Journal Pre-proof

Performances of Auto-CPAP devices under real-life leak patterns: A high-fidelity modeling approach

Mathéo Richard Emeline Fresnel Jean-Pierre Mallet Raphael Gilson Adrien Kerfourn Maxime Patout Carole Philippe Dany Jaffuel Marius Lebret

Performances of Auto-CPAP devices under real-life leak patterns: A high-fidelity modeling approach

Mathéo Richard[1], Emeline Fresnel[1], Jean-Pierre Mallet[2,3], Raphael Gilson[4], Adrien Kerfourn[1], Maxime Patout[5,6], Carole Philippe[5], Dany Jaffuel[2,3] & Marius Lebret[1]

1 KerNel Biomedical, 18 Rue Marie Curie Bâtiment ANIDER, Rouen, France
2 Department of Respiratory Diseases, CHU Montpellier, Montpellier, France
3 PhyMedExp, University of Montpellier, INSERM U1046, CNRS UMR 9214, Montpellier, France
4 Groupe ADENE, Rue de Chambert, Parc Euromédecine 2, Montpellier
5 AP-HP, Groupe Hospitalier Universitaire APHP-Sorbonne Université, site Pitié-Salpêtrière, Service des Pathologies du Sommeil (Département R3S), Paris, France
6 Sorbonne Université, INSERM, UMRS1158 Neurophysiologie Respiratoire Expérimentale et Clinique, France
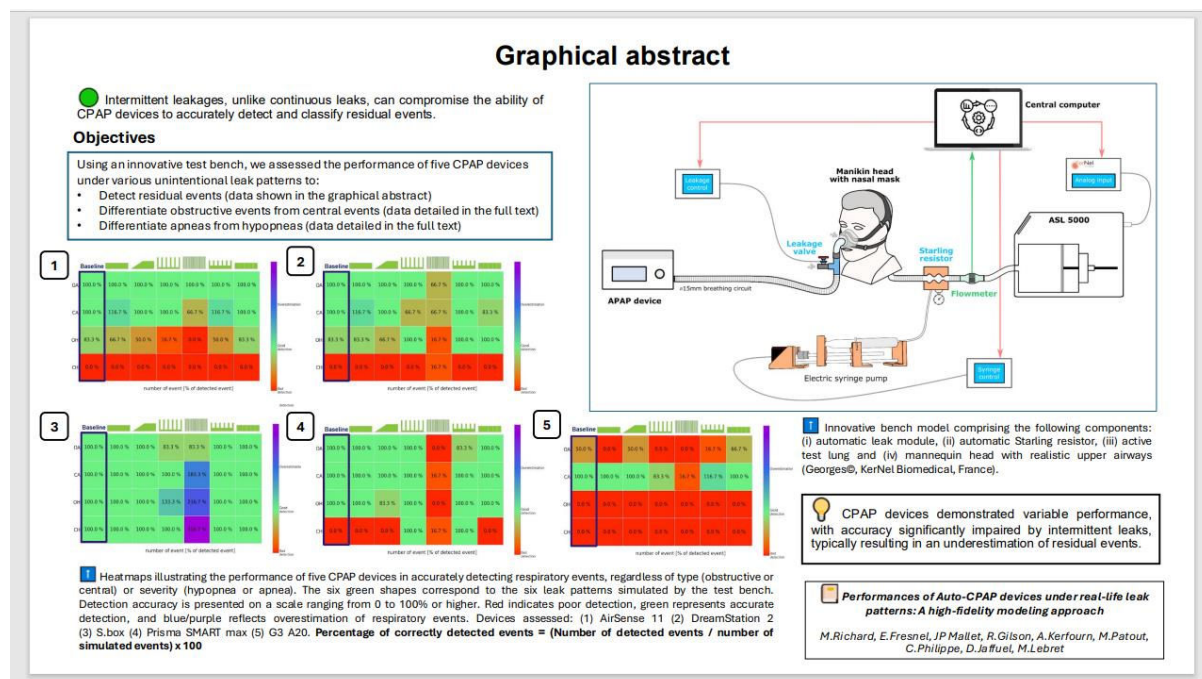
Corresponding author:
Marius Lebret
Kernel Biomedical
18 rue Marie Curie, 76000 Rouen (France)
email: mariuslebret@gmail.com

Graphical Abstract

The algorithms used by CPAP machines to identify and characterize residual respiratory events vary between brands and between models from the same brand. Additionally, CPAP manufacturers employ diverse methods for estimating and reporting air leaks which complicates the analysis of patient data from CPAP built-in software reports (1). This issue is particularly relevant in the context of the growing role of telemonitoring in managing these patients (2). Moreover, while leak intensity - in liters per minute - is a critical parameter well-recognized for its significant impact on PAP device performance (2–4), the leak pattern itself is another often-underestimated aspect that can significantly affect PAP performance (5,6). In a previous publication (7), we proposed a nomenclature for leak patterns based on real-life polygraphic recordings. Building on this work (7) and subsequent studies (3,8–10), we have developed a sophisticated bench model that realistically emulates sleep apnea syndrome, capable of replicating a wide range of leak patterns, closed-loop upper-airway obstruction, and ventilatory efforts. Our primary aim was to assess the performance of CPAP devices to accurately detect the respiratory events in the context of various unintentional leak patterns. Our secondary aims were to assess the performance of CPAP devices to classify the nature of the respiratory events (obstructive/central) and their severity (hypopnea/apnea).

The sophisticated bench model used a ASL 5000 (Ingmar Medical, Pittsburgh, PA, USA) and was able to simulate the patient's active breathing (figure 1). We used a mannequin head with realistic upper airways (Georges©, KerNel Biomedical, France) as the interface between the CPAP device and the ASL 5000. An automatic Starling resistor was used to simulate airway obstructions. The bench setup was engineered to simulate obstructive apneas for 25 seconds and obstructive hypopneas for 50 seconds. The ASL5000 was used to simulate central apneas with a gradual decrease and resumption of respiratory effort. Four main respiratory scripts were created for each type of event: obstructive apneas, central apneas, obstructive hypopneas, and central hypopneas. Each script was played in association with each of the six previously described unintentional leak patterns (7), as well as without leakage (baseline condition). 5 CPAP devices all set in continuous pressure mode at 10 cmH2O were evaluated: (1) ResMed AirSense 11, (2) Philips DreamStation 2, (3) Sefam S.box, (4) Lowenstein Prisma SMART max, and (5) BMC G3 A20. Detailed methods are available in the supplementary material.

Performance metrics of the 5 devices are presented in Figures 2, e-figure 4, e-figure 5, e-figure 6; e-figure 7 and e-figure 8. E-figure 4 illustrates the six leakage patterns used in the simulations along with the detailed leak flow data recorded by each device on the test bench. In the absence of unintentional leaks, the detection capability varied from one device to another. Four out of 5 devices detected 100% of apneas, regardless of their origin (devices 1, 2, 3 and 4). Only device 3 was able to detect central hypopneas in the absence of unintentional leak (figure 2). Overall, leak pattern number 4, characterized by a pure intermittent leak pattern, led to a significant and systematic decrease in the devices' ability to accurately identify and classify respiratory events. This pattern led predominantly to an underestimation of the number of events in devices 1, 2, 4, and 5. It led to an overestimation in device 3. When central hypopnea was generated, device 3 was still able to detect the events in the presence of leakage, while device 4 partially detected them when the UL was intermittent, likely because it interpreted intermittent leakage events as central hypopnea events. When comparing the mean residual Apnea Hypopnea Index (AHI) reported by built-in software under baseline conditions (no leaks) to that reported in the presence of intermittent leak patterns, the residual AHI was underestimated by an average of 33.8 ± 11.2% (p<0.001).

Similarly, with continuous leak patterns, the AHI was underestimated by an average of $5.1 \pm 4.5\%$ ($p<0.001$) compared to baseline. Details on the ability of the 5 devices to characterize the nature and severity of the respiratory events are presented in supplementary material.

We demonstrated that CPAP devices exhibited inconsistent performance in event detection and that their accuracy was significantly compromised by leak patterns. As previously described (11,12), our bench test study corroborates that devices generally succeeded in identifying simulated obstructive apnea events but encounters challenges with simulated obstructive hypopnea events. The failure of some CPAP devices to accurately identify hypopneas can be attributed to several factors: i) manufacturer-specific definitions of the hypopnea threshold (13,14) or the choice to exclude non-obstructive hypopneas from detection (device 1), ii) the use of linearized flow rather than raw flow to score events (15), or iii) the presence of leaks (13,15). In the case of leaks, depending on the type of simulated leak patterns, event detection can be intentionally suppressed by some devices to reduce false positives, which may explain the reduced performance of certain machines in detecting events with inhomogeneous leak patterns. These CPAP difficulties in identifying hypopneas are further accentuated in real-world when comparing the AHI measured by CPAP to that measured by polysomnographers according to the American Academy of Sleep Medicine (AASM) definition (16) (29). Consistent with previous studies comparing AHI measured by the CPAP (AHI_{flow}) versus AHI measured by polysomnography (AHI_{PSG}) (17–19), the correlation between device-detected and manually scored breathing events is higher for apneas than for hypopneas.

A recent meta-analysis (16) revealed greater dispersion around the pooled mean hypopnea index bias compared to the AHI bias, supporting the idea that there is significant variability in how different devices detect hypopnea events. It is crucial to note that the studies included in this meta-analysis do not address the straightforward questions faced by clinicians, since they do not clarify whether patients with central sleep apnea and/or significant CPAP leakage (i.e., substantial leakage leading to inaccurate AHI measurement and/or counteracting obstructive respiratory events) were excluded. Given the difficulties CPAP devices have in identifying the central nature of hypopneas, some manufacturers' decision to exclude these central events from the AHI_{flow} calculation may lead clinicians to

underestimate the residual AHI_{flow}. Some concerns may arise from this choice, such as limiting the clinician to detect treatment-emergent central sleep apnea via telemedicine and CPAP monitoring program.

Our results show that some devices consistently misclassify central apneas as obstructive apneas. In auto-adjusting mode, we can assume that this misclassification would likely lead to an inappropriate increase in auto-adjusting pressure. The observed variability in device performance is likely to arise from the inherent functioning of the algorithms, including thresholds for flow or tidal volume or the different technologies used by the devices to distinguish between obstructive and central respiratory events. It's important to note that even when two devices use the same technology, their algorithms can still differ significantly in how they operate.

In this study, we opted to perform the tests in fixed mode to focus on evaluating the impact of continuous versus intermittent leaks on the device's performances to detect and classify respiratory events. Testing the devices' ability to adjust pressure in response to these events was not within the scope of this work and will be addressed in future research. Furthermore, allowing the devices to freely adjust pressure in automatic mode would have introduced bias, as our test bench is designed to normalize simulated respiratory events with adequate positive pressure, potentially hindering the primary objective of this study.

This study has several strengths and limitations. As a bench modeling study, it benefits from standardized procedures and controlled testing conditions, allowing us to evaluate devices under scenarios we precisely define. However, bench data may lack the complexity and nuances observed in in vivo settings due to the limited range of test conditions. Despite this, our innovative model closely reflects real-life scenarios. First, the leakage patterns simulated in this study were archetypal, derived from overnight CPAP polysomnography, and previously published (7). In simulating leakage patterns, we couldn't maintain the exact 90th percentile of leakage across all patterns: by nature, each pattern has a unique leakage distribution, even though the total leakage volume remains consistent. While average and total leakage are held constant for each pattern, the peaks and lows differ due to each pattern's specific distribution characteristics. Second, in modeling respiratory events, we ensured that at least one of the five devices tested could detect and classify these events, thereby validating the recognizability of our simulations. We applied a strict threshold of 50%

flow reduction and an event duration of 50 seconds—more stringent than the AASM hypopnea criteria. This deliberate choice enhanced detectability and minimized false negatives, highlighting the challenges devices face in event detection. Importantly, the study's primary focus was to assess how varying leak patterns affect CPAP performance in event detection, rather than to evaluate the intrinsic scoring accuracy of the devices themselves. Third, our model successfully replicated true ventilatory recoveries with increased post-event respiratory drive, adding physiological realism to the simulations. Furthermore, we characterized obstructive events with a U-shaped flow pattern, effectively mimicking real-life obstructive respiratory events.

The 2019 AASM guidelines recommended starting therapy with auto-CPAP for adults without comorbidities (20). Additionally, telemonitoring-guided interventions were suggested during the initial phase of therapy. Accurate detection of residual respiratory events by devices is crucial in this context. Our study demonstrated that intermittent leaks significantly reduced CPAP accuracy, often leading to an underestimation of residual events. This highlights the need for caution when interpreting residual AHI data from built-in software or telemonitoring systems. However, clinical studies are necessary to assess CPAP performance in comparison to polysomnography under real-life intermittent leak patterns

Artificial intelligence involvement: The authors confirm that no part of this material was produced with the assistance of artificial intelligence software or tools.

Authors' Contributions:

Literature research: MR, DF, ML, EF, DJ; Study concept and design: MR, DF, ML, EF, DJ; Acquisition of data: ML, EF, MR; Analysis and interpretation of data: ML, MR, EF, DJ; Statistical analysis: MR, EF; Drafting of the manuscript: ML, DJ, RG, JPM, CP ; Critical revision of the manuscript for important intellectual content: ML, EF, CP, DJ.

ML and EF had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis

References

1. Schwab RJ, Badr SM, Epstein LJ, Gay PC, Gozal D, Kohler M, et al. An Official American Thoracic Society Statement: Continuous Positive Airway Pressure Adherence Tracking Systems. The Optimal Monitoring Strategies and Outcome Measures in Adults. Am J Respir Crit Care Med 2013;188:613- 20.
2. Zhu K, Rabec C, Gonzalez-Bermejo J, Hardy S, Aouf S, Escourrou P, et al. Combined effects of leaks, respiratory system properties and upper airway patency on the performance of home ventilators: a bench study. BMC Pulm Med 2017;17:145.
3. Lebret M, Fresnel E, Prouvez N, Zhu K, Kerfourn A, Richard J-C, et al. Responses of Bilevel Ventilators to Unintentional Leak: A Bench Study. Healthc Basel Switz 2022;10:2416.
4. Sogo A, Montanyà J, Monsó E, Blanch L, Pomares X, Lujàn M. Effect of dynamic random leaks on the monitoring accuracy of home mechanical ventilators: a bench study. BMC Pulm Med 2013;13:75.
5. Luján M, Sogo A, Pomares X, Monsó E, Sales B, Blanch L. Effect of leak and breathing pattern on the accuracy of tidal volume estimation by commercial home ventilators: a bench study. Respir Care 2013;58:770- 7.
6. Luján M, Sogo A, Grimau C, Pomares X, Blanch L, Monsó E. Influence of dynamic leaks in volume-targeted pressure support noninvasive ventilation: a bench study. Respir Care 2015;60:191- 200.
7. Richard M, Jaffuel D, Mallet J-P, Gilson R, Patout M, Fresnel E, et al. General nomenclature of air leak patterns: A proposal. Sleep Med 2023;112:188- 90.
8. Fresnel E, Muir J-F, Letellier C. Realistic human muscle pressure for driving a mechanical lung. EPJ Nonlinear Biomed Phys 2014;2:1- 18.
9. Fresnel E, Caillard C, Lebret M, Razakamanantsoa L, Kerfourn A, Dupuis J, et al. Comparison of oronasal and nasal mask in home mechanical ventilation: observational cohort and bench study. Eur Respir J 2024:2302010.
10. Jolly G, Razakamanantsoa L, Fresnel E, Gharsallaoui Z, Cuvelier A, Patout M. Defining successful non-invasive ventilation initiation: Data from a real-life cohort. Respirol Carlton Vic 2021;26:1067- 75.
11. Fasquel L, Yazdani P, Zaugg C, Barras A, Michotte J-B, Correvon N, et al. Impact of Unintentional Air Leaks on Automatic Positive Airway Pressure Device Performance in Simulated Sleep Apnea Events. Respir Care 2023;68:31- 7.
12. Zhu K, Roisman G, Aouf S, Escourrou P. All APAPs Are Not Equivalent for the Treatment of Sleep Disordered Breathing: A Bench Evaluation of Eleven Commercially Available Devices. J Clin Sleep Med JCSM Off Publ Am Acad Sleep Med 2015;11:725- 34.
13. Bertelli F, Suehs CM, Mallet JP, Rotty MC, Pepin JL, Gagnadoux F, et al. Apnoea-hypopnoea indices determined via continuous positive airway pressure (AHI-CPAPflow) versus those determined by polysomnography (AHI-PSGgold): a protocol for a systematic review and meta-analysis. BMJ Open 2021;11:e044499.
14. Johnson KG, Johnson DC. Treatment of sleep-disordered breathing with positive airway pressure devices: technology update. Med Devices Auckl NZ 2015;8:425- 37.
15. Jaffuel D, Mallet J-P, Sabil A. Accuracy of continuous positive airway pressure devices: the devil is in the details, the best is yet to come. Sleep Breath Schlaf Atm 2023;27:1651- 4.
16. Iftikhar IH, BaHammam A, Jahrami H, Ioachimescu O. Accuracy of residual respiratory event detection by CPAPs: a meta-analysis. Sleep Breath Schlaf Atm 2023;27:1759- 68.
17. Nigro CA, González S, Arce A, Aragone MR, Nigro L. Accuracy of a novel auto-CPAP device to evaluate the residual apnea-hypopnea index in patients with obstructive sleep apnea. Sleep Breath 2015;19:569- 78.
18. Berry RB, Kushida CA, Kryger MH, Soto-Calderon H, Staley B, Kuna ST. Respiratory Event Detection by a Positive Airway Pressure Device. Sleep 2012;35:361- 7.
19. Ueno Y, Nakanishi N, Oto J, Imanaka H, Nishimura M. A bench study of the effects of

leak on ventilator performance during noninvasive ventilation. Respir Care 2011;56:1758- 64.

20. Patil SP, Ayappa IA, Caples SM, Kimoff RJ, Patel SR, Harrod CG. Treatment of Adult Obstructive Sleep Apnea with Positive Airway Pressure: An American Academy of Sleep Medicine Clinical Practice Guideline. J Clin Sleep Med JCSM Off Publ Am Acad Sleep Med 2019;15:335- 43.
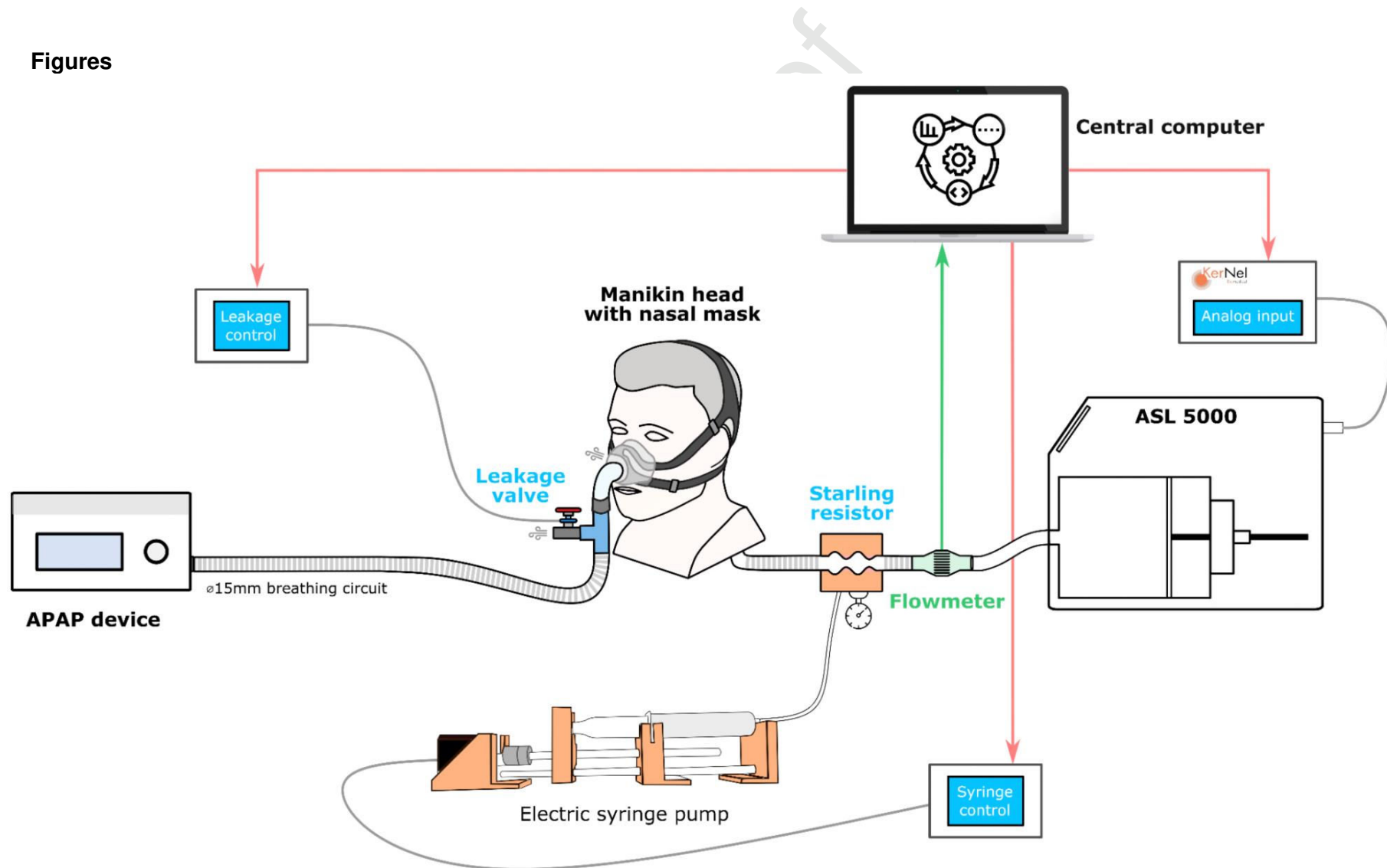
Figures

Figure 1. Comprehensive Bench Model including the following components: i. The leak module connected to its dedicated micro-computer. ii. The Starling resistor and the automatic electric syringe pump, linked to their dedicated micro-computer, together constituting the upper airway module. iii. The test lung (ASL 5000), connected to an analog module capable of varying muscular pressure and respiratory frequency. All three elements are managed by the central control computer, which executes the scripts and stores the data. The interface was placed on a mannequin head with realistic upper airways (Georges©, KerNel Biomedical, France).

Figure 2. Heatmaps illustrate the ability of the 5 CPAP devices to accurately detect the presence of respiratory events, regardless of their type (obstructive or central) or severity (hypopnea or apnea). The detection capability is represented on a scale from 0 to 100% or above. A 0% indicates that the device failed to detect any instances of a particular type of respiratory event with a specific leakage pattern The green shapes above each heatmap represent the leakage patterns tested during the study. A 100% indicates that the device detected all 6 instances of respiratory events. A score >100% indicates that the device detected more than 6 instances of respiratory events of any type or severity with a specific leakage pattern. For example, if a device detects 7 events instead of the actual 6, the subscore would be 116.7%. The red color indicates poor detection, the green color signifies accurate detection, and blue/purple represents an overestimation of respiratory events. Device 1: AirSense 11; device 2: DreamStation 2, Device 3: S.box, Device 4: Prisma SMART max and device 5: G3 A20. Obstructive apnea (OA); Central apnea (CA); Obstructive hypopnea (OH) and Central hypopnea (CH).

Ethics in publishing

1. Does your research involve experimentation on animals?:

No

2. Does your study include human subjects?:

No

3. Does your study include a clinical trial?:

No

4. Are all data shown in the figures and tables also shown in the text of the Results section and discussed in the Conclusions?:

Yes