

Quantitative validation of the suprasternal pressure signal to assess respiratory effort during sleep

Citation for published version (APA):

Cerina, L., Papini, G. B., Fonseca, P., Overeem, S., van Dijk, J. P., van Meulen, F., & Vullings, R. (2024). Quantitative validation of the suprasternal pressure signal to assess respiratory effort during sleep. *Physiological Measurement*, 45(5), Article 055020. <https://doi.org/10.1088/1361-6579/ad4c35>

Document license:

CC BY

DOI:

[10.1088/1361-6579/ad4c35](https://doi.org/10.1088/1361-6579/ad4c35)

Document status and date:

Published: 01/05/2024

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

PAPER • OPEN ACCESS

Quantitative validation of the suprasternal pressure signal to assess respiratory effort during sleep

To cite this article: Luca Cerina *et al* 2024 *Physiol. Meas.* **45** 055020

View the [article online](#) for updates and enhancements.

You may also like

- [Chaplygin Gas Inspired Warm Inflation and Swampland Conjecture Through Various Scalar Potentials](#)
abdul jawad, Shama Sadiq, Nadeem Azhar et al.
- [Development of a novel, concentric micro-ECOG array enabling simultaneous detection of a single location by multiple electrode sizes](#)
Ian R. Akamine, Jonathan V Garich, Daniel W. Gulick et al.
- [Computed chest radiography for total body irradiation: image quality and clinical feasibility](#)
Quentin Bouchez, Dirk Adelin Vandembroucke, Geert Pittomvils et al.

Breath Biopsy Conference

BREATH
BIOPSY

Join the conference to explore the **latest challenges** and advances in **breath research**, you could even **present your latest work!**



5th & 6th November
Online



Main talks



Early career sessions



Posters

Register now for free!



PAPER

OPEN ACCESS




RECEIVED
6 November 2023REVISED
6 May 2024ACCEPTED FOR PUBLICATION
15 May 2024PUBLISHED
29 May 2024

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Quantitative validation of the suprasternal pressure signal to assess respiratory effort during sleep

Luca Cerina^{1,*} , Gabriele B Papini^{1,2} , Pedro Fonseca^{1,2} , Sebastiaan Overeem^{1,3}, Johannes P van Dijk^{1,3}, Fokke van Meulen^{1,3} and Rik Vullings¹

¹ Electrical Engineering, Technische Universiteit Eindhoven, Eindhoven, Noord Brabant, The Netherlands

² Philips Research, Eindhoven, Noord Brabant, The Netherlands

³ Center for Sleep Medicine, Kempenhaeghe Foundation, Heeze, Noord Brabant, The Netherlands

* Author to whom any correspondence should be addressed.

E-mail: l.cerina@tue.nl, gabriele.papini@philips.com, pedro.fonseca@philips.com, overeems@kempenhaeghe.nl, dijkh@kempenhaeghe.nl, meulenf@kempenhaeghe.nl and r.vullings@tue.nl

Keywords: sleep disordered breathing, polysomnography, suprasternal notch pressure, esophageal pressure, respiratory effort

Supplementary material for this article is available [online](#)

Abstract

Objective. Intra-esophageal pressure (Pes) measurement is the recommended gold standard to quantify respiratory effort during sleep, but used to limited extent in clinical practice due to multiple practical drawbacks. Respiratory inductance plethysmography belts (RIP) in conjunction with oronasal airflow are the accepted substitute in polysomnographic systems (PSG) thanks to a better usability, although they are partial views on tidal volume and flow rather than true respiratory effort and are often used without calibration. In their place, the pressure variations measured non-invasively at the suprasternal notch (SSP) may provide a better measure of effort. However, this type of sensor has been validated only for respiratory events in the context of obstructive sleep apnea syndrome (OSA). We aim to provide an extensive verification of the suprasternal pressure signal against RIP belts and Pes, covering both normal breathing and respiratory events. **Approach.** We simultaneously acquired suprasternal (207) and esophageal pressure (20) signals along with RIP belts during a clinical PSG of 207 participants. In each signal, we detected breaths with a custom algorithm, and evaluated the SSP in terms of detection quality, breathing rate estimation, and similarity of breathing patterns against RIP and Pes. Additionally, we examined how the SSP signal may diverge from RIP and Pes in presence of respiratory events scored by a sleep technician. **Main results.** The SSP signal proved to be a reliable substitute for both esophageal pressure (Pes) and respiratory inductance plethysmography (RIP) in terms of breath detection, with sensitivity and positive predictive value exceeding 75%, and low error in breathing rate estimation. The SSP was also consistent with Pes (correlation of 0.72, similarity 80.8%) in patterns of increasing pressure amplitude that are common in OSA. **Significance.** This work provides a quantitative analysis of suprasternal pressure sensors for respiratory effort measurements.

1. Introduction

Respiration is the physiological pillar of human life: a remarkably efficient process emerging from the complex interaction of the nervous system, smooth muscles, lungs, blood vessels, and multiple mechanical and chemical receptors. The same complexity implies that different physiological, pathological, or external factors may disturb this delicate equilibrium and trigger abnormal respiratory patterns or apneic events in predisposed individuals. Sleep can be one of these factors as it naturally alters breathing activity in terms of respiratory rate, ventilation, upper airway resistance, and neural control of pharyngeal and rib-cage muscles (Chokroverty 2017). The combination of these physiological alterations and common neural and anatomical risk factors can induce sleep-disordered breathing (SDB) pathologies. These pathologies are highly prevalent

in the general population, with obstructive sleep apnea (OSA) as the most common manifestation (Dempsey *et al* 2010, Heinzer *et al* 2015). In general, it would not be feasible to measure in detail all the components of respiration (from neural to muscular activity and gas exchanges) without altering physiological sleep patterns. Therefore, finding the simplest model to measure sleep and respiration is important in the clinical diagnostic processes for assessing suspected SDB and sleep disorders in general. The challenge is to obtain enough data to reach the proper clinical decisions with the most natural sleeping conditions possible. The current gold standard is a laboratory polysomnography (PSG) system which includes: respiratory effort, oronasal airflow, and pulse oximetry to monitor respiration, brain, cardiac, and eye activity to assess sleep stages and arousals, plus other sensors for muscular activity and body movements. Other systems, such as home sleep apnea tests (HSATs), can simplify multiple nights of recordings by measuring sleep at home and reduce the psychological lab effect, but compromise the number of signals available, depending on the HSAT class (Ferber *et al* 1994, Mann *et al* 2020).

In this scenario of SDB, measuring changes in the respiratory effort is one of the most valuable descriptors in the pathophysiology of OSA and respiratory effort-related arousals (RERA), but also less explored phenomena such as prolonged partial obstruction syndrome (Lévy *et al* 2015, Anttalainen *et al* 2016). According to American Association of Sleep Medicine (AASM) recommendations, the closest approximation of effort is the measure of intrathoracic pressure swings with an esophageal catheter (Berry *et al* 2012). Yet, clinicians seldomly use the esophageal pressure (Pes) sensor due to complicated placement, limited tolerance by patients, and potential effects on pharyngeal dynamics (Brochard 2014, Glos *et al* 2018). Consequently, respiratory inductance plethysmography (RIP) belts are the accepted alternative, with various combinations of abdominal and thoracic movements and their asynchrony as surrogates of respiratory effort. They are frequently referred to as effort belts and signals (Hammer and Newth 2009), but these may be imperfect definitions. Earlier works on RIP and subsequent technical recommendations from AASM explicate how the altered respiratory effort is evident only during obstruction, with a consequent paradoxical motion of the abdomen and rib cage (Staats *et al* 1984, Flemons *et al* 1999, Redline *et al* 2007). In all other cases, RIP belts measure only chest wall motion and relative changes in tidal volume, which indicate, and do not quantify, the physical effort exerted by the act of respiration (particularly if not calibrated, Zimmerman *et al* 1983).

A promising alternative measure is suprasternal notch pressure (SSP): a pressure sensor or a microphone encased in an airtight capsule, which sits on the suprasternal fossa. Respiration contracts the skin, shifting pressure in the capsule, so the measured signals cover either low-frequency pressure swings (0.2–20 Hz: Glos *et al* 2018, Sabil *et al* 2019) or tracheal sounds in the audible range (200–2000 Hz: Amaddeo *et al* 2020, Devani *et al* 2021), depending on the sensor's manufacturer. In this way, SSP sensors might have the usability and comfort of RIP belts while measuring respiratory effort in a modality closer to esophageal sensors. Both suprasternal pressure and tracheal sounds are reliable measures of respiratory effort in the presence of apneas and hypopneas, but some discrepancies with RIP signals exist (Sabil *et al* 2019). Additionally, the current scope of research about SSP remains limited to complementing manual scoring of PSG, or in the case of automatic scoring, to the classification of broad OSA severity classes. What is missing is a precise characterization of the typologies, timing, and duration of respiratory events and the analysis of normal breathing.

We propose here a quantitative validation of the SSP, specifically the pressure signal, over the entire night of sleep with the inclusion of regular breathing, and compare it with synchronized Pes and RIP signals. The study's primary aim is to understand if the SSP is a valid substitute for respiratory belts and esophageal catheters in a broad set of breath-by-breath analyses. The secondary aim is to measure the similarity of SSP to Pes signal in patterns of multiple subsequent breaths. Lastly, we will highlight similarities and differences between sensors from a qualitative point of view and technical challenges associated with sleep-related experimental settings.

2. Methods

2.1. Dataset

To validate the SSP sensor, we employed data from Sleep Apnoea Monitoring with Non-Invasive Applications (SOMNIA) project (van Gilst *et al* 2019), a clinical database designed to facilitate research on unobtrusive monitoring of sleep and sleep disorders. Clinicians at the Center for Sleep Medicine Kempenhaeghe (Heeze, the Netherlands) collected all the recordings, where participants underwent PSG for suspicion of sleep disorders (including, but not only, OSA) as a third-line referral. Specifically, we started from a subset of 240 single-night PSG recordings collected between 2017 and 2022 where the SSP sensor was available along standard PSG sensors. The Pes sensor was available in 27 recordings. Following the hospital procedures, intra-esophageal pressure measurement was employed in specific cases (e.g. suspicion of central disorders,

Table 1. Demographics of the participants.

Variable		mean \pm std or median, IQR	(range)
Whole dataset (207)			
Sex	[#]	80 female / 127 male	
Age [†]	[years]	44 \pm 19	(3 : 94)
Body-mass index (BMI)	[kg m ⁻²]	26.1 \pm 5.1	(12.9 : 45.1)
Total sleep time (TST)	[minutes]	408.1 \pm 74	(167 : 607)
Apnea-hypopnea index (AHI) ⁺	[events/hour]	10.7, 19.9	(0 : 107.4)
- Apnea Index	[events/hour]	0.3, 1.3	(0 : 85.2)
- Hypopnea Index	[events/hour]	10.1, 15.9	(0 : 88.8)
Pes subset (20)			
Sex	[#]	5 female / 15 male	
Age	[years]	48 \pm 13	(26 : 69)
Body-mass index (BMI)	[kg m ⁻²]	28.1 \pm 4.2	(18.6 : 45.1)
Total sleep time (TST)	[minutes]	370 \pm 72	(219 : 489)
Apnea-hypopnea index (AHI) ⁺	[events/hour]	12.7, 12.9	(0.7 : 99.6)
- Apnea Index	[events/hour]	0.6, 2.4	(0 : 85.1)
- Hypopnea Index	[events/hour]	12.9, 9.6	(0.7 : 31.4)

[†] 23 participants younger than 18, mean 11 years.

⁺ Expressed as median and IQR due to skewed distribution.

clinical signs of OSA but a negative previous sleep study, etc). Each recording is associated with a unique participant.

After investigating the signal quality, we opted to discard some recordings under these criteria: Complete failure of the SSP sensor (17), Pes (1), or RIP (2), no overlap in the signals (2, e.g. split recording with different sensors), unrecoverable noise in RIP e.g. due to excessive belt tension (7) and excessive noise in the SSP signal, e.g. due to air leaks in the sensor's chamber (4). The final subset contains 207 recordings with SSP signal, and 20 include the Pes signal. Demographics of participants are shown in table 1 together with total sleep time (TST) and apnea-hypopnea index (AHI).

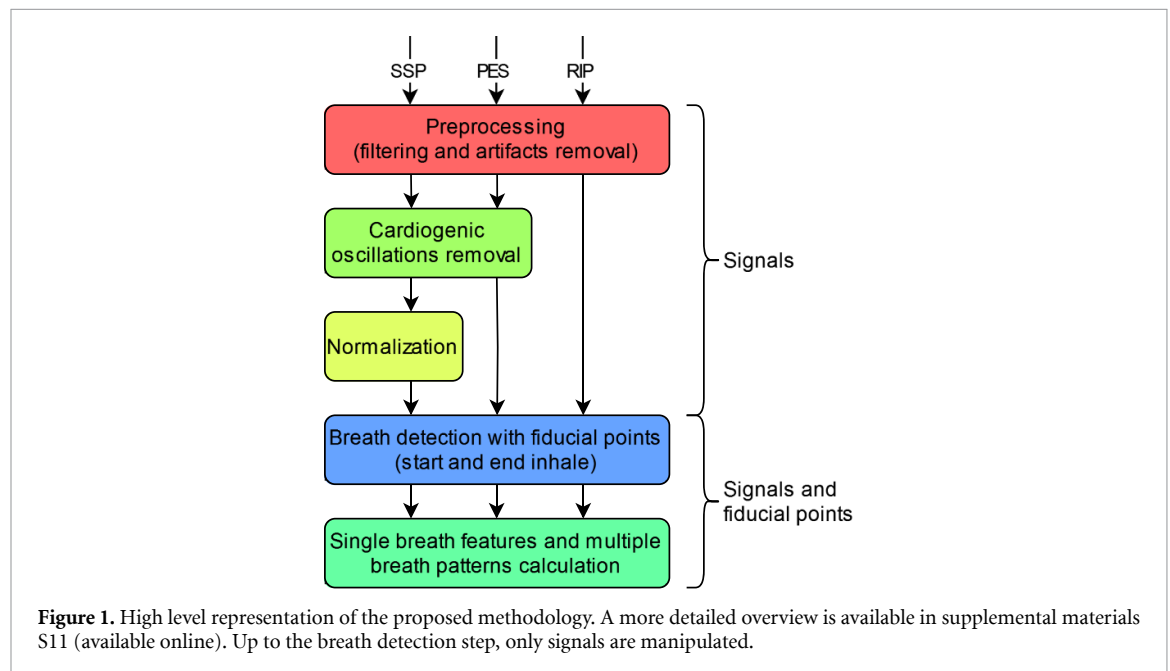
A trained sleep technician scored every recording according to the AASM rules v2.5 (Berry *et al* 2012), with a 3% oxygen drop threshold for hypopneas. Then a clinician classified potential sleep disorder according to International Classification of Sleep Disorders (ICSD-3) rules. SDB was the most common diagnosis with 52% of participants (108), followed by insomnia (35.2%, 73), sleep-related movement disorders (14.4%, 30) and parasomnias (12.5%, 26). As a third-line referral center, only a few participants did not present a sleep disorder (3.8%, 8). All participants with diagnosed SDB had one or more co-morbidities. The most common co-morbidity was insomnia in 15.6% of SDB participants. The percentage was lower than general prevalence of insomnia symptoms in literature (30%–50%; Sweetman *et al* 2019). Other sleep disorders were present as well, but in less than 5% of the participants. In the Pes subset, 80% of the participants (16) received a diagnosis of SDB.

The esophageal pressure was recorded according to AASM technical specifications with a sampling frequency of 128 Hz and a filter DC-15 Hz. The SSP signal was recorded with a sampling frequency f_s of 1024 Hz with a filter DC-285 Hz. The sensor's front-end automatic gain control and filters were disabled to avoid unwanted alterations in the signal's dynamic range. The SSP signal was re-sampled at 128 Hz with a polyphase FIR resampling filter to align it with the sampling frequency of RIP belts and Pes signals.

Table 2 includes the technical details of the PSG sensors employed in this project. Sleep was monitored with a Graef PSG system (Compumedics, USA). All sensors' channels are recorded with 24bit resolution through low noise, high impedance uni-polar inputs with ± 300 mV range, data save rate up to 2048 sps and linear frequency response up to 580 Hz. The position sensor does not output the tri-axial accelerations, but a pre-computed value corresponding to upright, supine, left, right, or prone position. The SSP is an in-house adaptation of a commercial sensor (Sensym SDX010IND4, Honeywell, North Carolina) placed on the suprasternal notch with a fixation adhesive (Provox Optiderm Round 7255, Atos Medical, Sweden) and encased in a custom-made 3D-printed adapter. Exact build specifications are available upon request. Given the clinical setting RIP belts's output depends on how they are fitted to the participant and are uncalibrated. In the same way, both the Pes and SSP sensors are not calibrated in this dataset (for example, assuming the end of exhalation to be 0 mmHg). The details of the other PSG sensors employed in the clinical setting, but outside the focus of this work, are available in the original SOMNIA paper (van Gilst *et al* 2019).

Table 2. Technical specification of the PSG sensors.

Manufacturer	Sensor	Sampling rate	Filter bands	Accuracy
Electrocardiography (ECG)				
Kendall, Ireland	Electrode H34SG	512 Hz	0.3–70 Hz	35 μ V (from PSG)
Respiratory inductance plethysmography (RIP)				
Compumedics, Australia	Dual RIP bands	128 Hz	0.1–15 Hz	NA/uncalibrated
Esophageal pressure (Pes)				
Gaeltec, Ireland	Gaeltec CTO-1	128 Hz	0.1–15 Hz	0.07 mmHg
Suprasternal notch pressure (SSP)				
Honeywell, North Carolina	Sensym SDX010IND4	1024 Hz	0–285 Hz	0.03 mmHg
Body position				
Sleepsense, Illinois	Grael position sensor	16 Hz	NA	upright, supine, left, right, prone



2.2. Methodology overview

Figure 1 provides an overview on how our respiratory signals of interest are processed before comparing their characteristics. All signals present similar noise sources and baseline wanders that are corrected or minimized in the pre-processing step (section 2.3). After that, the RIP belts' signal is good enough, while SSP and Pes signals require further processing. Specifically, both contains cardiogenic oscillations caused by the vibration of pulmonary arteries that we removed using the method proposed in Cerina *et al* (2023) with the ECG signal as the source of the heart rate (section 2.4). The SSP signal may suffer from drastic amplitude changes, therefore we added another normalization step before the breath detection (section 2.5.1). Finally, we detect fiducial points of breaths in all the signal with a two-step approach: first the end of inhalation, section 2.5, then the start of inhalation, section 2.5.2. The fiducial points are used in combination with the signals to extract breath-by-breath features (amplitude and area) and multiple breath patterns that are then compared across signals (section 2.6). Supplemental materials S11 include a more extensive overview of the algorithms to help with the replication of the proposed methodology, including relevant libraries, model formulations, computing requirements and algorithm complexity.

2.3. Pre-processing

We observed a variety of noise sources present in all our signals of interest, specifically baseline wanders, sensor disconnections, signal clipping, and large amplitude artefacts (probably caused by body motions).

These similarities between sensors allowed us to perform unified pre-processing for all the signals, with few additional sensor-specific steps, which we will present separately. First, we detected large amplitude spikes that may be present near the beginning or end of the recording due to clinical protocol (e.g. movement artefacts before going to sleep, here called borders). To do so, we used the following formula:

$$c_T = \sum_{t=0}^T \begin{cases} 1 & \text{if } |x(t)| > \text{threshold} \\ 0 & \text{if } |x(t)| \leq \text{threshold} \end{cases} \quad (1)$$

where c_T is the cumulative number of samples in the signal $x(t)$ (with t being time) up to time T exceeding a sensor-dependent fixed border's threshold (98% of the sensor's maximum range). If there are no spikes, the cumulative sum will always equal 0. If spikes are present, the cumulative sum will have one or two step increases (depending if the problem is at the beginning of the signal, end or both). We define the borders as the position of these steps and fix the value of the respiratory signals outside the borders using the edge values inside: the first signal's sample after a beginning border and the last sample before an ending border.

Then, we detected clipping samples in the signal of the using a sensor-dependent fixed *clipping* threshold (95% of the sensor's maximum range to account for partially clipped breaths) and set them to 0. We also detected sensor disconnections as all the samples with an absolute amplitude lower than a *disconnection* threshold of 0.0001 (to be robust to fluctuations around 0). The threshold is sensor-independent and unit-less, and we selected it to be 10 times below the average amplitude of the signal during central apneas in all sensors. We consider the signal not present or the sensor disconnected if a block of at least 1 second is smaller than the *disconnection* threshold. Each recording also contains a manual annotation of the *lights-on* periods, or any moment during the night that is ignored in the scoring phase because the participant was not sleeping, or out of bed, or if a sleep technician or nurse was required. We considered those phases equivalent to a sensor disconnection and discarded them in subsequent analyses. Lastly, we removed baseline wanders and high-frequency noise using a 4th order Butterworth band-pass filter with the cutoff frequencies recommended by AASM, i.e. between 0.1 Hz and 15 Hz.

2.4. Removal of cardiogenic oscillations

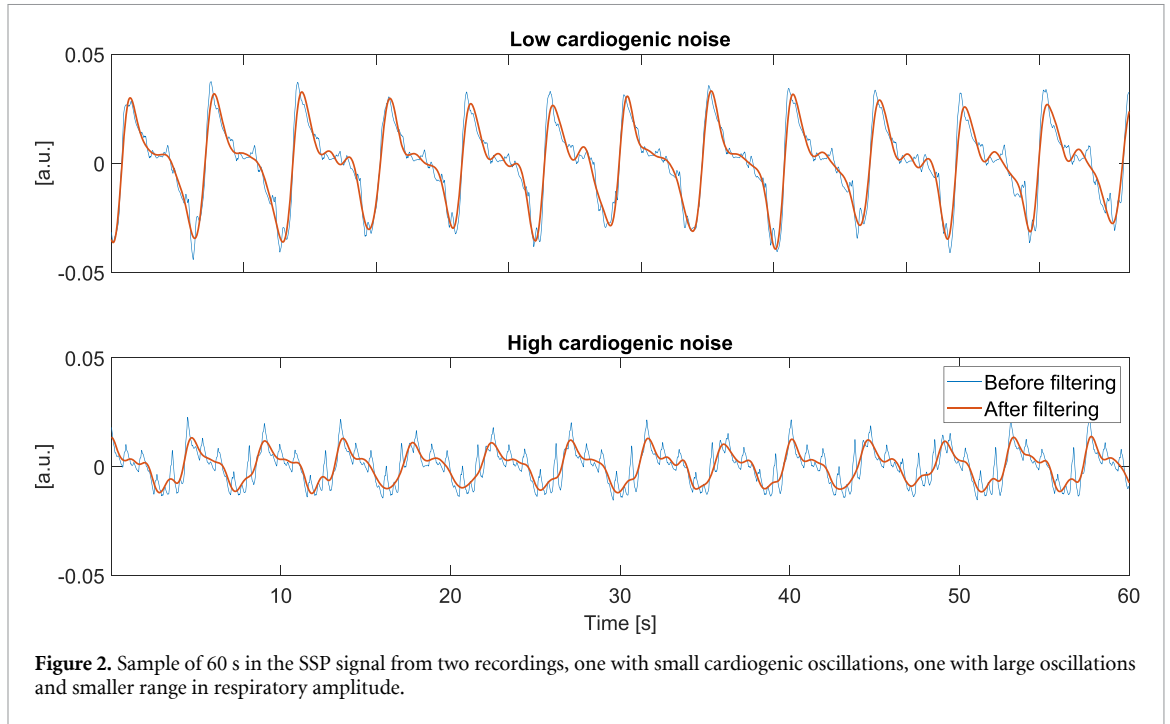
Another source of noise in SSP and Pes is cardiogenic oscillations: the vibration of the pulmonary arteries that causes a measurable pressure shift in the airways (Suarez-Sipmann *et al* 2012). These oscillations are generally negligible in RIP or oronasal airflow signals (except in the absence of respiratory drive, i.e. in central apneas). Still, they can drastically alter the morphology of breaths in pressure signals, affecting breath detection. Filtering cardiogenic oscillations is a non-trivial problem because the heart rate during sleep can be low enough to fall between the harmonic content of respiration (Cerina *et al* 2023). In some analyses, like detecting breathing rate, the fundamental frequency would be enough. Conversely, we aim to detect single breaths in respiratory signals and some morphological features, therefore preserving the whole frequency content is a necessity. To minimize the effect of cardiogenic noise, we estimated the instantaneous heart rate signal from ECG R-peaks detected with the method from Rodrigues *et al* (2021) defined as:

$$\text{HR}(t) = 1 / (t(\text{ECG}_{\text{peak}}(n)) - t(\text{ECG}_{\text{peak}}(n-1))). \quad (2)$$

With $t(*)$ the time of the n th detected ECG R-peak after the first one. The heart rate and its first two harmonics tuned a time-varying infinite impulse response (IIR) notch filterbank (updated every 0.5 s), which we used to filter SSP and Pes signals. The coefficients for the cardiac bandstop filters, given a certain heart rate $\text{HR}(t)$ and its harmonics, and a sampling rate fs are:

$$\begin{aligned} \alpha &= 0.999 \\ \omega_0(t) &= 2\pi * \text{HR}(t) * i / fs, \text{ with } i \in [1, 3] \\ H(z, t) &= \frac{1 - 2 * \cos(\omega_0(t)) * z^{-1} + z^{-2}}{1 - 2 * \alpha * \cos(\omega_0(t)) z^{-1} + \alpha^2 * z^{-2}}. \end{aligned} \quad (3)$$

Then, we applied a global low-pass 4th order Butterworth filter to remove residual noise with a cutoff frequency equal to the 10th percentile of the extracted heart rate. If the cutoff frequency is lower than 0.78 Hz, we multiply it by 1.2 to preserve higher harmonics of respiration (assuming a maximum breathing rate during sleep close to 23–25 breaths per minute). We selected these parameters trying to maximize our performance indexes (section 2.7) of SSP-Pes comparison in a subset of recordings that were highly degraded by cardiogenic noise (from a qualitative point of view). Figure 2 shows two examples of the SSP signal with small and large oscillations before and after filtering.



2.5. Breath detection

Similar to the pre-processing phase, there are common characteristics in the morphology of the respiratory signals that align the breath detection process across sensor modalities. Considering the morphology of each breath across the signals of interest, we opted for the detection of the end of inhalation as an initial fiducial point as it is less affected by noise or the breath's morphology than the start of inhalation (an highlight in figure 3). The end of inhalation corresponds to a local maximum in RIP signals (largest expansion) and a local minimum in pressure signals (largest negative pressure). For sake of conciseness, we will refer to both as peaks ($peaks_{end}$). We also calculated the spectral respiratory signal-to-noise ratio (SNR) of the pre-processed signals, derived from the power spectral density (PSD) calculated through short-time Fourier transform (STFT) and defined as:

$$SNR(n) = \frac{PSD_{resp}(n)}{PSD_{notresp}(n)} \quad (4)$$

With n being the n th segment over 30 s windows with 5 s steps of the power spectral density and the $resp$ range as the frequencies between 0.15 Hz and 0.41 Hz, where the main respiratory rate is expected, and $notresp$ being as the rest of the spectrum. We will use the SNR to threshold segments of the signal (-5 dB minimum) that are too noisy for a reliable detection of breaths' fiducial points.

The morphology of breaths can vary drastically during the night, between patients, and during respiratory events (apneas or abnormal breathing patterns like Cheyne–Stokes). Therefore, we opted for a general and almost sensor-independent approach with a funnel selection mechanism on all the signals. First, we detected candidate $peaks_{end}$ with a coarse unspecific peak detector, using the Automatic Multiscale Peak Detection (AMPD) algorithm (Scholkmann et al 2012) with a moving window of 4 s and a step of 2 s.

The initial detection is refined with a set of rules over $peaks_{end}$ distances and signal amplitudes, then we determined the candidate start of inhalation points ($peaks_{start}$) and further removed breaths with rules on inhalation amplitude and length. The rules to discard $peaks_{end}$ candidates are:

- during the *lights-on* period of the recordings, close to disconnected segments (0.1 s before and after), or with amplitude close to 0 (margin 2.5×10^{-4}). Potential causes for these peaks are sensor noise, fluctuations near sensors' disconnections, or residual cardiogenic oscillations.
- in segments with respiratory SNR < -5 dB. Potential causes are sensors' noise, body movements, respiratory events, or coughs.
- two $peaks_{end}$ closer than 2.05 s under our assumptions of maximum breathing rate during sleep (with a margin higher than the 25 breaths-per-minute mentioned before). The second $peaks_{end}$ is removed

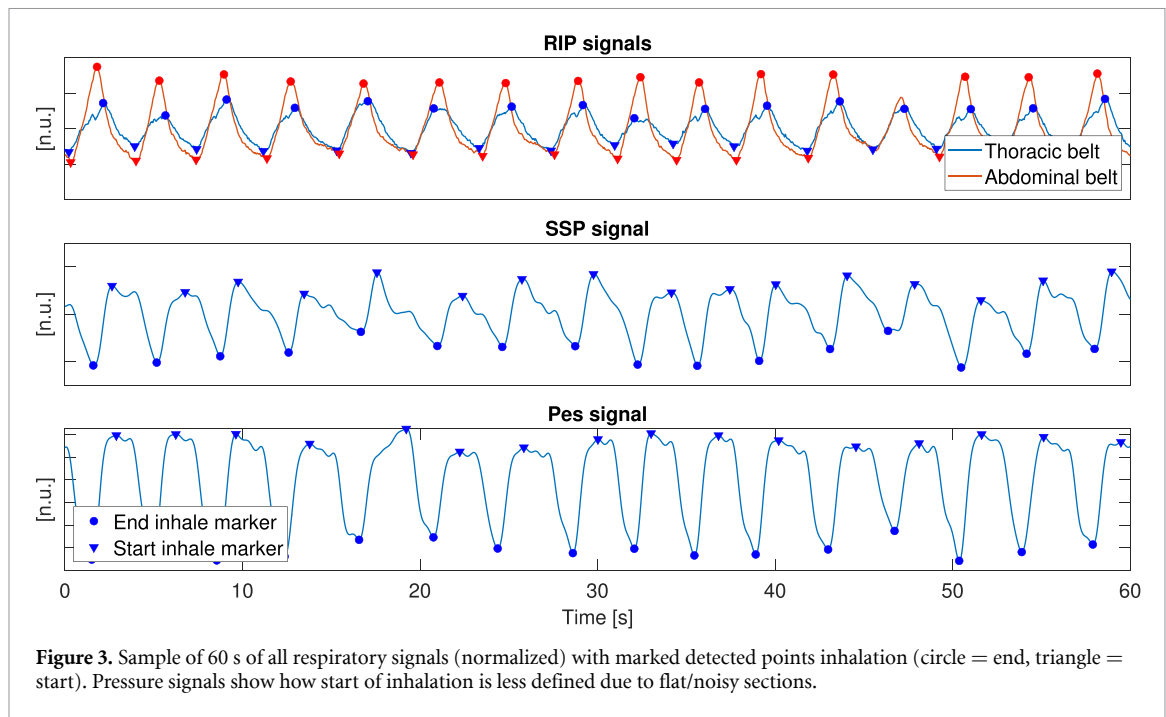


Figure 3. Sample of 60 s of all respiratory signals (normalized) with marked detected points inhalation (circle = end, triangle = start). Pressure signals show how start of inhalation is less defined due to flat/noisy sections.

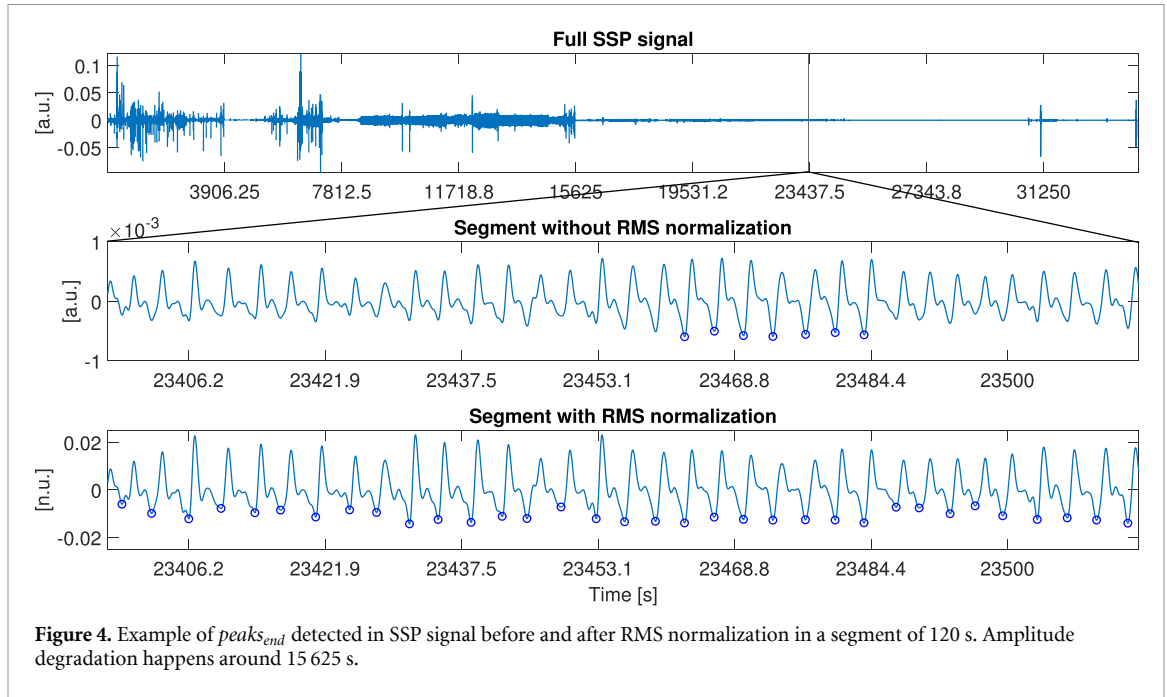
- with a negative (RIP) or positive (SSP, Pes) value in the signal, as they conflict with our assumptions on end of inhalation being a local maxima (RIP expansion) or minima (negative intrathoracic pressure, as in figure 3).
- with a $peaks_{end}$ amplitude above 99th percentile in case of heavily skewed distributions (skewness > 2.5) and other potential outliers detected with Iglewicz–Hoaglin transform (Iglewicz and Hoaglin 1993).

In our case, the Iglewicz–Hoaglin transform is preferable to the standard Z-score because the amplitude distribution of $peaks_{end}$ may be far from normal due to pathophysiological mechanisms or general sensor issues. Additionally, we removed isolated peaks that are distant more than 15 s from their direct neighbors and isolated clusters of peaks. Most of the time they represent spurious detections or a group of peaks surrounded by noisy signal sections. Other times they may be caused by periodic breathing, with around 4 to 7 breaths interleaved by apneas. We considered as a *cluster* a group of N_c peaks, with $N_c = \{1, \dots, 6\}$, with an inter-peak distance less than 10 s for the majority of them (75th percentile) and with the first and last peak of the cluster separated by the neighboring ones by more than 15 s. If the percentage of clusters with $N_c = \{4, \dots, 6\}$ was greater than 1% of total peaks detected, we removed only clusters of 3 peaks or less. Lastly, after detecting $peaks_{start}$ with a sensor-specific model (see section 2.5.2), we removed all breaths with an inhalation time too short or too long (we assumed a minimum of 0.5 s up to 5.25 s as physiologically plausible), and all breaths with an amplitude lower than 2nd percentile, as potentially too small to be physiological.

After breath detection, we estimated the breathing rate as the inverse of the average distance between $peaks_{end}$ considering 30 s windows with a step of 1 s. We discarded detected breaths with an inter-breath distance larger than 7.5 s. We then discarded a window with fewer than two valid breaths or if the sensor is disconnected for more than 25% of the window.

2.5.1. SSP signal normalization

We observed that the average root mean square (RMS) amplitude of RIP and Pes signals remain relatively stable during the night, and it is only partially affected by changes in body position or displacements of the sensor. On the other hand, the SSP sensor requires an airtight separation between the capsule and the environment to function correctly. If the adhesive keeping the sensor in place loosens, it worsens the transmissions of pressure changes from the trachea to the transducer, drastically altering the amplitude of breaths. The problem is not frequent, but it may affect some of the rules in section 2.5 that employ the statistical distribution of SSP signal at $peaks_{end}$, discarding valid peaks. In figure 4, we can see a signal that suffered a loss of contact with a large decrease in RMS after 15 625 s (or 4h 20 min). Without normalization, the selection process discarded 25% of breaths present in this specific recording.



For this reason, we normalized the SSP signal to obtain a constant RMS amplitude (0.0075 in our specific sensor, around -4.8 dB). We calculate the RMS over 10 s (step of 1 sample) with the formula:

$$RMS = \sqrt{\frac{1}{n} \sum_{t=0}^n a_t * x(t)^2} \quad (5)$$

with n the number of samples (1280 with $f_s = 128$ Hz) and a_t the correction coefficient we need. In the original signal $a_t = 1$. Inverting the formula and fixing the desired output RMS_{out} we obtain an amplification coefficient for each sample (considering the same windows of equation (5)) equal to:

$$a = \sqrt{\frac{n * RMS_{out}^2}{\sum_{t=0}^n x(t)^2}}. \quad (6)$$

If the RMS of the original signal is too small or disconnected (< -10 dB) a is kept equal to 1 to avoid artifacts caused by excessive corrections or numerical errors. We want to specify that the normalized SSP signal is employed only to detect breaths, and we used the SSP signal pre-normalization to detect patterns of respiratory effort presented in section 2.6.

2.5.2. Detection of inhalation start

The harmonic content of RIP belts is limited compared to pressure signals, almost sinusoidal, and the detection of inhalation start is straightforward. For each potential end inhalation point $peaks_{end}(i)$ (with $i = \{1, \dots, N_{peaks}\}$), we detected the associated start inhalation as the relative minimum found between $peaks_{end}(i-1)$ and $peaks_{end}(i)$, or using a window of 4 s in case the two breaths were distant more than 8 s (e.g. because of apneas).

Conversely, respiratory pressure signals morphology is either a biphasic, flat, or slowly decreasing section followed by a steeper decrease, potentially indicative of the action of respiratory muscles active with different timings (see figures 2 and 3). We decided to characterize each breath between two end inhalations as a trapezoid and the local maxima in the signal closer in time to its upper-left angle as the candidate start of inhalation. For each $peaks_{end}(i)$, we fit a trapezoid using least squares optimization and constraining the lower-left angle as $peaks_{end}(i-1)$, the lower-right angle as $peaks_{end}(i)$ and the upper side to be flat and with a length between 1.25 s and 2 s. Also, in this case, each pair of $peaks_{end}$ is excluded if they are separated in time by more than 8 s.

2.6. Respiratory effort patterns

To examine if different sensors follow similar amplitude patterns of the recorded signals, we considered overlapping sequences of five breaths as the basic element of our comparison. We considered a sequence

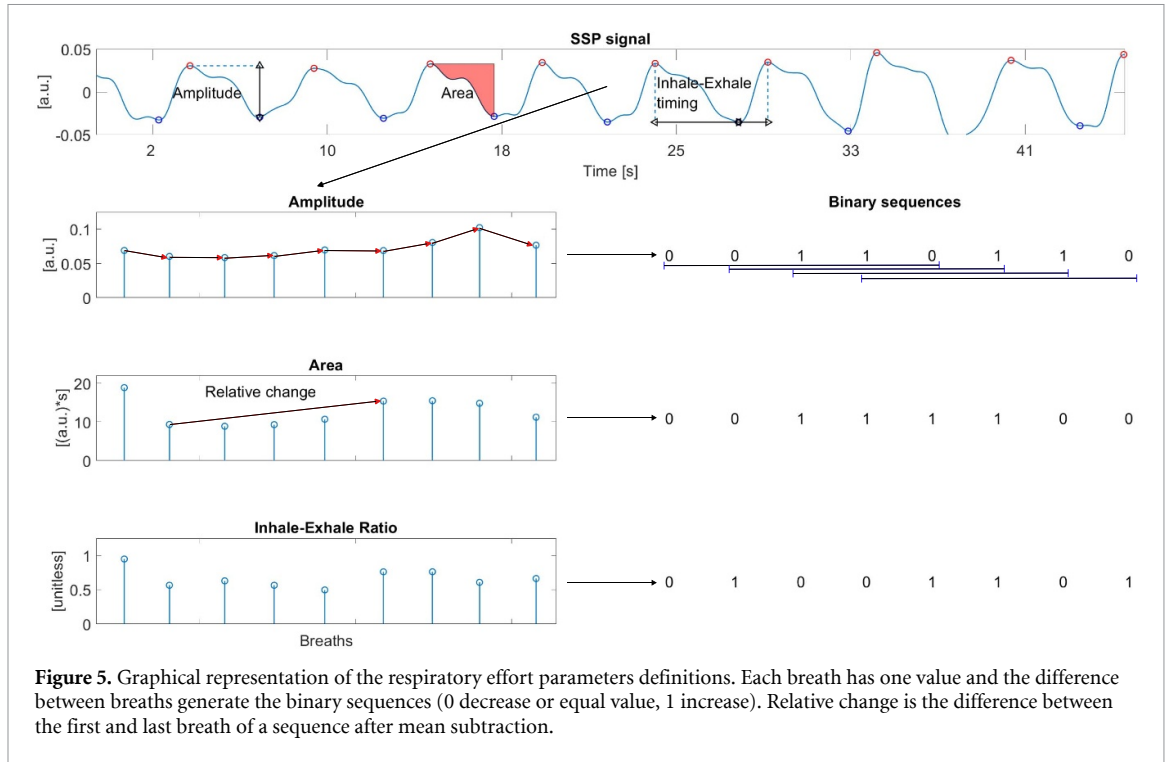


Figure 5. Graphical representation of the respiratory effort parameters definitions. Each breath has one value and the difference between breaths generate the binary sequences (0 decrease or equal value, 1 increase). Relative change is the difference between the first and last breath of a sequence after mean subtraction.

valid using the same clustering considerations in section 2.5 to avoid patterns overarching sensor disconnections or mis-detected breaths. In our specific case, we chose five breaths according to the average length of respiratory events in our dataset (inter-quartile range four breaths). For each sequence, we calculated the amplitude, area under the curve and the inhalation ratio of breaths as:

$$\text{amplitude}(i) = |x(\text{peaks}_{\text{end}}(i)) - x(\text{peaks}_{\text{start}}(i))| \quad (7)$$

$$\text{area}(i) = \sum_{t=\text{peaks}_{\text{start}}(i)}^{\text{peaks}_{\text{end}}(i)} |x(t) - x(\text{peaks}_{\text{start}}(i))| \quad (8)$$

$$\text{ratio I:E}(i) = \frac{T_{\text{inhale}}(i)}{T_{\text{exhale}}(i)} = \frac{\text{peaks}_{\text{end}}(i) - \text{peaks}_{\text{start}}(i)}{\text{peaks}_{\text{end}}(i+1) - \text{peaks}_{\text{end}}(i)}. \quad (9)$$

For each sequence of breaths and each parameter (amplitude, area, inhalation ratio), we extracted binary patterns of direction of change between breaths, assigning a value of 1/+ in case of increase or equivalence, and 0/- for decrease. We also measured the relative change of each parameter for each sequence as the difference between the last and first breath, after subtracting the mean. An example of these patterns is visible in figure 5.

2.7. Performance evaluation

To understand if SSP is a valid alternative to other respiratory sensors, we performed multiple comparisons over different aspects.

Starting from breath detection, we analysed the SSP signal using RIP belts and Pes as reference. We also compared RIP belts with Pes and between each other. Using $\text{peaks}_{\text{end}}$ as the fiducial point, we considered a breath in the tested signal as correct if another breath is detected in the reference signal closer than 1.05 s (a limit we determined heuristically in a subset of signals). To avoid spurious degradation of the performance, we checked if the reference signal can contain a breath in the range, ignoring breaths where the reference is disconnected or has a low SNR. This is an extra precaution, since excessively degraded signals were excluded *a-priori* from the comparison. The coverage of potential common breaths, or the percentage of non-disconnected signal in both sensors, always exceeded 95%.

For each recording, we counted the number of reference breaths (n_{ref}), test breaths (n_{test}) and the number of correctly detected breaths in the test signals (n_{correct}). Breath detection performance was measured according to the following metrics:

$$\text{Sensitivity}(\%), \quad Se = \frac{n_{\text{correct}}}{n_{\text{ref}}} \quad (10)$$

$$\text{Positive predictive value (PPV)(\%),} \quad PPV = \frac{n_{\text{correct}}}{n_{\text{test}}} \quad (11)$$

$$\text{F1-score(\%),} \quad F1 = \frac{2 * PPV * Se}{PPV + Se}. \quad (12)$$

For each correct breath, we calculated the temporal distance with the reference breath, both for $peaks_{end}$ and $peaks_{start}$. This distance is a usual measure of localization error, but in our case the signals are naturally shifted due to different timings of respiratory components. The timings could also be influenced by various factors, such as the breathing rate, airflow speed, or anatomical dimensions of airways. Unfortunately, none of the pairs of signals considered here can have a distance equal to zero and we cannot measure the localization error alone. Therefore, we measured the average distance and its standard deviation pooling all correct breaths in our population, and assessed potential effects on F1-score as a linear function of detection spread. Results are reported as the slope of linear regression and the correlation as adjusted R^2 (R^2_{adj} in the text).

We then evaluated differences in breathing rate estimation as the mean absolute percentage error (MAPE) for each recording and the Bland–Altman agreement aggregating results on the whole population. To have a perspective on how often the estimation may fail (due to situations described in section 2.5) we measured the breathing rate coverage. Coverage is defined as the number of segments valid in both reference and test signal over the total number of segments.

We evaluated patterns in respiratory effort after the construction of breaths' sequences. The first index of the performance is the number of valid sequences, which can be seen as an extension of single breath detection sensitivity.

Then, we calculated the distance correlation (Székely *et al* 2007) between the extracted parameters for each sequence (referred here as sequence-by-sequence correlation) and averaged it at recording level to allow for paired statistical comparisons. We opted for distance correlation because it provides a more robust interpretation in case of potential non-linear relationships and small samples (as in our case) compared to the widespread Pearson's or Spearman's correlations. Distance correlation is also always positive in the range 0 to 1. After reducing each sequence to a single value through relative change we measured the global correlation over the whole night (relative change correlation in the results). At this level, Spearman's and distance correlation show similar values, but we opted to present results with the latter to be consistent with other results.

Other than measuring how patterns are proportional to each other in magnitude, it is also important to understand if they change in the same direction (e.g. the Pes and another signal both increasing in amplitude during an apnea). To do so we calculated the Hamming's distance between test and reference sequences. This distance count the number of different symbols in a string, therefore it is a direct quantification of the breaths with different directions in our binary sequences. In our specific case, the distance can range from 0 (when all breaths have the same pattern) to 4 (completely diametrically opposite direction). To account for noise in the data, we assumed sequences with a distance up to 2 as 'close' (e.g. both 1011 and 1101 are generally increasing). Over the whole recording, the similarity score for patterns is given by:

$$\text{Similarity}(\%) = \frac{\text{sequences}_{\text{close}}}{\text{sequences}_{\text{total}}}. \quad (13)$$

In all the comparisons above, we measured how the sleep context (sleep stage, body position, presence of respiratory events) and demographic factors can affect the sensors annotating each breath and recording with the relevant context. If a breath starts and ends across a transition (such as for sleep stages), we associate the most prevalent class (for example 3 s in REM and 1 s in N1 will be annotated as REM). It must be noted that due to AASM 30 s epochs, the transition instant between sleep stages is not annotated exactly. Conversely, body position and respiratory events are sampled with sub-second granularity. In the case of respiratory events, we consider a range of 5 s before and after the event to account for unexpected discrepancies in respiratory signals (for example, the resumption of breathing after an obstructive apnea). In respiratory patterns, differently from breath-by-breath analysis, the state corresponds to the most frequent in the sequence. For example, a sequence with 3 hypopneic breaths and 2 undisturbed will be marked as hypopneic ('H').

We assessed the effect of sleep context on breath detection's sensitivity using with two proportion Z -test with $\alpha = 0.05$. The effect on respiratory patterns' metrics was assessed with Wilcoxon rank-sum test with $\alpha = 0.05$. For sake of conciseness we focused on breaths' amplitude patterns as they are a common feature of interest in medical literature (for example, Sforza *et al* 1998). Unless stated otherwise, the results can be generalized to the other pattern parameters considered in our analyses. Complete results of this analysis are available in supplemental materials S7.

Lastly, we assessed the effects of sex, age, BMI, and AHI on our performance metrics. Unless stated otherwise, all hypotheses were tested with Mann–Whitney U two-tailed test, with a threshold at $\alpha = 0.05$ and multiple tests were corrected with Benjamini–Hochberg false discovery rate procedure.

3. Results

Before delving in the quantitative results, we want to outline a conceptual difference among respiratory sensors that can support their interpretation. From previous literature we know that a perfect agreement on respiratory events scored manually with RIP belts, Pes or SSP is lower than 100% (Sabil et al 2019). Specifically, the authors in Sabil et al (2019) reported a sensitivity of 97% for both SSP and RIP against Pes regarding obstructive apneas, but for central and mixed apneas it reduces to around 94% for SSP and 87% for RIP belts. We did not re-score our dataset using Pes or SSP so currently we cannot provide a quantitative comparison. Nevertheless, the underlying problem may still be present and figure 6 shows an example of it.

Here, the first central apnea (1) could be potentially scored as obstructive observing the Pes, and vice-versa the last obstructive apnea (5) could be central looking at the SSP. Some hypopneas (2, 3) show very small fluctuations in the SSP and the thoracic RIP, but well-visible breaths in Pes and abdominal RIP. A single breathing pause is present in all four signals around the 107 s mark (4), but it is likely shorter than 10 s and it was not scored as an event. In all the cases above, the number of detected breaths will be inconsistent between signals, with a lower performance that is independent from our proposed method.

3.1. Breath detection

Table 3 shows the performance of breath detection in the whole population and by reference sensor. The Pes and SSP signals were used only as reference and test, respectively. For sake of conciseness, only statistical differences with SSP are reported. Complete results can be found in supplemental materials S1 to S3.

The SSP signal obtained good results against all reference signals, with overall sensitivity and PPV between 78% and 82%. The differences in performance of SSP against reference signals and with RIP belts against Pes were not significant. Only the abdominal RIP was significantly better than SSP against Pes ($p < .05$). The results did not differ considering only the 20 recordings with all sensors. Even if the median F1-scores were not significantly different, the 1st quartile of SSP against Pes was lower than for RIP belts. We observed that 10/20 SSP recordings had a F1-score lower than 80% against Pes, but only 2 out of 20 with RIP belts as test sensor. There is a potential (although not significant) negative effect of Pes presence on SSP against RIP. After matching the 20 recordings with Pes with other 20 without Pes for Age, AHI and male/female proportion, the median F1-score of SSP against thoracic RIP was 66.75% if Pes was present versus 74.68% ($p = 0.29$) if not, and 69.73% versus 77.27% with the abdominal belt as reference ($p = 0.65$).

Evaluating one RIP belt against the other as reference, they show internal consistency, with every performance index was above 90%, and only 8 recordings out of 207 with a F1-score lower than 80%. No significant differences were observed.

Table 4 shows the distribution of temporal differences of inhalation $peaks_{end}$ and $peaks_{start}$ between signals on the whole population. We observed a negative relationship between the spread of $peaks_{end}$ distance (measured as its standard deviation) and our performance indexes with Pes reference, even if we cannot discern the effects of $peaks_{end}$ localization error from physiological aspects of breaths' timing. The F1-score of SSP reduced by 2.9% on average per 100 ms of spread ($R^2_{adj} = 0.22$), and a stronger effect for RIP belts with -3.3% in abdominal RIP ($R^2_{adj} = 0.52$) and -8.1% in thoracic RIP ($R^2_{adj} = 0.71$). The correlation for other reference signals was below 0.1 and considered negligible. The start of inhalation distance showed a higher spread due to the localization jitter in pressure signals (as from example in figure 3). There were no significant correlations with $peaks_{start}$ because all performance indexes depend only on $peaks_{end}$. Complete results are available in supplemental materials S4.

3.2. Breathing rate estimation

Table 5 shows the comparison between sensors in terms of breathing rate estimation. We did not observe significant differences in MAPE both for SSP against other reference sensors and for SSP and RIP with respect to Pes. The Bland–Altman analysis reveals an average bias close to 0 between all sensors, with 95% limits of agreement under 3 breaths-per-minute. The SSP had an error below 5 breaths-per-minute for more than 98% of samples across all signals. The breathing rate coverage was comparable for SSP and RIP belts, with an average close to 85%. On average, the SSP showed a tendency of higher coverage with Pes compared to RIP, but the difference was not significant. The differences in MAPE and coverage, with and without Pes in a matched subset (see section 3.1 for matching rules), were negligible. There was a small effect of breaths' distance spread on performance, relevant only for RIP belts against Pes. The MAPE of abdominal RIP increased by 0.56% per 100 ms of spread ($R^2_{adj} = 0.19$) and by 1.58% for the thoracic RIP ($R^2_{adj} = 0.43$).

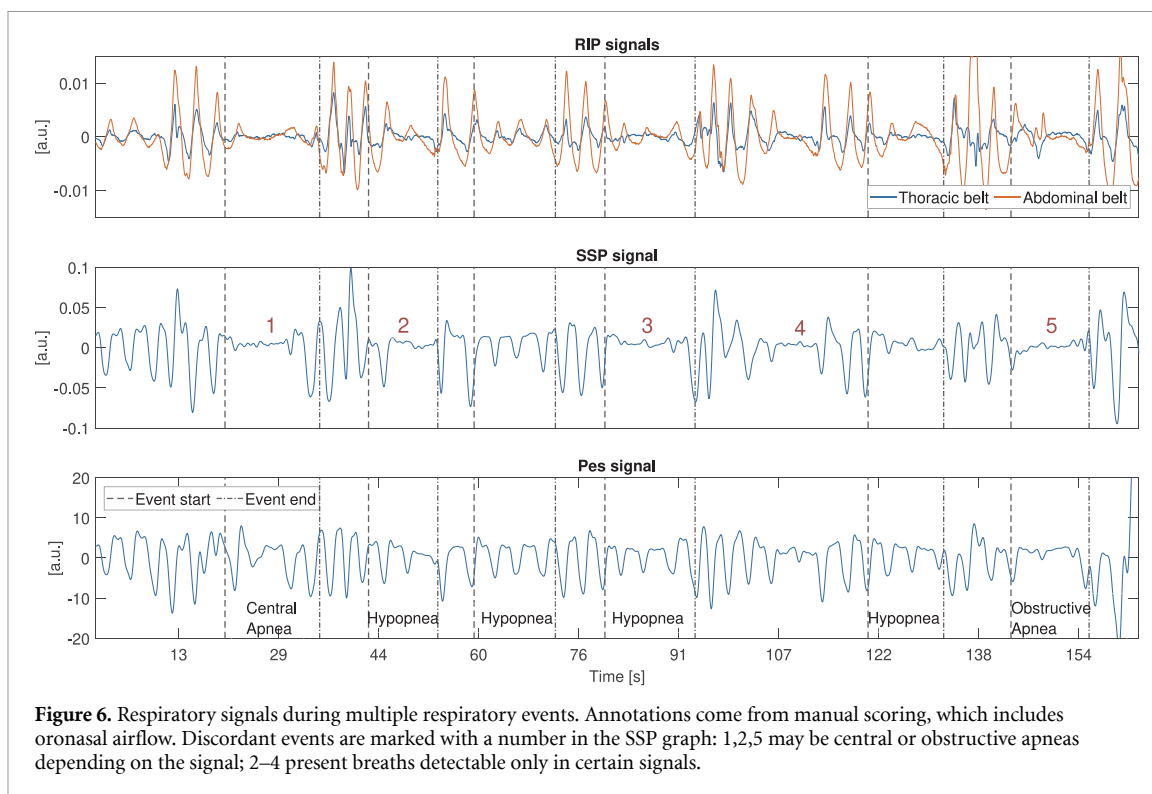


Table 3. Breath detection performance indexes.

Test signal	Reference signal (# recordings)		
	Pes (20)	Abdomen RIP (207)	Thorax RIP (207)
Sensitivity (%)			
SSP	81.60 (66.43 : 89.06)	79.83 (65.32 : 88.62)	78.21 (61.61 : 88.49)
Abdomen RIP	92.54* (85.15 : 94.81)	—	95.10 (90.98 : 96.93)
Thorax RIP	87.12 (81.55 : 91.27)	92.70 (88.80 : 94.87)	—
Positive predictive value (%)			
SSP	78.96 (68.87 : 83.86)	81.69 (68.48 : 90.16)	78.04 (62.58 : 87.48)
Abdomen RIP	85.95* (83.18 : 88.96)	—	93.06 (89.75 : 94.9)
Thorax RIP	86.50 (81.66 : 89.35)	95.26 (92.04 : 96.98)	—
F1-score (%)			
SSP	80.58 (66.87 : 87.69)	80.88 (67.47 : 88.92)	77.62 (62.76 : 88.10)
Abdomen RIP	89.56* (84.33 : 91.74)	—	93.61 (90.28 : 95.63)
Thorax RIP	86.97 (81.13 : 90.00)	93.65 (90.65 : 95.51)	—

Data presented as median (first quartile : third quartile). The number of recordings is the number of valid test/reference signal pairs. The ‘*’ indicates $p < .05$ in Mann–Whitney U test against the results in the SSP:Pes cell.

3.3. Patterns in respiratory effort

Table 6 shows all the relevant performance metrics against Pes on the subset of 20 recording where esophageal pressure was available. We tested significant differences with Wilcoxon one-sided signed-rank tests with $\alpha = 0.025$.

Table 4. Temporal distance between the end of inhalation measured with different signals.

Test signal	Reference signal (# recordings)		
	Pes (20)	Abdomen RIP (207)	Thorax RIP (207)
End inhalation distance (s)			
SSP	-0.078 ± 0.477	-0.322 ± 0.436	-0.375 ± 0.440
Abdomen RIP	0.360 ± 0.225	—	-0.090 ± 0.228
Thorax RIP	0.435 ± 0.275	0.090 ± 0.228	—
Start inhalation distance (s)			
SSP	-0.171 ± 0.856	0.008 ± 1.360	-0.138 ± 1.322
Abdomen RIP	0.861 ± 0.713	—	-0.058 ± 0.444
Thorax RIP	0.840 ± 0.748	0.058 ± 0.444	—

Data presented as mean \pm standard deviation (overall skewness < 1). The number of recordings is the number of valid test/reference signal pairs.

Table 5. Performance indexes in breathing rate estimation.

Test signal	Reference signal (# recordings)		
	Pes (20)	Abdomen RIP (207)	Thorax RIP (207)
Median Absolute Percentage Error (MAPE)(%)			
SSP	4.18 (3.71 : 5.61)	4.18 (3.34 : 5.27)	4.14 (3.42 : 5.32)
Abdomen RIP	3.52 (2.96 : 4.43)	—	2.83 (2.20 : 3.55)
Thorax RIP	3.36 (2.94 : 5.52)	2.89 (2.24 : 3.63)	—
Bland–Altman Bias and Limits-of-Agreement (bpm)			
SSP	$-0.04 (-2.87 : 2.78)$	$-0.04 (-2.79 : 2.72)$	$-0.09 (-2.87 : 2.68)$
Abdomen RIP	$0.00 (-2.47 : 2.46)$	—	$-0.05 (-2.26 : 2.17)$
Thorax RIP	$0.10 (-2.71 : 2.90)$	$0.05 (-2.17 : 2.26)$	—
Coverage (%)			
SSP	77.90 (67.00 : 85.14)	85.28 (68.36 : 91.63)	84.26 (70.00 : 90.54)
Abdomen RIP	71.90 (62.54 : 86.16)	—	85.09 (68.35 : 91.63)
Thorax RIP	75.34 (61.83 : 87.78)	85.09 (68.35 : 91.63)	—

Data presented as median (first quartile : third quartile) for MAPE and coverage, and median bias (95% limits of agreement) for Bland–Altman. The number of recordings is the number of valid test/reference signal pairs. No significant differences found in SSP comparisons.

By construction, sequences employ only correctly detected breaths, thus the valid sequences performance is proportional sensitivity in table 3. Overall, the SSP performance was slightly lower than RIP, but the difference was significant only versus the abdominal RIP.

The performance of sequence-by-sequence correlation was comparable across sensors, with ranges from 0.67 to 0.76. The SSP correlates significantly better with Pes compared to thoracic RIP in breaths' sequence amplitude and inhale ratio, but less than abdominal RIP in the area. All other differences were not significant.

While the correlation of single sequences was fairly high, the relative change of patterns over the night was lower, with all comparisons having an average correlation of less than 0.5. Nevertheless, the three sensors are comparable: the SSP had a significantly higher correlation with Pes in the inhale ratio, and slightly higher, but not significantly, in the breaths' amplitude. Despite a low correlation, the similarity results tell that all three sensors tend to follow Pes in the direction of change. The SSP was not significantly different from RIP belts regarding amplitude and inhale ratio, but under-performed in terms of breaths' area.

Considering the distance spread of $peaks_{start}$ and $peaks_{end}$ between the respiratory sensors and Pes we observed slightly negative linear relationships with the patterns metrics. The effect was overall quite small, with a reduction around 0.01 per 100 ms of spread in correlations, and less than 2% per 100 ms of spread in similarity metrics. The R_{adj}^2 of each model is around 0.37. Complete results are available in supplemental materials S5. The RIP belts were influenced significantly only in the inhale ratio metrics, with a reduction in sequence-by-sequence and relative correlation for the thoracic belt due to $peaks_{start}$ spread, and a reduction in similarity for the abdominal belt due to $peaks_{end}$ spread. The effect on SSP was broader, with all

Table 6. Comparison of local respiratory patterns metrics with Pes signal as reference (n recordings = 20).

Parameter	Test signal		
	SSP	Abdomen RIP	Thorax RIP
Valid sequences (%)			
Sequences	76.46 (59.41 : 84.26)	84.69 [†] (82.73 : 89.93)	80.41 (75.72 : 88.73)
Sequence-by-sequence correlation			
Amplitude	0.724 (0.689 : 0.760)	0.738 (0.704 : 0.767)	0.692* (0.687 : 0.708)
Area	0.688 (0.670 : 0.710)	0.720 [†] (0.700 : 0.751)	0.713 (0.697 : 0.724)
Inhale:Exhale	0.700 (0.684 : 0.771)	0.691 (0.678 : 0.714)	0.685* (0.672 : 0.702)
Relative change correlation			
Amplitude	0.470 (0.311 : 0.553)	0.446 (0.341 : 0.553)	0.340 (0.276 : 0.441)
Area	0.318 (0.263 : 0.402)	0.366 (0.306 : 0.473)	0.355 (0.304 : 0.401)
Inhale:Exhale	0.295 (0.210 : 0.412)	0.238* (0.186 : 0.310)	0.248* (0.193 : 0.295)
Similarity (%)			
Amplitude	80.79 (75.10 : 84.21)	82.51 (78.83 : 86.91)	78.24 (76.99 : 79.58)
Area	77.73 (75.61 : 82.35)	83.59 [†] (79.54 : 85.70)	83.36 [†] (80.51 : 84.79)
Inhale:Exhale	83.32 (77.93 : 87.52)	82.88 (79.94 : 85.52)	82.53 (79.30 : 84.29)

Data presented as median (first quartile : third quartile).

A ^{*} indicates a better performance of SSP compared to RIP, [†] a worse performance with Wilcoxon paired test at $\alpha = 0.025$.

sequence-by-sequence correlations, all the similarity, and the inhale ratio relative correlation significantly influenced by both $peaks_{start}$ and $peaks_{end}$ spread.

3.4. Effect of sleep-related parameters on breath detection

To analyze how respiratory events, sleep stages, and sleeping position may affect the performance of SSP we considered only the subset of 20 recordings with Pes signal available. In this way, breaths can be pooled and the SSP performance is comparable across all reference sensors. This selection implies some quantitative differences with results in table 3. In figure 7 we report only a graphical representation of differences in sensitivity, full results with significance levels are available in supplemental materials S6. The number of breaths analyzed is reported in table 7. We note that due to the 5 s range around events to annotate breaths, and the events' discrepancies mentioned above, the number of breaths in central apneas can be higher than 0.

With respect to respiratory events, we observed significant differences between every pair of events in each reference sensor. Only central and mixed apneas in Pes and abdominal RIP, and all types of apneas in thoracic RIP are not significant. Differences across reference sensors were overall negligible. The SSP only slightly under-performed against thoracic RIP compared to abdominal RIP and Pes during obstructive and mixed apneas.

All sleep stages differed significantly from each other in every reference sensor. The sensitivity increased from wake to deeper sleep, and slightly decreased in REM. There were no notable differences across reference sensors, except a variation close to 14% in REM for thoracic RIP compared to the others, and a 7% reduction with Pes and thoracic RIP references compared to abdominal RIP.

Regarding the sleeping position, the sensitivity was the lowest in the supine position, increased in lateral positions, and the highest in prone position. All increases were statistically significant. The changes in RIP belts were contained in a range around 4%–5%, while Pes shows the largest differences, from 71% in supine position to 88% in prone position.

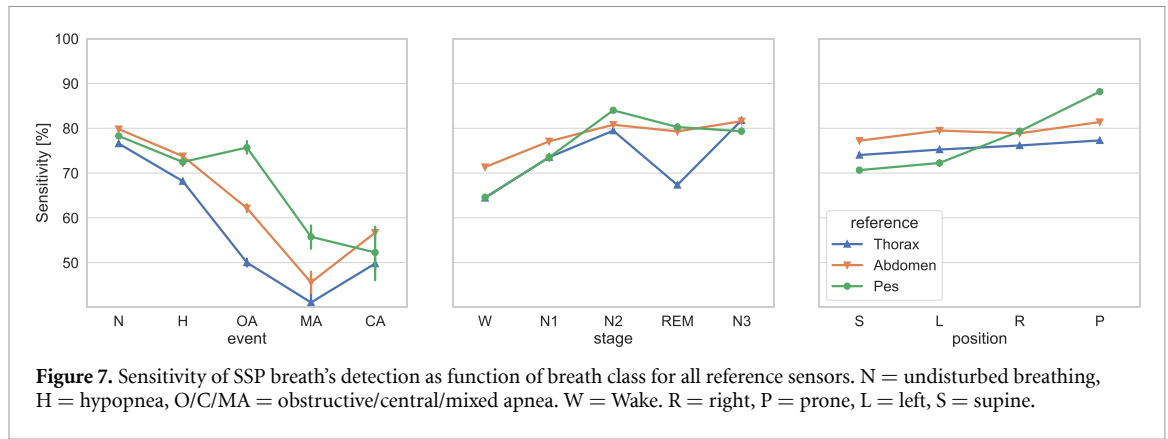


Table 7. Number of breaths analyzed in each sleep-related state. Total count may differ due to exclusion of unclassified sleep stages and positions.

Respiratory event		Sleep stage		Sleeping position	
N	325 680	W	70 911	R	91 605
H	30 729	N1	42 903	P	92 556
OA	9921	N2	146 760	L	62 655
MA	4047	N3	52 563	S	122 148
CA	1179	REM	52 908		

N = undisturbed breathing, H = hypopnea, O/C/MA = obstructive/central/mixed apnea. W = Wake.
R = right, P = prone, L = left, S = supine.

Table 8. Average number of sequence analyzed in each sleep-related state. Total count of sequences may differ due to exclusion of sequences in unclassified stages and positions.

Respiratory event		Sleep stage		Sleeping position	
N	71 032	W	10 272	R	20 167
H	5600	N1	7953	P	20 204
OA	1321	N2	37 106	L	20 167
MA	119	N3	12 641	S	23 751
CA	9	REM	9645		

N = undisturbed breathing, H = hypopnea, O/C/MA = obstructive/central/mixed apnea. W = Wake.
R = right, P = prone, L = left, S = supine.

3.5. Effect of sleep-related parameters on effort patterns

Similarly to section 3.4 we associated each local sequence to a respiratory state, sleep stage or position to examine their effects on breath characterization performance. Due to sensors' discrepancies in sensitivity, the number of sequences analyzed in each state is reported in table 8 as the average between sensors. By construction, the number of valid sequences depends on breath detection's sensitivity, so the results of section 3.4 apply also here and will not be repeated.

Starting from sequence-by-sequence correlation in figure 8, the abdominal RIP was significantly better than SSP in all respiratory events. The SSP was comparable with thoracic RIP in all apneas, and significantly better during normal breathing and hypopneas.

Across sleep stages, the SSP had a significantly lower correlation during wake, but the difference decreased in deeper sleep and REM. Overall, the SSP had a lower performance than abdominal RIP, but was better than thoracic RIP. We did not observe significant differences caused by sleeping position. With regards to patterns similarity, the results mimic those of sequence-by-sequence correlation, and we omitted them from the main discussion (complete results in the supplemental materials S7). These results were not surprising, as we can expect that a high correlation in a single sequence is indicative of two signals changing in similar directions. Conversely, diverging patterns will have low correlation and similarity. The differences due to sleep context between sensors are more evident looking at the correlation of relative changes in figure 9.

In the SSP and abdominal RIP (but not the thoracic one) the correlation increased from normal breathing to hypopneas and obstructive apneas. The correlation during central apneas was high, but the small sample size does not allow strong hypotheses. Overall, there were no statistically significant differences between sensors. Considering sleep stages, none of the differences between sleep stages were significant (except for N2 and N3 versus REM in Thoracic RIP), but the patterns are still interesting. The RIP belts

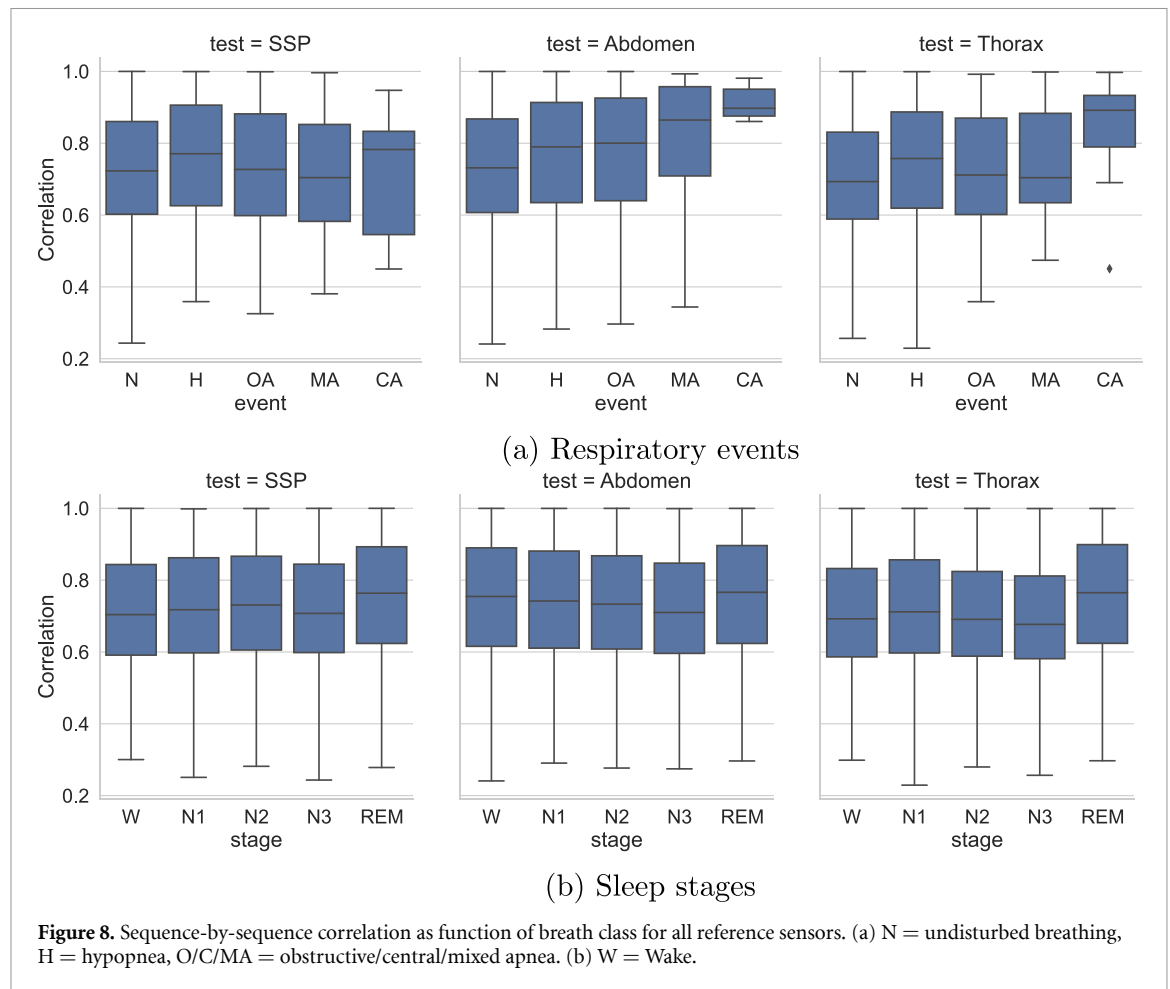


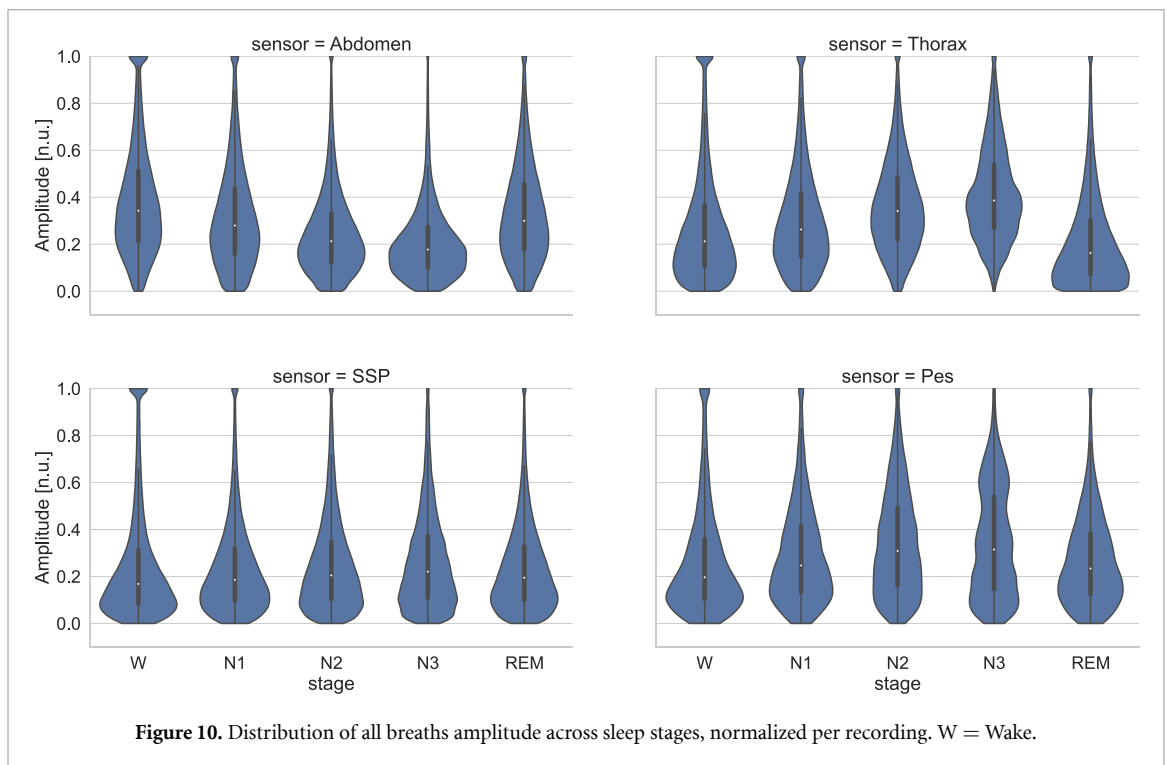
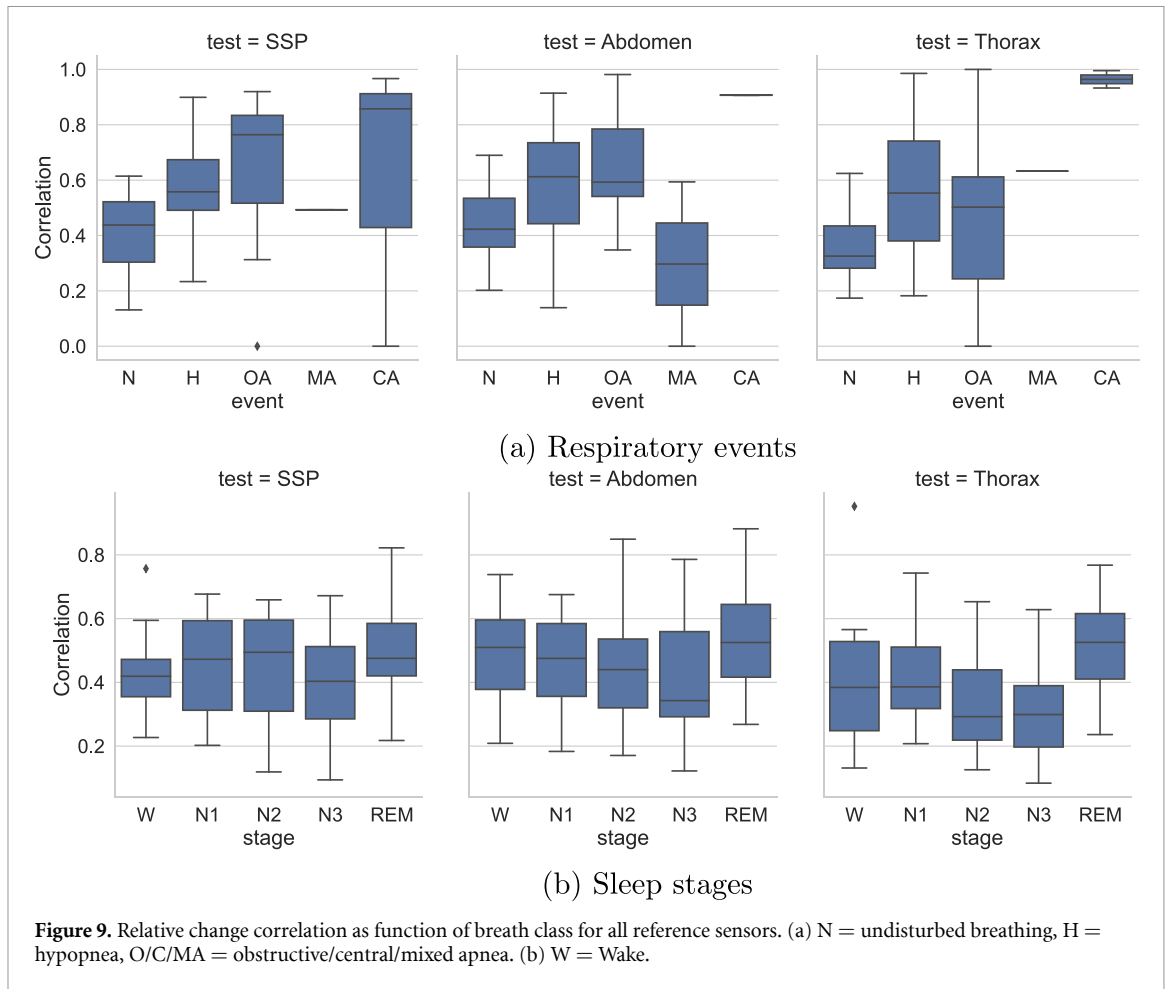
Figure 8. Sequence-by-sequence correlation as function of breath class for all reference sensors. (a) N = undisturbed breathing, H = hypopnea, O/C/MA = obstructive/central/mixed apnea. (b) W = Wake.

shown a decreasing correlation from wakefulness to deeper stages, followed by an increase during REM. The SSP sensor instead seemed to be more stable across sleep stages, and had the same increase in REM. In line with sequence-by-sequence correlation, we did not observe relevant changes in performance caused by different sleeping positions (not shown here).

While analyzing the effects of sleep stages on patterns, we observed peculiar relationships across sensors in the amplitude of breaths. We pooled all breaths in each recording and normalized their amplitude between the 1st and 99th percentile, then plotted their distribution across sleep stages as reported here in figure 10. We can see that the two RIP belts had an opposite behavior going from wakefulness to deeper stages, with the abdominal belt decreasing in amplitude, while the thoracic one increased. This mechanism seems to reflect in the amplitude of P_{es} , with an average increase in deeper stages, but also a larger spread, while the SSP seems to be less influenced.

3.6. Effects of demographic factors

Considering the F1-score and breathing rate MAPE, the SSP sensor performed significantly better ($p < 0.01$) in female participants against both RIP belts, but the difference was not significant against P_{es} . However, in the P_{es} /no- P_{es} subset mentioned in section 3.1 the difference between male and female was not significant. We did not observe significant correlations of F1-score and MAPE with age or BMI. Interestingly, there were no significant differences in performance for participants younger than 18 years compared to adults in RIP belts (not tested on P_{es} because the whole subset is adults). Regarding the effect of AHI, the impact on performance was limited to a small variation of -0.15% and $+0.05\%$ on average per events/hour for F1-score and MAPE, respectively. However, we did not observe significant differences when grouping participants by OSA severity classes according to the standard thresholds (5, 15, and 30). Regarding performance metrics in respiratory patterns, there were no significant differences in any performance metric for sex or age. BMI had a small effect on sensitivity in RIP belts (around 1.5% per unit of BMI) and similarity (around 1.1% in area and 1.7 in inhale ratio), but not significant in SSP and any other pattern metrics. Only the sequences' sensitivity was negatively affected by the presence of respiratory events, as represented by the



AHI, with thoracic RIP being the most affected (-0.581% per event/hr), followed by SSP (-0.351%) and abdominal belt (-0.221%) in line with breath detection results. The tables with complete results for each test are available in the supplemental materials S8 and S9.

4. Discussion

We present a direct comparison of suprasternal notch pressure (SSP), with sensors commonly employed in sleep research: respiratory inductance belts (RIP) and esophageal catheters (Pes). The main goal was to understand if the suprasternal sensor can substitute other sensors and if it is reliable for a whole night of recording. A secondary goal was to identify if SSP can be used to analyze short patterns of respiratory effort in place of the more invasive esophageal catheter. Since RIP belts are the de-facto standard in clinical practice, we also compared them with Pes to have a frame of reference for SSP performance. To achieve these goals we designed algorithms to detect breath fiducial points (start and end of inhalation) for every type of sensor. The algorithms were tested on 207 PSG recordings from participants undergoing a sleep test for suspected sleep disorders.

4.1. Qualitative advantages of SSP sensor

Considering current practices for PSG, there are some clear advantages of the SSP sensor over esophageal catheters. Regarding usability, the SSP sensor is simply placed on the skin with a medical-grade adhesive and does not require the same clinical expertise necessary for esophageal manometry. The SSP sensor shows also a higher acceptance from patients compared to the discomfort of the catheter passing through the nose. Since the esophageal catheter cannot be fixed in place with a balloon (compared to, for example, patients with mechanical ventilation), it is prone to movements; whereas the SSP sensor is more stable during changes in sleeping position. The sensor remains outside of the body, so it is also easier for a sleep technician to fix it in case of displacements or disconnections during the night.

From a sleep scoring point of view, we observed in the previous sections some discrepancies with RIP and Pes, and that the SSP and RIP signals do remain a proxy of intra-thoracic pressure and respiratory effort. At the same time, the examples in figures 3 and 6 show that the characteristics of the two signals are quite similar, so a sleep technician trained to score Pes events will be able to score also SSP signals (demonstrated also in the comparison of manual scorers by Sabil *et al* 2019). Lastly, we measured how the SSP signal closely follows Pes, with a shorter latency compared to RIP signals (see table 4) or other surrogate measures of respiratory effort, making it a valid alternative for real-time analyses.

4.2. Breath detection and breathing rate

First, we analyzed the performance of our breath detection algorithms. We were able to correctly detect close to 80% of breaths on average in the SSP signal against all the other signals. However, RIP belts outperformed SSP in breath detection against Pes with an average difference up to 10% and a large performance spread. A potential culprit may be the technological readiness of the SSP sensor itself. Compared to well-tested respiratory belts and esophageal catheters, the SSP is still a research device, and some drawbacks of it are still unexplored or sparsely documented. In our dataset, 7% of SSP recordings had to be excluded *a-priori* due to missing signal compared to less than 0.01% of other signals. While all respiratory sensors had different issues and present missing bouts of data during the night, the SSP signal may degrade if the capsule is not sealed (see example in figure 4) or if it moved abruptly (for example, the participant unintentionally touching it). We believe that improvements in the physical design of the sensor (for example a soft patch instead of a rigid capsule) may reduce sealing issues. We should note that the presence of the catheter itself may be detrimental for the SSP signal as it alters the physiological characteristics of the upper airway. In a subgroup of recordings matched for demographic factors, we observed a reduction of the performance of SSP with respect to RIP belts if the Pes was also present compared to when it was absent. A proper split-test in controlled settings is required to better understand the effect of the catheter on both SSP and respiratory belts.

After the detection of breaths, we measured the temporal distance between breaths from each sensor. Generally, the distance of each correctly detected point with the reference is named localization error (e.g. Charlton *et al* 2022), but in our case a physiological delay exists between reference and test signals. This delay can be explained by their relative position in the airway and the different activation timing of respiratory muscles. Interpreting table 4, the inhalation starts at the diaphragm, but the first sensor to measure the end of inhalation is the SSP, located above the trachea. The Pes is almost synchronized to SSP before the abdomen ends its expansion, and the rib-cage is the latest one to be activated. A higher spread of the end inhalation distance was associated with a generally slightly lower F1-score performance.

The detection of breaths allowed a direct estimation of breathing rate over the entire night, which can be an index as important as counting respiratory events in the determination of cardiovascular risks (Baumert *et al* 2019). The SSP sensor achieved an average percentage discrepancy around 4.25% and an absolute discrepancy around 2.7 bpm, both comparable with RIP belts internal comparison, but under-performed them in the Pes subset. Nevertheless, the SSP obtained a higher breathing rate coverage than RIP when compared with Pes. We must remember that in our specific dataset the reference signals are not indicative of

an exact ground-truth, so a low MAPE indicates high agreement between estimates rather than an absolute measure of the error.

4.3. Patterns in respiratory effort

We extracted different features from each breath and analyzed them over patterns of 5 consecutive breaths. In our analysis, we chose 5 breaths to focus on the average length of respiratory events. In broader research settings, the patterns of other physiological phenomena of interest may happen at shorter time-scales (e.g. apneas lasting less than 10 s or 2–3 breaths) or longer, such as prolonged partial obstructions (longest sequence with increasing effort) or RERAs. In all indexes considered, the SSP performance was comparable with RIP belts. However, the lower performance in breath detection against Pes reflected in a lower number of sequences correctly identified. Also the indexes relative to the breaths' inhalation area were slightly lower versus RIP. Overall, all tested sensors followed the same direction of Pes in increasing or decreasing patterns. This result opens new paths in the automatic detection of certain respiratory events, such as RERA, or a better characterization of hypopneas (compared to RIP belts). Differences arose between the sequence-by-sequence correlation and how relative change of patterns correlate over the whole night.

The lower correlation in relative change could be explained through sensors' characteristics. The RIP belts are known to shift during the night, particularly if a participant moves in his/her sleep. The SSP signal may suffer from drastic amplitude shifts, as we have seen in figure 4. In the same way, the Pes signal itself was not immune to noise and abrupt changes (in accordance with current literature), particularly in sleep studies where the catheter is not fixed in-place with a balloon, making it a moving target rather than a fixed, stable reference signal. All these small discrepancies may accumulate during the night and degrade the whole performance. Limiting the correlation to short sequences partially hid this problem, but it remains necessary to understand what are the longest sequences that can be correlated reliably, and how they could be used in clinical decisions.

The inhale/exhale ratio metric also shown the lowest correlations if compared with amplitude and area. This can be explained looking at the example in figure 3. Generally the SSP and Pes signals have an almost flat top followed by a decrease in intrathoracic pressure. This means that amplitude and area were less influenced by errors in $peaks_{start}$, while the inhale/exhale ratio was. Improving the quality of the signals better modeling of breaths' fiducial points in SSP and Pes can improve these results.

4.4. Sleep-related effects

We wanted to assess if the detection of breathing and respiratory patterns were reliable over the whole night. Depending on the metrics of interest, we observed that the sleep context, namely the presence of respiratory events, sleep stages and body position may have some influences.

In some cases, the reduction in performance was expected for sleep-related states that may introduce noise in the signal (e.g. arousals and body movements). Here the reference signals may be influenced by the same factors, with a combined reduction of the performance that does not depend only on the SSP itself, and some technical solutions are foreseeable. The combination other signal sources, for instance measuring capsule movements with inertial sensors and integrating them in the pre-processing could compensate or account for these errors. The same solutions may be applicable also to respiratory belts and esophageal catheters.

Similarly, respiratory events and the return to normal breathing after an event may be associated to noise in the signals, and we measured an effect on the detection of breaths. Interestingly, the SSP maintained a high sensitivity with Pes during obstructive apneas, but not against RIP belts. The reduction in sensitivity could be dependent on noisy signals in RIP, but as we observed in figure 6, it could be that some breaths may be visible in certain signals, but not in others. Clinicians routinely employ multiple signals to discriminate ambiguous events, so it might be not possible to design a *one-sensor-fits-all* detection method for respiratory events. We also observed that the amplitude of respiratory effort seemed to be more correlated during the events compared to normal breathing. The current focus is on respiratory events scored with AASM rules, while the so-called *Pes events* are ignored (described by Guilleminault and Chowdhuri 2000). In future studies, we will consider electroencephalography (EEG) or autonomic arousals extracted from other PSG sensors to examine if those events may show a different relationship between SSP and Pes. Among pressure sensors, the discrepancy between SSP and Pes may also depend on obstruction happening in the airway space between the two sensors (potentially explaining missing breaths in SSP but not Pes in figure 6). However, we cannot verify this with the data available to us. At the moment, we can only hypothesize the existence of an upper boundary of agreement between respiratory sensors (even in presence of signals without noise), but not how high it may be and how a larger disagreement may influence clinical processes. Another open field of study may regard the distinction of hypopneas of obstructive or central origin, and with or without airway collapse

(e.g. Mooney *et al* 2012) which currently are not discernible with RIP and airflow signals alone, but may have distinct patterns in SSP.

The progression of sleep stages shown a positive, although small, effect on the detection of breaths. The most likely causes may be a reduction of body movements in deeper stages, or less respiratory events in non-REM sleep stages and certain sleeping positions. We opted to not define a model evaluating the effect of all factors together because it was not the main focus of this research, but future studies may provide a better assessment of it.

The depth of sleep also exerted different effects in the correlation of respiratory effort patterns. Where the SSP signal remained relatively consistent with Pes, the RIP belts did not. From the data in figure 10 we can hypothesize that the amplitude of breaths in the SSP may be less influenced by sleep stages, compared to RIP belts' amplitude that evolve in opposite directions. Oversimplifying, we can assume that RIP belts measure the activity of single anatomical compartments, which reflects on a simpler morphology. That is why certain respiratory events, such as paradoxical breathing, need both belts to be described appropriately. Conversely, the morphology of each breath in the pressure signals is the combination of multiple elements. On the one hand, one single pressure signal provides a richer view of overall respiratory dynamics. On the other hand, it is difficult to know the exact contribution of abdominal and thoracic expansion, the effect of occlusions in the upper airway, the breathing rate, and other physiological mechanisms on the pressure signals. Earlier studies tested how abdominal and thoracic breathing patterns may influence the Pes signal, but only in relation to multiple sensing locations in the esophagus (Irvin *et al* 1984). Other researchers analyzed how the body position influences the relation between esophageal and pleural pressure, with applications in assisted ventilation (Pasticci *et al* 2020). Other than these, the field of research around the components of Pes signal (and by proxy, the SSP) remain open and largely unexplored. Future studies may be able to tackle this interesting problem. Overall, we can state that the SSP proved to be a good alternative to Pes in the detection of breaths and tracking of respiratory patterns are consistent over the whole night.

4.5. Effect of cardiogenic oscillations and breath detection spread

Another open research problem is the proper management of cardiogenic oscillations in pressure signals. The filtering method presented here relies on a clean ECG signal and precise heart rate estimation, which is not always possible. If an ECG signal is not present, the heart rate can be estimated from other cardiac signals, such as photoplethysmography (PPG), or from the SSP signal itself (Cerina *et al* 2023). Other works in Pes literature employed adaptive filters (Schuessler *et al* 1998), singular spectrum analysis (Mukhopadhyay *et al* 2020) or template subtraction (Graßhoff *et al* 2017). In all cases, the filtering cannot prevent small shifts in the fiducial points of the Pes and SSP signal caused by cardiogenic oscillations (as visible in figure 2), with consequences on all the features of breaths' morphology that we analyzed. An improved detection method, based on better models of single breaths in SSP and Pes, may help in correcting the position of detected $peaks_{end}$ and $peaks_{start}$. Another future development could be using an additional sensor that is not sensitive to respiratory swings, but only cardiogenic oscillations (e.g. placing the sensor in a different position on the body) in combination with a source separation algorithm. Additionally, cardiogenic oscillations are the result of multiple overlapping factors, such as structure of the upper airway, heart, and pulmonary arteries, body composition between the trachea and the SSP sensor, or cardio-respiratory coupling mechanisms and could represent important physiological mechanisms that are currently under-studied.

Other than cardiogenic oscillations, we noticed that the overall localization spread of $peaks_{end}$ and $peaks_{start}$ can influence how well respiratory sensors are correlated. Generally, the spread was higher in $peaks_{start}$ rather than $peaks_{end}$, including the spread between the RIP belts. A higher spread between the end inhalation points of two sensors was associated to a lower F1-score in breaths detection and breathing rate estimation. Our hypotheses for the causes of spread may be uncontrolled noise sources (such as varying tension in the belts), but also physiological conditions, including variations in the action of respiratory muscles across sleep stages or the breathing rate. More importantly, the spread can directly affect the calculation of breaths' features such as amplitude, area, and inhale/exhale ratio. The minimization of the spread caused by noise or localization issues will unveil its physiological cause, and future studies could examine potential applications of analyzing the natural asynchronies of respiratory sensors (e.g. asynchronies in children Bronstein *et al* 2018).

4.6. Demographic factors

The SSP sensor demonstrated to be reliable irrespective of sex, age, and number of respiratory events during the night. The SSP proved to operate as a reliable alternative to RIP also in children under the age of 18. This result opens a new line of research to verify if results hold for Pes. It was proven that esophageal pressure swings are useful in the analysis of pediatric sleep disorders, but the catheters were also poorly tolerated (accepted by less than 65% Chervin *et al* 2012).

Considering the F1-score in breath detection, we observed significantly worse performance of the SSP signal against thoracic RIP belt compared to SSP against abdominal ones. Grouping by sex, the difference remained significant for females and not for males. We cannot prove with the data available to us if the difference is caused by physiological or physical mechanisms, and it remains an open problem in literature. Despite known differences in airway mechanics (Pillar *et al* 2000), biological sex is not considered as a factor in the design of RIP belts (Montazeri *et al* 2021). By its own design, the SSP sensor could be less influenced by sex differences, but other studies are necessary to examine the variety of body characteristics that may affect the quality of the recording.

The usage of a clinical dataset also presents significant differences with designed and well-controlled laboratory settings, with both pros (such as designing more robust methods) and cons. For example, the Pes subset contains very few cases of apneic events, with a large predominance of hypopneas. This is a consequence of the clinical process, by which Pes is employed if suspected respiratory events are not evident in other respiratory sensors. This limitation combined with fewer valid breath sequences (table 6) leads to a very small number of sequences with mixed and central apneas (table 8). For this reason, it is not possible to make statistically robust conclusions on the behavior of SSP versus RIP belts during these respiratory events. Even with these limitations, this is the first work on SSP analysis (to the best of our knowledge) where a population using the Pes sensor could be compared with a matched subset without it, giving us some insights on the effect that Pes may have on the other respiratory sensors. At the moment, we can only assume the same contraindications of Pes presented in previous literature, with potential disturbances on sleep and breathing caused by the presence of the catheter in the airway (Kushida *et al* 2002).

5. Conclusions

In this paper, we present an extensive comparison of the signal from a suprasternal notch pressure (SSP) sensor against current clinical standard for the measurement of respiration and respiratory effort during sleep. While under-performing in some situations, the performance of the SSP is comparable with respiratory inductance plethysmography (RIP) belts and a promising alternative to esophageal catheters. Despite limitations, we believe that the SSP signal is closer to the esophageal pressure signal and may be a better surrogate of respiratory effort than RIP belts. The discrepancy in performance may be caused by conceptual differences between the signals. This, however, needs to be examined with more controlled experiments and could be resolved by improving some minor technical issues in the application of the SSP sensor. We believe that a variety of details regarding the analysis of respiratory dynamics with different sensors and the underlying physiology are still unexplored, and that sleep research may benefit from a shift in focus to better understand physiological mechanisms before pathological ones.

Data availability statement

The SOMNIA data used in this study are available from the Sleep Medicine Centre Kempenhaeghe upon reasonable request. The data can be requested by presenting a scientific research question and by fulfilling all the regulations concerning the sharing of the human data. The details of the agreement will depend on the purpose of the data request and the entity that is requesting the data (e.g. research institute or corporate). Each request will be evaluated by the Kempenhaeghe Research Board and, depending on the request, approval from independent medical ethical committee might be required. Access to data from outside the European Union will further depend on the expected duration of the activity; due to the work required from a regulatory point of view, the data is less suitable for activities that are time critical, or require access in short notice. Specific restrictions apply to the availability of the data collected with sensors not comprised in the standard PSG set-up, since these sensors are used under license and are not publicly available. These data may however be available from the authors with permission of the licensors. For inquiries regarding availability, please contact Merel van Gilst (M.M.v.Gilst@tue.nl). The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

This work was performed within the IMPULS framework of the Eindhoven MedTech Innovation Center (e/MTIC, incorporating Eindhoven University of Technology, Philips Research, and Sleep Medicine Center Kempenhaeghe), including a PPS-supplement from Dutch Ministry of Economic Affairs and Climate Policy. Additional support by STW/IWT in the context of the OSA+ Project (No. 14619).

ORCID iDs

Luca Cerina  <https://orcid.org/0000-0001-8571-3971>

Gabriele B Papini  <https://orcid.org/0000-0002-5752-9226>

Pedro Fonseca  <https://orcid.org/0000-0003-2932-6402>

References

- Amaddeo A, Sabil A, Arroyo J O, De Sanctis L, Griffon L, Baffet G, Khirani S and Fauroux B 2020 Tracheal sounds for the scoring of sleep respiratory events in children *J. Clin. Sleep Med.* **16** 361–9
- Anttalainen U, Tenhunen M, Rimpilä V, Polo O, Rauhala E, Himanen S L and Saaresranta T 2016 Prolonged partial upper airway obstruction during sleep—an underdiagnosed phenotype of sleep-disordered breathing *Eur. Clin. Respir. J.* **3** 1–10
- Baumert M, Linz D, Stone K, McEvoy R D, Cummings S, Redline S, Mehra R and Immanuel S 2019 Mean nocturnal respiratory rate predicts cardiovascular and all-cause mortality in community-dwelling older men and women *Eur. Respir. J.* **54** 1802175
- Berry R B et al 2012 Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events: deliberations of the sleep apnea definitions task force of the American Academy of Sleep Medicine *J. Clin. Sleep Med.* **8** 597–619
- Brochard L J 2014 Measurement of esophageal pressure at bedside: pros and cons *Curr. Opin. Crit. Care* **20** 39–46
- Bronstein J Z, Xie L, Shaffer T H, Chidekel A and Heinle R 2018 Quantitative analysis of thoracoabdominal asynchrony in pediatric polysomnography *J. Clin. Sleep Med.* **14** 1169–76
- Cerina L, Papini G B, Fonseca P, Overeem S, van Dijk J P and Vullings R 2023 Extraction of cardiac-related signals from a suprasternal pressure sensor during sleep *Physiol. Meas.* **44** 1–21
- Charlton P H, Kotzen K, Mejía-Mejía E, Aston P J, Budidha K, Mant J, Pettit C, Behar J A and Kyriacou P A 2022 Detecting beats in the photoplethysmogram: benchmarking open-source algorithms *Physiol. Meas.* **43** 1–20
- Chervin R D, Ruzicka D L, Hoban T F, Fetterolf J L, Garetz S L, Guire K E, Dillon J E, Felt B T, Hodges E K and Giordani B J 2012 Esophageal pressures, polysomnography and neurobehavioral outcomes of adenotonsillectomy in children *Chest* **142** 101–10
- Chokroverty S 2017 Physiological changes of sleep *Sleep Disorders Medicine: Basic Science, Technical Considerations and Clinical Aspects* (Springer) pp 153–94
- Dempsey J A, Veasey S C, Morgan B J and O'Donnell C P 2010 Pathophysiology of sleep apnea *Physiol. Rev.* **90** 47–112
- Devani N, Pramono R X A, Imtiaz S A, Bowyer S, Rodriguez-Villegas E and Mandal S 2021 Accuracy and usability of AcuPebble SA100 for automated diagnosis of obstructive sleep apnoea in the home environment setting: an evaluation study *BMJ Open* **11** e046803
- Ferber R et al 1994 ASDA standards of practice: portable recording in the assessment of obstructive sleep apnea *Sleep-Lawrence* **17** 378
- Flemons W W et al 1999 Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research *Sleep* **22** 667–89
- Glos M, Sabil A, Jelavic K S, Schöbel C, Fietze I and Penzel T 2018 Characterization of respiratory events in obstructive sleep apnea using suprasternal pressure monitoring *J. Clin. Sleep Med.* **14** 359–69
- Graßhoff J, Petersen E, Eger M, Bellani G and Rostalski P 2017 A template subtraction method for the removal of cardiogenic oscillations on esophageal pressure signals 2017 39th Annual Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBC) (IEEE) pp 2235–8
- Guilleminault C and Chowdhuri S 2000 Upper airway resistance syndrome is a distinct syndrome *Am. J. Respir. Crit. Care Med.* **161** 1412–3
- Hammer J and Newth C 2009 Assessment of thoraco-abdominal asynchrony *Paediat. Respir. Rev.* **10** 75–80
- Heinzer R et al 2015 Prevalence of sleep-disordered breathing in the general population: the HypnoLaus study *Lancet* **3** 310–8
- Iglewicz B and Hoaglin D C 1993 *How to Detect and Handle Outliers* vol 16 (Quality Press)
- Irvin C, Sampson M, Engel L and Grassino A 1984 Effect of breathing pattern on esophageal pressure gradients in humans *J. Appl. Physiol.* **57** 168–75
- Kushida C A, Giacomini A, Lee M K, Guilleminault C and Dement W C 2002 Technical protocol for the use of esophageal manometry in the diagnosis of sleep-related breathing disorders *Sleep Med.* **3** 163–73
- Lévy P, Kohler M, McNicholas W T, Barbé F, McEvoy R D, Somers V K, Lavie L and Pépin J L 2015 Obstructive sleep apnoea syndrome *Nat. Rev. Dis. Primers* **1** 1–21
- Mann E A et al 2020 Study design considerations for sleep-disordered breathing devices *J. Clin. Sleep Med.* **16** 441–9
- Montazeri K, Jonsson S A, Agustsson J S, Serwatko M, Gislason T and Arnardottir E S 2021 The design of rip belts impacts the reliability and quality of the measured respiratory signals *Sleep Breath.* **25** 1–7
- Mooney A M, Abounasr K K, Rapoport D M and Ayappa I 2012 Relative prolongation of inspiratory time predicts high versus low resistance categorization of hypopneas *J. Clin. Sleep Med.* **8** 177–85
- Mukhopadhyay S K, Zara M, Telias I, Chen L, Coudroy R, Yoshida T, Brochard L and Krishnan S 2020 A singular spectrum analysis-based data-driven technique for the removal of cardiogenic oscillations in esophageal pressure signals *IEEE J. Trans. Eng. Health Med.* **8** 1–11
- Pasticci I et al 2020 Determinants of the esophageal-pleural pressure relationship in humans *J. Appl. Physiol.* **128** 78–86
- Pillar G, Malhotra A, Fogel R, Beauregard J, Schnall R P and White D P 2000 Airway mechanics and ventilation in response to resistive loading during sleep: influence of gender *Am. J. Respir. Crit. Care Med.* **162** 1627–32
- Redline S et al 2007 The scoring of respiratory events in sleep: reliability and validity *J. Clin. Sleep Med.* **3** 169–200
- Rodrigues T, Samoutphonh S, Silva H and Fred A 2021 A low-complexity R-peak detection algorithm with adaptive thresholding for wearable devices 2020 25th Int. Conf. on Pattern Recognition (ICPR) (IEEE) pp 1–8
- Sabil A, Glos M, Günther A, Schöbel C, Veauthier C, Fietze I and Penzel T 2019 Comparison of apnea detection using oronasal thermal airflow sensor, nasal pressure transducer, respiratory inductance plethysmography and tracheal sound sensor *J. Clin. Sleep Med.* **15** 285–92
- Scholkmann F, Boss J and Wolf M 2012 An efficient algorithm for automatic peak detection in noisy periodic and quasi-periodic signals *Algorithms* **5** 588–603
- Schuessler T F, Gottfried S B, Goldberg P, Kearney R E and Bates J H 1998 An adaptive filter to reduce cardiogenic oscillations on esophageal pressure signals *Ann. Biomed. Eng.* **26** 260–7

- Sforza E, Krieger J and Petiau C 1998 Nocturnal evolution of respiratory effort in obstructive sleep apnoea syndrome: influence on arousal threshold *Eur. Respir. J.* **12** 1257–63
- Staats B A, Bonekat H W, Harris C D and Offord K P 1984 Chest wall motion in sleep apnea *Am. Rev. Respir. Dis.* **130** 59–63
- Suarez-Sipmann F, Santos A, Peces-Barba G, Bohm S H, Gracia J L, Calderón P and Tusman G 2012 Pulmonary artery pulsatility is the main cause of cardiogenic oscillations *J. Clin. Monitor. Comput.* **27** 47–53
- Sweetman A, Lack L and Bastien C 2019 Co-morbid insomnia and sleep apnea (COMISA): prevalence, consequences, methodological considerations and recent randomized controlled trials *Brain Sci.* **9** 371
- Székely G J, Rizzo M L and Bakirov N K 2007 Measuring and testing dependence by correlation of distances *Ann. Stat.* **35** 2769–94
- van Gilst M M *et al* 2019 Protocol of the SOMNIA project: an observational study to create a neurophysiological database for advanced clinical sleep monitoring *BMJ Open* **9** e030996
- Zimmerman P, Connellan S, Middleton H, Tabona M, Goldman M and Pride N 1983 Postural changes in rib cage and abdominal volume-motion coefficients and their effect on the calibration of a respiratory inductance plethysmograph *Am. Rev. Respir. Dis.* **127** 209–14