, analysis pipelines, and figure generation, is available on https://github.com/NAIL-NeurologyAILab/RLC_CrossSectional.

References

- Ivanov PC, Bartsch RP. Future of sleep medicine. *Sleep Med Clin*. 2025;**20**(1):149–164. <https://doi.org/10.1016/j.jsmc.2024.10.013>
- Manconi M, Zavalko I, Bassetti CL, Colamartino E, Pons M, Ferri R. Respiratory-related leg movements and their relationship with periodic leg movements during sleep. *Sleep*. 2014;**37**(3):497–504. <https://doi.org/10.5665/sleep.3484>
- Mogavero MP, Lanza G, Bruni O, et al. Periodic limb movements during sleep and progression of sleep apnea: a longitudinal analysis from the Wisconsin sleep cohort. *Sleep Breath*. 2025;**29**(6):314. <https://doi.org/10.1007/s11325-025-03484-z>
- Krzyzaniak K, Partinen E, Partinen M, Sieminski M. Patients with a higher number of periodic limb movements have higher nocturnal blood pressure. *J Clin Med*. 2022;**11**(10):2829. <https://doi.org/10.3390/jcm11102829>
- Sieminski M, Pyrzowski J, Partinen M. Periodic limb movements in sleep are followed by increases in EEG activity, blood pressure, and heart rate during sleep. *Sleep Breath*. 2017;**21**(2):497–503. <https://doi.org/10.1007/s11325-017-1476-7>
- Siddiqui F, Strus J, Ming X, Lee IA, Chokroverty S, Walters AS. Rise of blood pressure with periodic limb movements in sleep and wakefulness. *Clin Neurophysiol*. 2007;**118**(9):1923–1930. <https://doi.org/10.1016/j.clinph.2007.05.006>
- Pennestri MH, Montplaisir J, Fradette L, Lavigne G, Colombo R, Lanfranchi PA. Blood pressure changes associated with periodic leg movements during sleep in healthy subjects. *Sleep Med*. 2013;**14**(6):555–561. <https://doi.org/10.1016/j.sleep.2013.02.005>
- Yang CK, Jordan AS, White DP, Winkelman JW. Heart rate response to respiratory events with or without leg movements. *Sleep*. 2006;**29**(4):553–556. <https://doi.org/10.1093/sleep/29.4.553>
- Malkiewicz MA, Grzywinska M, Malinowski KS, et al. Time-dependent autonomic dysregulation and co-activation induced by periodic limb movements in sleep. *J Clin Med*. 2025;**14**(6):1940. <https://doi.org/10.3390/jcm14061940>
- Aritake S, Blackwell T, Peters KW, et al. Prevalence and associations of respiratory-related leg movements: the MrOS sleep study. *Sleep Med*. 2015;**16**(10):1236–1244. <https://doi.org/10.1016/j.sleep.2015.06.012>
- Esmaili N, Aritake S, Redline S, Winkelman JW, Azarbarzin A. Respiratory-related leg movement as a predictor of mortality: the osteoporotic fractures in men sleep study. *SLEEP*. 2026;**49**(4):zsaf310. <https://doi.org/10.1093/sleep/zsaf310>
- Ferri R, Fulda S, Allen RP, et al. World Association of Sleep Medicine (WASM) 2016 standards for recording and scoring leg movements in polysomnograms developed by a joint task force from the International and the European Restless Legs Syndrome Study Groups (IRLSSG and EURLSSG). *Sleep Med*. 2016;**26**:86–95. <https://doi.org/10.1016/j.sleep.2016.10.010>
- Huang Z, Lobbezoo F, de Vries N, Aarab G, Hilgevoord AAJ. Temporal association between leg movements and respiratory events in patients with obstructive sleep apnea: description and differences between the AASM and WASM scoring criteria. *Sleep Breath*. 2023;**27**(6):2295–2304. <https://doi.org/https://doi.org/10.1007/s11325-023-02844-x>
- Westover MB, Moura Junior V, Thomas R, et al. The Human Sleep Project. Brain Data Science Platform. 2023. <https://doi.org/10.60508/QJBV-HG78>
- Gramfort A, Luessi M, Larson E, et al. MNE software for processing MEG and EEG data. *Neuroimage*. 2014;**86**:446–460. <https://doi.org/10.1016/j.neuroimage.2013.10.027>
- Larson E, Gramfort A, Engemann DA, et al. MNE-python. Published online June. 2024;**15**. <https://doi.org/10.5281/ZENODO.592483>
- Adámek K, Novotný J, Armour W. A polyphase filter for many-core architectures. *Astron Comput*. 2016;**16**:1–16. <https://doi.org/10.1016/j.ascom.2016.03.003>
- Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat Methods*. 2020;**17**(3):261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Nasiri S, Ganglberger W, Nassi T, et al. CAISR: achieving human-level performance in automated sleep analysis across all clinical sleep metrics. *SLEEP*. 2025;**48**(8):zsaf134. <https://doi.org/10.1093/sleep/zsaf134>
- Einizade A, Nasiri S, Sardouie SH, Clifford GD. ProductGraph-SleepNet: sleep staging using product spatio-temporal graph learning with attentive temporal aggregation. *Neural Netw*. 2023;**164**:667–680. <https://doi.org/10.1016/j.neunet.2023.05.016>

21. Perslev M, Darkner S, Kempfner L, Nikolic M, Jennum PJ, Igel C. U-sleep: resilient high-frequency sleep staging. *Npj Digit Med*. 2021;**4**(1):72. <https://doi.org/10.1038/s41746-021-00440-5>
22. Ferri R, Zucconi M, Manconi M, et al. Computer-assisted detection of nocturnal leg motor activity in patients with restless legs syndrome and periodic leg movements during sleep. *Sleep*. 2005; **28**(8):998–1004. <https://doi.org/10.1093/sleep/28.8.998>
23. Mietus JE, Peng CK, Ivanov PC, Goldberger AL. Detection of obstructive sleep apnea from cardiac interbeat interval time series. In: *Computers in Cardiology 2000*. IEEE; 2000;**27**:753–756. <https://doi.org/10.1109/CIC.2000.898634>
24. Ivanov PC, Rosenblum MG, Peng CK, et al. Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis. *Nature*. 1996;**383**(6598):323–327. <https://doi.org/10.1038/383323a0>
25. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python. *J Mach Learn Res*. 2011;**12**(85): 2825–2830. <http://jmlr.org/papers/v12/pedregosa11a.html>
26. Centers for Disease Control and Prevention. International Classification of Diseases, ninth revision, clinical modification (ICD-9-CM). *Ctr Dis Control Prev*. 2013;
27. Box GEP, Cox DR. An analysis of transformations. *J R Stat Soc Ser B Methodol*. 1964;**26**(2):211–243. <https://doi.org/10.1111/j.2517-6161.1964.tb00553.x>
28. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;**57**(1):289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
29. Hosmer DW Jr, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. 3rd ed. Hoboken, NJ: John Wiley & Sons; 2013. <https://doi.org/10.1002/9781118548387>
30. R Core Team. R: A language and environment for statistical computing. In: *R Foundation for Statistical Computing*. 2024: <https://www.R-project.org/>.
31. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *J Open Source Softw*. 2019;**4**(43):1686. <https://doi.org/10.21105/joss.01686>
32. Ferri R, Koo BB, Picchietti DL, Fulda S. Periodic leg movements during sleep: phenotype, neurophysiology, and clinical significance. *Sleep Med*. 2017;**31**:29–38. <https://doi.org/10.1016/j.sleep.2016.05.014>
33. Ferri R. The time structure of leg movement activity during sleep: the theory behind the practice. *Sleep Med*. 2012;**13**(4): 433–441. <https://doi.org/10.1016/j.sleep.2011.10.027>
34. Landry SA, Andara C, Terrill PI, et al. Ventilatory control sensitivity in patients with obstructive sleep apnea is sleep stage dependent. *Sleep*. 2018;**41**(5):zsy040. <https://doi.org/10.1093/sleep/zsy040>
35. Han H, Seong MJ, Hyeon J, Joo E, Oh J. Classification and automatic scoring of arousal intensity during sleep stages using machine learning. *Sci Rep*. 2024;**14**(1):5983. <https://doi.org/10.1038/s41598-023-50653-9>
36. Ivanov PC. The new field of network physiology: building the human physiome. *Front Netw Physiol*. 2021;**1**:711778. <https://doi.org/10.3389/fnetp.2021.711778>
37. Bartsch RP, Liu KKL, Bashan A, Ivanov PC. Network physiology: how organ systems dynamically interact. *PLoS One*. 2015;**10**(11):e0142143. <https://doi.org/10.1371/journal.pone.0142143>
38. Liu KKL, Bartsch RP, Lin A, Mantegna RN, Ivanov PC. Plasticity of brain wave network interactions and evolution across physiologic states. *Front Neural Circuits*. 2015;**9**:62. <https://doi.org/10.3389/fncir.2015.00062>
39. Rizzo R, Wang JWJL, DePold HA, Holsapple JW, Vaou OE, Ivanov PC. Dynamic networks of cortico-muscular interactions in sleep and neurodegenerative disorders. *Front Netw Physiol*. 2023;**3**:1168677. <https://doi.org/10.3389/fnetp.2023.1168677>