

# UC San Diego

## UC San Diego Previously Published Works

### Title

Remote Monitoring of Positive Airway Pressure Data Challenges, Pitfalls, and Strategies to Consider for Optimal Data Science Applications

### Permalink

<https://escholarship.org/uc/item/7f89127q>

### Journal

CHEST Journal, 163(5)

### ISSN

0012-3692

### Authors

Bottaz-Bosson, Guillaume

Midelet, Alphanie

Mendelson, Monique

et al.

### Publication Date

2023-05-01

### DOI

10.1016/j.chest.2022.11.034

Peer reviewed

# Remote Monitoring of Positive Airway Pressure Data



## Challenges, Pitfalls, and Strategies to Consider for Optimal Data Science Applications

Guillaume Bottaz-Bosson, PhD; Alphanie Midelet, PhD; Monique Mendelson, PhD; Jean-Christian Borel, PhD; Jean-Benoît Martinot, PhD; Ronan Le Hy, PhD; Marie-Caroline Schaeffer, PhD; Adeline Samson, PhD; Agnès Hamon, PhD; Renaud Tamisier, PhD; Atul Malhotra, PhD; Jean-Louis Pépin, PhD; and Sébastien Bailly, PhD

Over recent years, positive airway pressure (PAP) remote monitoring has transformed the management of OSA and produced a large amount of data. Accumulated PAP data provide valuable and objective information regarding patient treatment adherence and efficiency. However, the majority of studies that have analyzed longitudinal PAP remote monitoring have summarized data trajectories in static and simplistic metrics for PAP adherence and the residual apnea-hypopnea index by the use of mean or median values. The aims of this article are to suggest directions for improving data cleaning and processing and to address major concerns for the following data science applications: (1) conditions for residual apnea-hypopnea index reliability, (2) lack of standardization of indicators provided by different PAP models, (3) missing values, and (4) consideration of treatment interruptions. To allow fair comparison among studies and to avoid biases in computation, PAP data processing and management should be conducted rigorously with these points in mind. PAP remote monitoring data contain a wealth of information that currently is underused in the field of sleep research. Improving the quality and standardizing data handling could facilitate data sharing among specialists worldwide and enable artificial intelligence strategies to be applied in the field of sleep apnea.

CHEST 2023; 163(5):1279-1291

**KEY WORDS:** data management; OSA; positive airway pressure; remote monitoring; time series

OSA is a highly prevalent chronic disease with nearly one billion adults aged 30 to 69 years affected worldwide.<sup>1</sup> OSA is a systemic disease that presents multiple clinical phenotypes<sup>2</sup> and is associated independently with cardiovascular comorbidities, decreased

quality of life, diminished neurocognitive function, and depression.<sup>3,4</sup> Positive airway pressure (PAP), the first-line therapy for moderate-to-severe OSA, is used in millions of people worldwide<sup>1</sup> and improves symptoms and quality of life and may

**ABBREVIATIONS:** PAP = positive airway pressure; rAHI = residual apnea-hypopnea index

**AFFILIATIONS:** From the Laboratoire HP2, U1300 Inserm (G. B.-B., A. Midelet, M. M., J.-C. B., R. T., J.-L. P., and S. B.), and the Jean Kuntzmann Laboratory (G. B.-B., A. S., and A. H.), University Grenoble Alpes, Grenoble, France; Probayes (A. M., R. L. H., and M.-C. S.), Montbonnot-Saint-Martin, France; the AGIR à dom HomeCare Charity (J.-C. B.), Meylan, France; the Sleep Laboratory (J.-B. M.), CHU UCL Namur Site Sainte-Elisabeth, Namur, Belgium; the Institute of Experimental and Clinical Research (J.-B. M.), UCL, Bruxelles

Woluwe, Belgium; and the Division of Pulmonary, Critical Care and Sleep Medicine (A. Malhotra), University of California San Diego, La Jolla, CA.

**CORRESPONDENCE TO:** Sébastien Bailly, PhD; email: [sbailly@chu-grenoble.fr](mailto:sbailly@chu-grenoble.fr)

Copyright © 2022 American College of Chest Physicians. Published by Elsevier Inc. All rights reserved.

**DOI:** <https://doi.org/10.1016/j.chest.2022.11.034>

improve cardiovascular outcomes.<sup>4-6</sup> The effectiveness of PAP treatment relies on adherence,<sup>7</sup> with some data suggesting that a minimum of 4 h and probably 6 h of PAP per night is required to improve BP control in patients who are minimally symptomatic.<sup>6,8</sup>

Over the past 10 years, the development of communicating PAP devices and the willingness of manufacturers and home care providers to reshape follow-up care have enabled the emergence of remote monitoring platforms for the visualization of nightly data that are generated by hundreds of millions of patients worldwide.<sup>9</sup> This technology has transformed the way patients are followed and monitored by providing an opportunity for early interventions to improve disease management,<sup>10,11</sup> PAP adherence,<sup>12,13</sup> and augment telemedicine activities in sleep breathing disorders.<sup>14</sup> In some countries, such as France, remote monitoring of PAP adherence is a condition for health care insurance reimbursement of PAP devices, with rates of reimbursement proportionate to levels of adherence.

PAP data can be used to improve our understanding of OSA and to personalize treatment by the use of

innovative statistical approaches alongside artificial intelligence for the identification and prediction of the trajectories of patients who are treated with PAP.<sup>15,16</sup> PAP monitoring data include daily aggregated measurements of adherence<sup>17</sup> and treatment efficacy (residual apnea-hypopnea index [rAHI] and leaks) (Fig 1).

The majority of studies that are based on longitudinal PAP remote monitoring summarize the data trajectories in static and simplistic metrics for PAP adherence and rAHI by using mean or median values.<sup>10,18-20</sup> However, much of the value of these data lies in the complexity of their evolution and variability over time, which is very relevant for patient care and treatment. Because data are collected daily, they can be analyzed as time series and modeled with the use of multiple deterministic and stochastic components. Indeed, these data may reveal, for example, trends, cyclic components, or disruptions in the observed behaviors.<sup>21-23</sup>

Yet currently, the reliability of summarized data remains questionable, in particular because of to the absence of standardization of PAP device-generated indicators among the various PAP models and manufacturers,

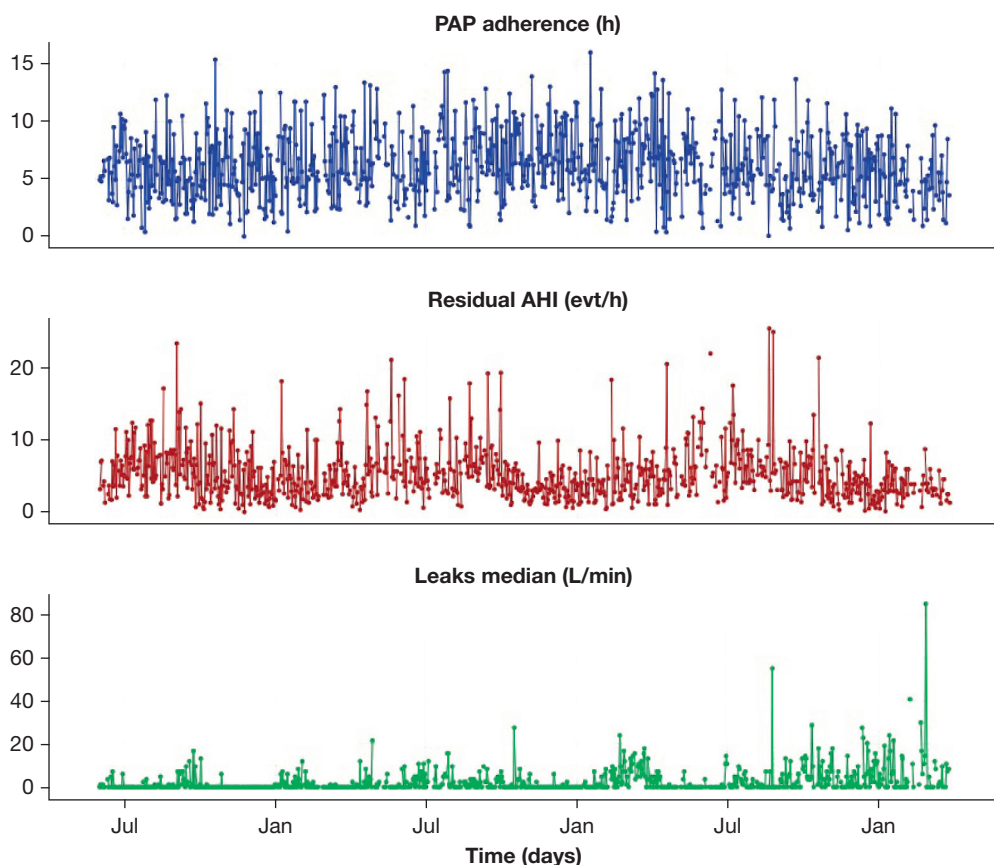


Figure 1 – Daily indicators from PAP device records. AHI = apnea-hypopnea index; evt = event; PAP = positive airway pressure.

different approaches to the management of missing values,<sup>24</sup> and the involvement of different intermediaries (ie, home care providers or private digital health companies) that process original data with some approximations or reduction in dimensionality. Clinicians and data scientists who access these data are often unaware of the entire processing pipeline that could impact interpretation of the data. Thus, there is a need to develop standardized methods, definitions, and

terminology for cleaning, processing, reporting, and interpreting PAP data.

The aim of this article is to suggest directions for improving the data cleaning and processing and to highlight major concerns for data science applications that rely on daily remote monitoring data from PAP devices (including fixed PAP and automated PAP). The recommendations are intended not only for researchers who analyze the PAP data, but also for home care

**TABLE 1 ] Summary of Problems, Impacts, and Solutions**

Problem to Address	Potential Impact	Proposed Solution
<b>Reliability of the residual apnea-hypopnea index</b>		
The residual apnea-hypopnea index, when measured over short periods of use, is a poor statistical estimator and may not be representative of the whole night.	The residual apnea-hypopnea index measured over short periods of use may not be reliable or may be biased.	Set a minimum device usage threshold (eg, > 1 h) for consideration of the residual apnea-hypopnea index.
Drift of the flow curve is due to large number of leaks.	Respiratory events cannot be detected, and the reported residual apnea-hypopnea index may be distorted.	Do not interpret residual apnea-hypopnea index in cases with excessive leaks.
<b>Device replacement and absence of standardization</b>		
Transmission bugs can result from change in device model and/or manufacturer.	Data may not be transmitted, even though the patients use the device.	Missing values should not be replaced by zero values.
Parameters reported and summary statistics for leaks and residual apnea-hypopnea index are different among PAP manufacturers.	It is impossible to compare directly data from different devices.	Transition periods between two distinct devices require specific data preprocessing.
<b>Multiple records of the same device for a patient</b>		
Several records can be sent on the same day.	Multiple rows exist for a patient with different values.	Aggregate multiple rows by summing the usage values and compute the mean weighted residual apnea-hypopnea index.
Mismatch occurs between the machine identification number and the patient identification number.	The records from two distinct devices are assigned to the same patient.	Provide the device identification number or remove the period with duplicate devices.
Dataflow issue exists.	Duplicate rows exist for a patient.	Remove duplicates.
<b>Missing values</b>		
Transmission with external modem	Missing values because of connecting issues	Remove data recorded by PAP device with external modem.
Between two transmissions	Bias related to missing values	Missing values can be replaced by zero.
At the end of a patient's sequence	Bias related to missing values	Imputation strategy should be defined.
<b>Treatment interruption</b>		
With daily aggregated data, there is no distinction whether the patient did three 2-h naps or slept 6 h consecutively at night, for example.	The same aggregated value of CPAP adherence might reflect completely different CPAP usages that potentially are related to CPAP termination and prognosis (eg, insomnia)	Use raw data that provide the exact time of CPAP exposures during night and day.

providers and health services and those who develop PAP telemonitoring data visualization applications for the guidance and the alerts to clinicians, caregivers, and patients.

In the following sections, our objective is to provide an overview of the landscape of PAP remote monitoring data, including concerns regarding (1) conditions for rAHI reliability, (2) lack of standardization of the summarized data provided by PAP manufacturers, (3) handling of missing values, and (4) consideration of treatment interruptions. These issues, their impact, and potential solutions are summarized in [Table 1](#).

### Data Source

Raw data, including airflow and pressure, are collected at a high frequency (ie, 25 Hz for the airflow signal) by the PAP devices. Real-time respiratory event detection is

performed during the night by the proprietary device's embedded software. At the end of the night, raw records and summary statistics of the data are sent to the equipment provider and stored in the secure digital card of the PAP device. Then, summarized data are made available to health care providers. This information usually includes the time and duration of PAP usage, the estimated rAHI, and leaks (mean or median, 95th percentile, and/or maximum), depending on the PAP device model/manufacturer.

Depending on the source of the data to be analyzed, its life-cycle and nature may be different ([Fig 2](#)) and can vary among different countries and health systems. Data are available from (1) the manufacturers' databases, with mainly device-generated data and some limited clinical data collected through patient engagement applications,<sup>13</sup> (2) the home care providers' databases,

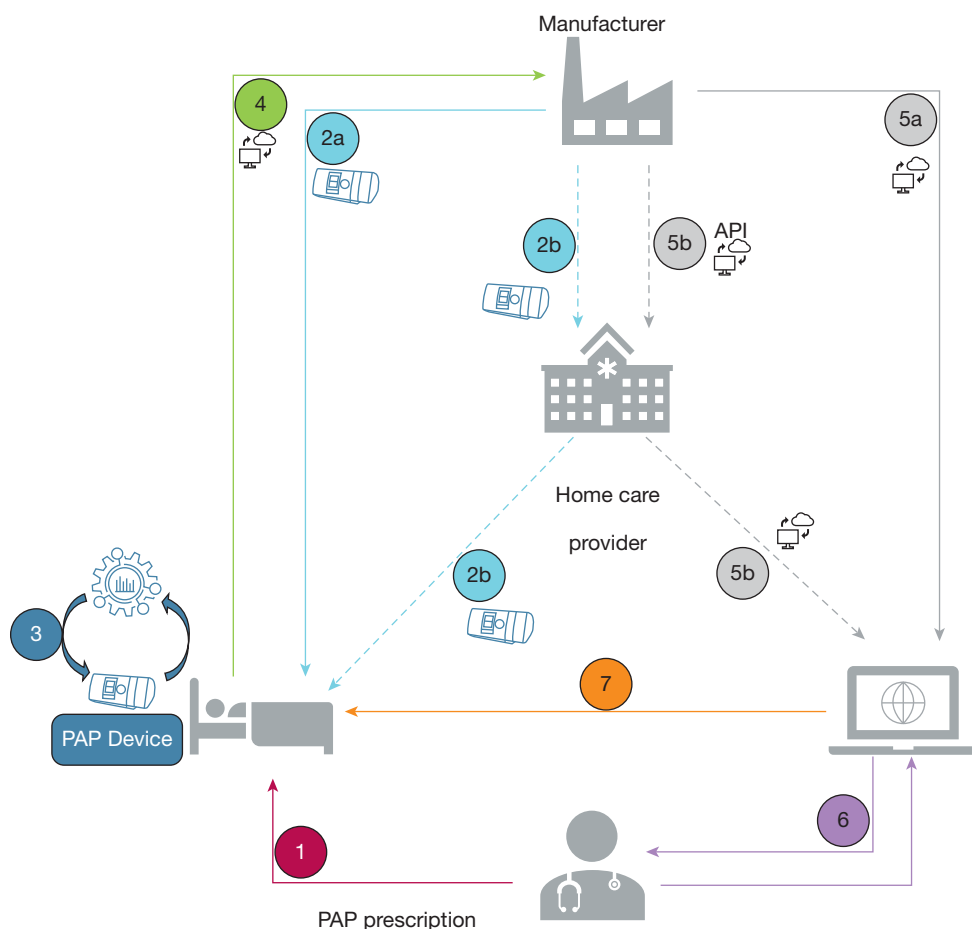


Figure 2 – From PAP Prescription to PAP Data Visualization: PAP Data Processing (1) PAP prescription from clinician to the patient. (2) PAP device distribution either directly from the manufacturer (2a) or via a home care provider (2b). (3) Real-time respiratory events detected by proprietary software embedded in PAP device. (4) Data transfer from the PAP device to the manufacturer. (5) PAP data are stored in a secured online platform either by the manufacturer (5a) or by the home care provider (with the use of an Application Programming Interface) who can make data transformations (5b). (6) Clinician can visualize and download PAP data for their patient with the use of a web platform. (7) The patient can visualize data on the device's screen or with the use of a web platform or a mobile application. API = application programming interface; PAP = positive airway pressure.

with additional information on accessories, age, sex, and sometimes BMI and AHI at diagnosis, (3) clinical databases, with primarily clinical data and limited device data from the manufacturers or home care providers. Raw data recorded by the PAP device can be transmitted via a modem, external to the PAP device, or a PAP device connected to the Wi-fi or Bluetooth that is downloaded from the device during home visits by the homecare provider's technician or directly obtained during PAP device recalls.

Advances in technology have made the connected PAP device the most common remote monitoring system, sending data on a daily basis, and keeping data records over several weeks (to be sent later in case of transmission failure [eg, because of an interrupted internet connection]). Some limitations have to be considered because of transmission systems or specific device constraints.<sup>25</sup> Respiratory events detection and the computation of summary statistics are performed according to each individual manufacturer's specifications. Data can be displayed on the manufacturer's platform or sent to the various health care providers' platforms via application programming interfaces as raw or aggregated data, one value that summarizes measurements made over 24 h (from midnight to midnight or from noon to noon) for a given patient (Fig 2).

### Reliability of the rAHI

Independently of the data transmission procedure, device usage and leaks can influence the reliability of the reported rAHI.

First, we recommend analyzing the distribution of rAHI values throughout 24 h, retaining in the final analysis only periods during which the device was used continuously for several hours and filtering out very short periods. We make this statement for several reasons. First, the empirical mean of the events detected over a short period is a poor statistical estimator of the true rAHI for the whole night. Second, the occurrence of residual events is influenced by many factors including body position, sleep stage, and time of night.<sup>26</sup> Sleep instability during the first hours of the night could impact the dynamics of the respiratory events. The risk of residual events, despite PAP therapy, increases during rapid eye movement sleep, whereas deep (N3) sleep is generally free of apneas and hypopneas.<sup>27-29</sup> Thus, the rAHI measured over 1 or 2 h of the night might not be representative of the whole

night and overall treatment efficacy. Accordingly, in previous studies that have analyzed rAHI big data, the investigators discarded periods with device usage of < 1 h.<sup>27,28</sup>

Furthermore, above a maximum leak threshold, the respiratory event detection is unreliable, and the rAHI can be distorted (Fig 3) and should probably not be retained.

### Device Change and Absence of Standardization of Reported Indicators

A change of PAP devices (eg, change from one manufacturer to another or even to a different model from the same manufacturer) can generate some transmission bugs that will last for a few days. This can cause missing values or periods with no recording, even though the patient has been using the device. Thus, all missing values cannot be interpreted as poor adherence to PAP therapy.

The algorithms used by each manufacturer are somewhat different, which leads to important variations in the reported indicators because there is no standardization in the computation nor in reporting methods between PAP manufacturers or brands. Indeed, the parameters included in summary statistics for rAHI and leaks and how they are reported are different among PAP manufacturers (ie, only nonintentional leaks or all leaks, median or mean, and/or 95<sup>th</sup> percentile and/or maximum value over the night).<sup>29</sup>

Thus, one might observe an apparent sudden change in the rAHI because of both the period of adaptation to the new device and/or a difference in the way of calculating the rAHI (Fig 4).<sup>24</sup>

As a consequence, for each patient, transition periods between two distinct devices require specific data preprocessing. To assess the impact of changes in PAP device at the individual level, we suggest that all device changes be recorded systematically and that the statistical distribution of the indicators and rates of missing and null usage values over a period before and after the change be compared. In the analysis of large databases for data science purposes, one solution is to consider data from only one device for each patient and to censor the second device. Another solution is to consider only devices from a single manufacturer to ensure coherence in the method used to calculate the summary statistics. Supplementing the data downloads with clinical history can also be quite informative at an individual level.

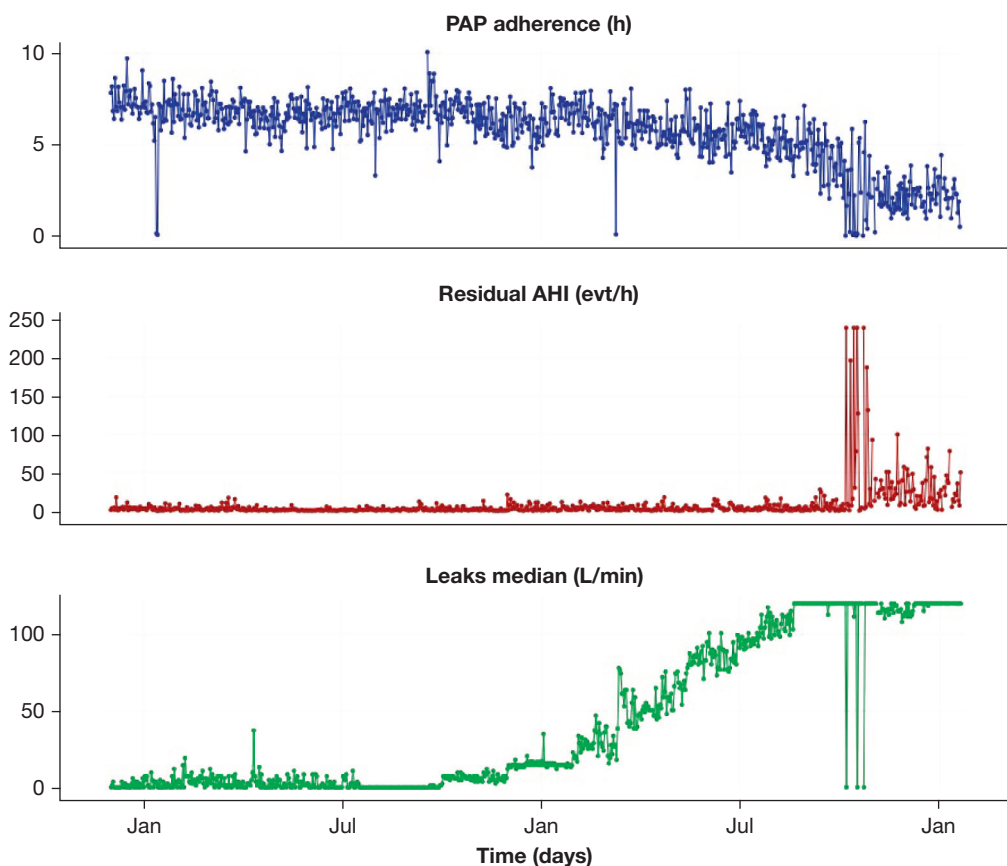


Figure 3 – Illustration of the Unreliability of the Reported Residual Apnea-Hypopnea Index in the Presence of Excessive Leaks. AHI = apnea-hypopnea index; evt = event; PAP = positive airway pressure.

### Multiple Records From the Same Patient

During database “cleaning,” one should ensure that it does not include duplicate recordings or errors in device attribution. After the device has been checked changes, this problem may be due to one or more of three scenarios.

First, several recordings might be sent in the same day if multiple sleep periods are transmitted because of the

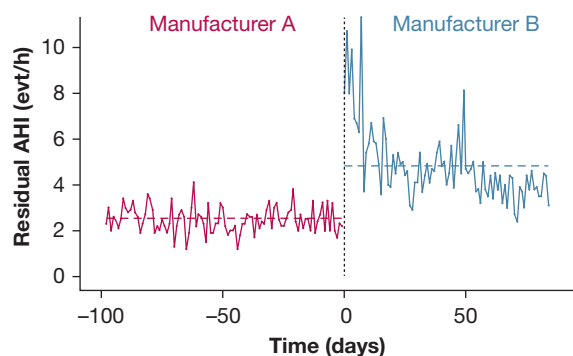


Figure 4 – Impact of Device Change in Reported Residual Apnea-Hypopnea Index. Dashed lines indicate the mean residual apnea-hypopnea index over time. AHI = apnea-hypopnea index; evt = event.

PAP device restarting several times during the night or when the patient takes several short daytime naps. We suggest aggregating them by summing the hours of use and computing the mean weighted rAHI values, bearing in mind that the rAHI is reliable probably only for longer periods of use.

Second, there might be mismatches between device identification numbers and patient identification that would result in the records from two distinct devices being attributed to the same patient, despite the use of only one PAP device (Fig 5A). In practice, this situation can occur when a device develops faults and when the patient returns it to the supplier. Mismatches can also occur when patients change their home care provider. The problem can be solved by asking the patient to always give the device identification number (Fig 5B and C) or by removing the period with duplicate devices from the database if it is impossible to identify them correctly (Fig 5D).

The third scenario is a dataflow issue that results in duplicate rows for a patient with the same device and



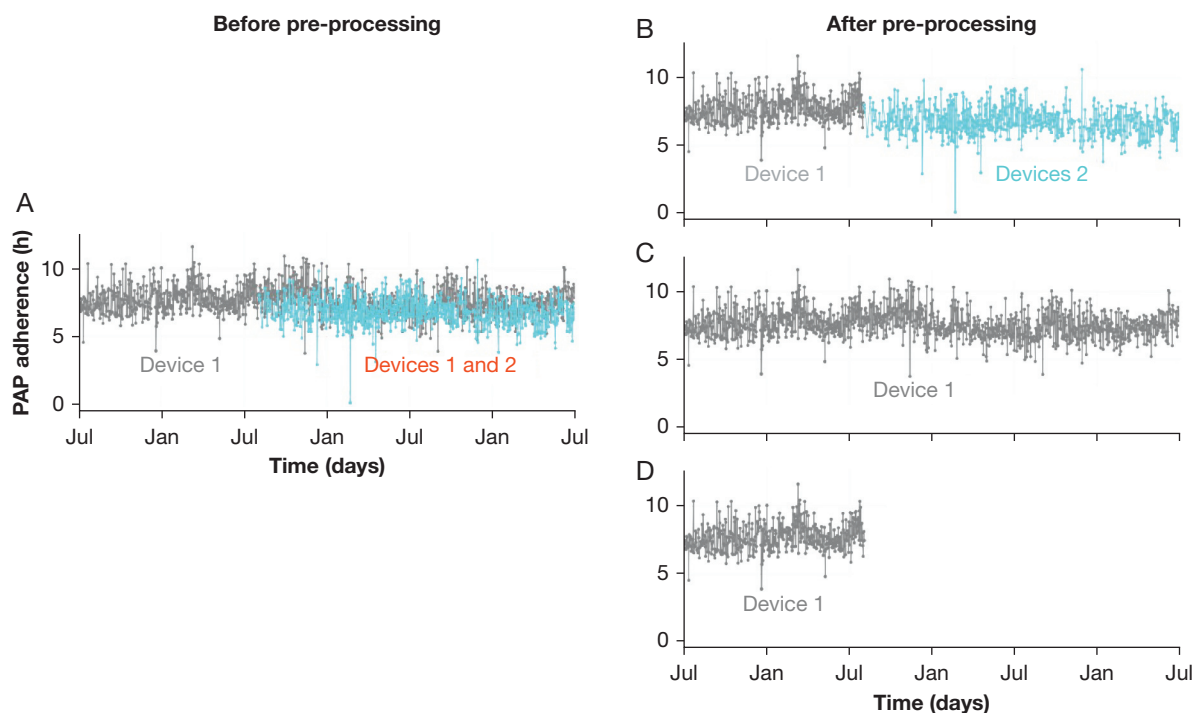


Figure 5 – A-D, Example of a Situation in Which Two PAP Devices Are Simultaneously Assigned to a Single Patient. A, Plot of the PAP usage of a patient who was assigned the records of two distinct devices. B, Usage records to keep for analysis if device 2 was used by the patient after device 1. C, Usage records to keep for analysis if the device 1 was used throughout the period. D, Usage records to keep for analysis if no information can be retrieved about which device actually was used by the patient after device 1. PAP = positive airway pressure.

collected values. This can be solved by the removal of duplicate rows where all the values are completely identical. This removal may be sufficient by keeping one row per patient per day.

### Missing Values

It is important to check the distribution of rates of missing values and null usage times (ie, records with a zero-usage time) so as to reveal inconsistencies. One possible source of missing data is due to PAP devices that transmit data with the use of an external modem, which may become disconnected or misconnected to the PAP device. These devices should be indicated in the database and preferentially excluded before analysis to avoid errors in interpretation of null or missing values. Moreover, depending on the country's data privacy regulations, patient consent, and the personalization of follow up, additional annotations may be collected to explain the nonuse of the devices (hospitalizations, vacations, travel, or discontinuation of PAP treatment) and thus reduce the rate of missing values.

### Imputation of Missing Values

In the recent literature, different methods have been used to deal with missing values in remote monitoring of

PAP adherence: (1) replacing missing values by zero,<sup>16,30-32</sup> (2) using an imputation procedure when the missing values are due to transmission failure,<sup>21,23</sup> (3) excluding patients with excessive missing data,<sup>23</sup> (4) not replacing data,<sup>33</sup> or (5) not reporting what has been done.<sup>34</sup> The relevance of each method depends on the situation: if the missing values are observed between two transmissions or if missing values are observed at the end of the patient's time series.

### Imputation of PAP Adherence Values in Case of Missing Records Between Two Transmissions

If we consider a patient with several consecutive missing records between two transmissions (Fig 6A and B). The storage capacity of their device is  $p$  days. Then the period with missing records can be imputed with zero values if the length of this period is inferior to  $P$  value (Fig 6A) and if the patient was not using another device (for example during a hospitalization).

When information on periods such as vacations or travel is known, it is possible to consider these as specific cases for which missing values can be replaced by zero, whatever the storage period of the device, providing that the patient confirms that they did not use their device over this time period (Fig 6A). Different reasons can



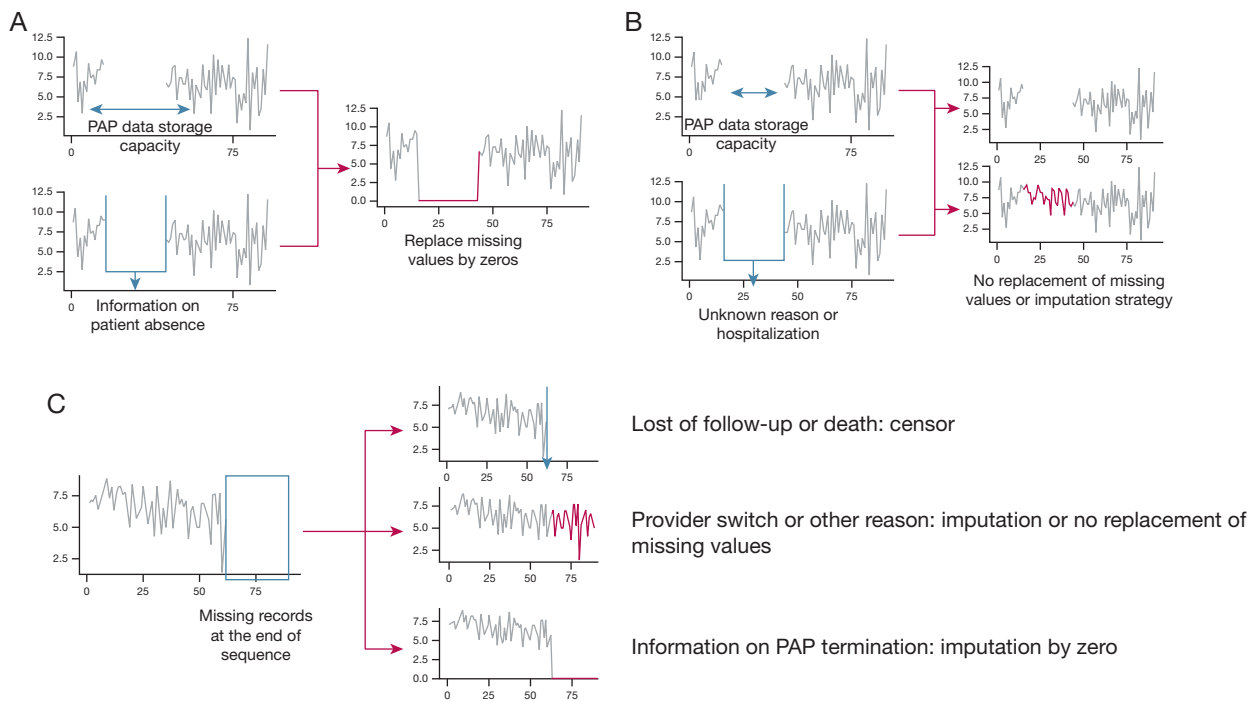


Figure 6 – A-C, Management of missing data. A, Missing values are observed within a patient sequence. Either the length of the missing value is lower than the data storage capacity of the PAP device without usage of another device (for example, during a hospitalization) or the patient provides a reason for non-use of any PAP device (for example, when on vacation or travelling). In these cases, missing values can be replaced by 0. B, Missing values are observed within a patient sequence, and the patient does not provide a reason for non-use of the PAP device or the length of missing values is higher than PAP storage capacity; missing values cannot be replaced by 0. An imputation method must be considered. C, Missing values are observed at the end of a patient sequence, and three possibilities can be considered according to the information about missing values. PAP = positive airway pressure.

lead to change in data reporting, such as a change of time-zone, mask changes, humidification, damage to the device.

Otherwise, missing values cannot be imputed with zero values and have to remain either as missing values or a specific imputation method should be considered if the patient effectively used a PAP device; however, this approach requires an understanding of complex adherence behaviors (Fig 6B).

### Imputation of PAP Adherence Values in Case of Missing Records at the End of a Patient's Sequence

Some patients' data can show missing values during the interval between the last available record and the date when the data were exported for analysis (Fig 6C). This can be due to PAP termination (eg, initiation of alternative therapies, change in PAP provider, or death). If it is known that the patient stopped the therapy, then the missing sequence should be imputed with zero values. If the device restitutions are caused by a change in PAP provider, then the patient is considered to have continued the therapy, so the sequences should not be filled with zero values. A sequence that is interrupted

because of death should be censored before it is included in any analysis.

### Example of Impact of Imputation of Missing Values on PAP Adherence Results

From a random sample of 407 patients with daily PAP adherence values collected from a home care provider over 3 months after PAP initiation, without any observed PAP termination, in our own data, 33 patients (8%) had missing values when PAP adherence was considered over the first month of use.

This proportion increased to 15% at month 2 and 19% at month 3. In the total sample, no change in the mean of the individual average adherence values was found, whether individual averaging was done with or without imputation (Fig 7). The greatest difference was observed if patients with missing values were removed from the sample (Table 2). When we considered only patients with missing values, the imputation strategy leads to changes because of the small sample size (N = 33, 60, and 78), especially if the missing values were replaced by zeros (Table 2, Fig 7). The proportion of missing values was related to both the sample size

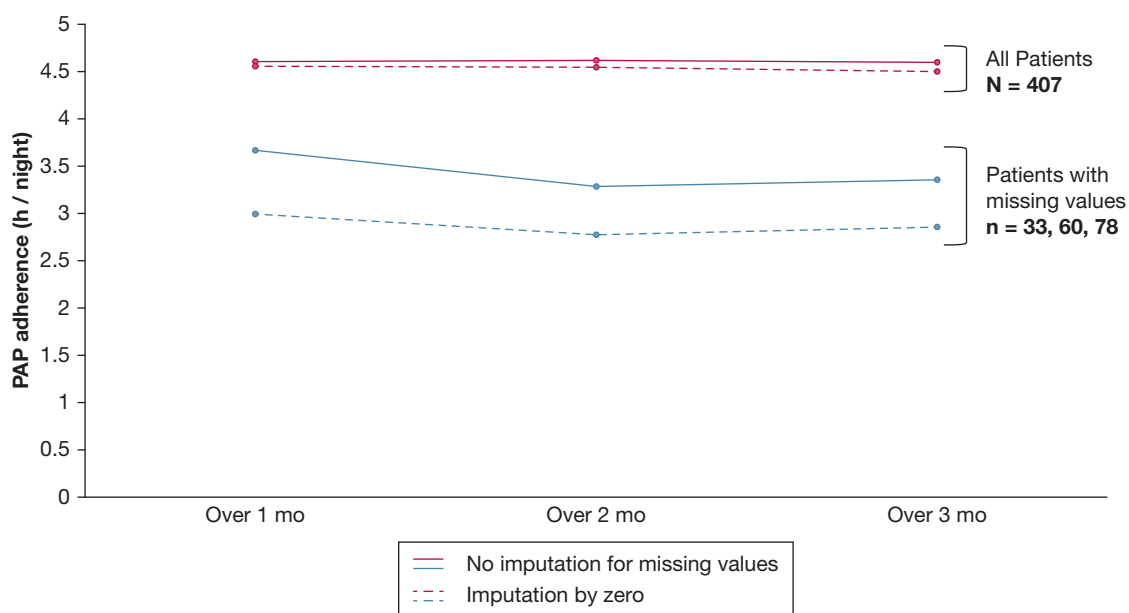


Figure 7 – Change in mean PAP individual average adherence over different numbers of months of therapy according to sample and missing value imputation strategy. PAP = positive airway pressure.

and the percentage of missing values within a given sample. Here, we present a large sample with a low frequency of missing values that resulted in small

differences whatever the methods used to handle missing values. This result contrasts to the proactive and efficient replacement of missing data in the case of

**TABLE 2 ]** Impact of Different Strategies for Imputing Adherence Missing Values on Mean Adherence Value and Rate of Patients With Mean PAP Use of  $\geq 4$  h/night

Variable	Study Period		
	1 mo	2 mo	3 mo
Missing values, mean (SD), No.	0.5 (2.6)	1.6 (5.9)	3.5 (11.2)
Maximum missing values, No.	29	58	86
<b>All samples</b>			
No strategy for missing values			
Individual average PAP adherence, mean (SD), h/night	4.61 (2.4)	4.62 (2.4)	4.6 (2.4)
PAP-adherent patients, <sup>a</sup> No. (%)	257 (63)	255 (63)	255 (63)
Replacing missing values by zeros			
Individual average PAP adherence, mean (SD), h/night	4.56 (2.4)	4.55 (2.4)	4.5 (2.4)
PAP-adherent patients, <sup>a</sup> No. (%)	252 (62)	249 (61)	251 (62)
Removing patients with missing values			
Individual average PAP adherence, mean (SD), h/night	4.69 (2.4)	4.85 (2.3)	4.89 (2.3)
PAP-adherent patients, <sup>a</sup> No. (%)	239 (64)	230 (66)	221 (67)
Patients with missing values, No. (%)	33 (8)	60 (15)	78 (19)
No strategy for missing values			
Individual average PAP adherence, mean (SD), h/night	3.67 (2.4)	3.29 (2.4)	3.36 (2.4)
PAP-adherent patients, <sup>a</sup> No. (%)	18 (55)	25 (42)	34 (44)
Replacing missing values by zeros			
Individual average PAP adherence, mean (SD), h/night	3.0 (2.1)	2.78 (2.2)	2.86 (2.3)
PAP-adherent patients, <sup>a</sup> No. (%)	13 (39)	19 (32)	30 (38)

Results are presented for two different populations: the entire sample (N = 407) and the subsample of patients for whom values are missing.

<sup>a</sup>With mean PAP use  $\geq 4$  h/night.

a small sample or a significant percentage of missing values (eg, > 20%).

In summary, if missing values occur between two transmissions, the suggestion is to replace them by zeros if the period is shorter than the PAP device's storage capacity and if no other PAP device was used (eg, at the hospital) and otherwise to leave the values as missing or to use an imputation strategy. If missing values are at the end of the exported sequence, the suggestion is to censor the data in case of loss to follow up or death, to leave the missing values or use an imputation strategy in case of change in provider, and to replace by zeros in the case of PAP termination.

To impute missing data, one solution is to consider each indicator as an independent time series and compare the performance of several methods: mean/mode imputation, last observation carried forward, next observation carried backward, mean of the last and the next observation, interpolation (linear, spline), and time series modeling and forecasting.<sup>35</sup> Another solution consists in imputing the multiple incomplete variables with an imputation method suited to multivariate longitudinal data, with the use of either joint modelling (multivariate normal imputation) or with fully conditional specification (sequential regression and multiple imputation by chained equations).<sup>36</sup>

### Short Treatment Interruptions Go Undetected in Aggregated Data

When only the mean or median of aggregated 24-h data are presented, there is no distinction between the patient using the PAP device for three 2-h daytime naps or using it continuously for 6 h at night. The following example shows that discontinuities in PAP use during the night, which are often due to simply removal of the mask, are frequent.

From an exploratory analysis of our ongoing research database, in a sample of 15.7 million night-time records from 2,607 patients who used PAP devices from the same manufacturer between June 2017 and April 2021, the majority of the interruptions in records (81.3%) corresponded to one or two-mask removals. In 15.9% of the records, the mask was removed between three and five times during the night. Moreover, 2.8% of the records were the aggregation of five to ten separate short periods of use. If data about mask removals and the distinct periods of PAP use over 24 h are available, then nights that are broken up into many parts could be identified.

Indeed, the same aggregated 24-h PAP adherence could hide completely different PAP usages, some potentially related to long-term PAP termination and poor prognosis. For example, in patients with comorbid insomnia and sleep apnea,<sup>37,38</sup> one of the most frequent OSA phenotypes is that numerous distinct periods of PAP use occur during the night and during naps in the day, which reflects the severity of insomnia and sleep deprivation. Likewise, the 24-h mean or median residual AHI does not reveal events that occur during only a particular phase of sleep, such as the increased AHI during rapid eye movement sleep, which has been associated with hypertension and diabetes mellitus.<sup>39,40</sup> Thus, it is preferable to use raw data that provide the exact times of PAP exposure during the night and the day. Furthermore, the timing of PAP use can provide information about the circadian rhythm, which can be valuable for some analyses.

Finally, when the analysis focuses on the indicators that were measured while the patients are using their PAP device (rAHI, leaks), only relatively unfragmented nights with less than five mask removals and several hours of continuous use should be retained for the analysis.

### Summary and Perspectives

The remote monitoring of PAP data provides the opportunity to access valuable and objective information regarding patient adherence, patterns of use, and treatment efficiency. Continuous monitoring enables clinicians, health care providers, and researchers to follow variations in several parameters over short and long periods of time. Although numerous studies have been based on summary statistics (24-h means or medians) of daily PAP usage and indicators, there is growing interest in the consideration of internal structure of the daily aggregated data and longer term trajectories that are considered in the framework of time series analysis.

To increase the comparability among studies and to avoid biases in computation, the modalities of PAP data management and processing should be considered carefully. First, excessive leaks (for which thresholds defining what constitutes a leak vary among the device manufacturers) can result in inaccurate rAHI estimations. Second, when PAP reported rAHI is analyzed, the inclusion of short periods of device usage should be avoided because they do not reflect the actual mean number of residual events. However, these short periods of usage should be kept when detailed PAP

adherence, disturbed sleep, and sleep patterns across the 24-h span are reported.<sup>37</sup> Third, if possible, the number of missing values should be minimized, and imputation strategies should be well described. However, future studies are needed to compare the effect of different imputation methods on PAP data to set guidelines. Fourth, ways of avoiding errors in PAP device/patient attribution or data transmission should be explored to prevent situations such as multiple devices being attributed to one patient (which can happen during a device change) or multiple sets of data for one patient in the same day.

In cross-sectional or short-term studies, different levels of granularity of data aggregation could be provided to reveal details such as mask removal during the night. Such data provide contextual annotations that are related to patient phenotypes that influence PAP device use, including nocturia, other sleep disorders such as insomnia or multiple sequential sleep periods, and daytime naps. To our knowledge, such detailed data currently are not considered in PAP data studies.

A change in device could impact long-term studies, particularly if the new device comes from a different manufacturer. The nonstandardization of indicator thresholds (leaks or rAHI) can lead to major biases in data analyses. To avoid such biases, it would be preferable to consider only records for patients with a single model of device. An alternative possibility is to use an external device that can ensure PAP data standardization regardless of the PAP manufacturer.<sup>41</sup> However, with current technology, precautions must be taken to avoid the problems that are encountered with external modem transmissions. The creation and implementation of international guidelines for standardization might help to solve this concern.

Missing data are a concern in all types of study. Contacting individual patients to determine the reason that the data are missing, duplicated, or inconsistent might be considered, but only if it is essential for the analysis. Indeed, this can be a tedious task that is feasible only in well-structured health services. This approach is probably unattainable in fragmented health care systems, which can occur in the United States, involving many private stakeholders.

The spread of PAP remote monitoring worldwide potentially increases the amount of data that could become available from a huge population of patients. Well-managed national and international collaborative digital databanks will be essential for the enhancement

of data science in this field and the facilitation of the move toward artificial intelligence approaches. There is a unique opportunity to develop and improve in-depth data analyses that will consider not only computed features at the trajectory level but also whole time series of daily records or even raw data sampled at a higher frequency, such as the raw airflow signal recorded by PAP devices. This approach could contribute to the development of personalized tools to improve PAP management, to ensure data quality control, to provide data for health insurance reimbursement purposes, and ultimately to provide better health of patients.

By developing standardized processes for data management of PAP records and by working on a standardization of metrics produced by PAP devices, it could be possible to compare different management patterns and to develop standardized tools for patient self-monitoring to improve PAP adherence and treatment efficiency.

The potential uses of PAP remote monitoring data are multiple and currently are underused in the field of sleep research. Improving the quality of data from PAP remote monitoring would encourage more widespread use. This approach would help promote personalized medicine through various applications that would include the identification of patient trajectories of PAP use or the variability of treatment efficiency based on rAHI measures.<sup>22,23,32,42</sup> The different clusters of PAP trajectories potentially are linked to patient-reported outcomes and long-term incident cardiovascular events and death.

However, considering only PAP data has some limitations. There are a number of other relevant parameters, such as the number of hours of sleep,<sup>43</sup> patient-reported outcomes measures,<sup>44</sup> technical alerts, or patient demographics and socioeconomic factors that can help to better characterize patient trajectories. Thus, the aggregation of clinical, technical, and sociologic data could help to improve personalized medicine in the field of sleep apnea by considering individual variability.<sup>30</sup>

### Funding/Support

G. B.-B., J.-L. P., S. B., and R. T. are supported by the “e-health and integrated care and trajectories medicine and MIAI artificial intelligence” Chairs of excellence from the Grenoble Alpes University Foundation (ANR-19-P3IA-0003) and the French National Research Agency in the framework of the “Investissements d’avenir” program (ANR-15-IDEX-02). A. Midelet is

supported by Probayes and MIAI (ANR-19- P3IA-0003) in the framework of a “Convention Industrielle de Formation par la Recherche” (CIFRE) PhD; her PhD is also supported by the French National Research Agency (grant 2020/0007). A.Malhotra is funded by National Institutes of Health.

## Financial/Nonfinancial Disclosures

None declared.

## Acknowledgments

**Author contributions:** G. B.-B., A. Midelet, M. M., J.-L. P., and S. B. contributed substantially to the concept, literature search, the initial writing of the manuscript. J.-C. B., J.-B. M., A. Malhotra, and R. T. substantially improved the manuscript from an expert clinical perspective. R. L. H., M.-C. S., A. S., and A. H. further improved the manuscript by providing expertise from a data science point of view.

**Role of sponsors:** The sponsor had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.

**Other contributions:** We thank Alison Foote, PhD (an independent medical writer based in Grenoble, France), for critical reading, partial rewriting, and editing of the manuscript.

## References

1. Benjafield AV, Ayas NT, Eastwood PR, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. *Lancet Respir Med*. 2019;7(8):687-698.
2. Bailly S, Grote L, Hedner J, et al. Clusters of sleep apnoea phenotypes: a large pan-European study from the European Sleep Apnoea database (ESADA). *Respirology*. 2021;26(4):378-387.
3. Javaheri S, Barbe F, Campos-Rodriguez F, et al. Sleep apnea: types, mechanisms, and clinical cardiovascular consequences. *J Am Coll Cardiol*. 2017;69(7):841-858.
4. Levy P, Kohler M, McNicholas WT, et al. Obstructive sleep apnoea syndrome. *Nat Rev Dis Primers*. 2015;1:15015.
5. Pépin JL, Bailly S, Rinder P, et al. Relationship between CPAP termination and all-cause mortality: a French nationwide database analysis. *Chest*. 2022;161(6):1657-1665.
6. Gerves-Pinque C, Bailly S, Goupil F, et al. Positive airway pressure adherence, mortality and cardiovascular events in patients with sleep apnea. *Am J Respir Crit Care Med*. 2022;206(11):1393-1404.
7. 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*. 2019;15(2):335-343.
8. Bratton DJ, Stradling JR, Barbe F, Kohler M. Effect of CPAP on blood pressure in patients with minimally symptomatic obstructive sleep apnoea: a meta-analysis using individual patient data from four randomised controlled trials. *Thorax*. 2014;69(12):1128-1135.
9. Dusart C, Andre S, Mettay T, Bruyneel M. Telemonitoring for the follow-up of obstructive sleep apnea patients treated with CPAP: accuracy and impact on therapy. *Sensors (Basel)*. 2022;22(7):2782.
10. Pepin JL, Jullian-Desayes I, Sapene M, et al. Multimodal remote monitoring of high cardiovascular risk patients with OSA initiating CPAP: a randomized trial. *Chest*. 2019;155(4):730-739.
11. Tamisier R, Treptow E, Joyeux-Faure M, et al. Impact of a multimodal telemonitoring intervention on CPAP adherence in symptomatic OSA and low cardiovascular risk: a randomized controlled trial. *Chest*. 2020;158(5):2136-2145.
12. Fox N, Hirsch-Allen AJ, Goodfellow E, et al. The impact of a telemedicine monitoring system on positive airway pressure adherence in patients with obstructive sleep apnea: a randomized controlled trial. *Sleep*. 2012;35(4):477-481.
13. Malhotra A, Crocker ME, Willes L, Kelly C, Lynch S, Benjafield AV. Patient engagement using new technology to improve adherence to positive airway pressure therapy: a retrospective analysis. *Chest*. 2018;153(4):843-850.
14. Verbraecken J. Telemedicine in sleep-disordered breathing: expanding the horizons. *Sleep Med Clin*. 2021;16(3):417-445.
15. McNicholas WT, Bassetti CL, Ferini-Strambi L, et al. Challenges in obstructive sleep apnoea. *Lancet Respir Med*. 2018;6(3):170-172.
16. Liu D, Armitstead J, Benjafield A, et al. Trajectories of emergent central sleep apnea during CPAP therapy. *Chest*. 2017;152(4):751-760.
17. Pepin JL, Bailly S, Tamisier R. Big data in sleep apnoea: Opportunities and challenges. *Respirology*. 2020;25(5):486-494.
18. Giampá SQC, Furlan SF, Freitas LS, et al. Effects of CPAP on metabolic syndrome in patients with OSA: a randomized trial. *Chest*. 2022;161(5):1370-1381.
19. Wohlgenuth WK, Chirinos DA, Domingo S, Wallace DM. Attempters, adherers, and non-adherers: latent profile analysis of CPAP use with correlates. *Sleep Med*. 2015;16(3):336-342.
20. Aardoom JJ, Loheide-Niesmann L, Ossebaard HC, Ripper H. Effectiveness of health interventions in improving treatment adherence for adults with obstructive sleep apnea: meta-analytic review. *J Med Internet Res*. 2020;22(2):e16972.
21. Aloia MS, Goodwin MS, Velicer WF, et al. Time series analysis of treatment adherence patterns in individuals with obstructive sleep apnea. *Ann Behav Med*. 2008;36(1):44-53.
22. Bottaz-Bosson G, Hamon A, Pepin JL, Bailly S, Samson A. Continuous positive airway pressure adherence trajectories in sleep apnea: clustering with summed discrete Frechet and dynamic time warping dissimilarities. *Stat Med*. 2021;40(24):5373-5396.
23. Babbin SF, Velicer WF, Aloia MS, Kushida CA. Identifying longitudinal patterns for individuals and subgroups: an example with adherence to treatment for obstructive sleep apnea. *Multivariate Behav Res*. 2015;50(1):91-108.
24. Midelet A, Borel JC, Tamisier R, et al. Apnea-hypopnea index supplied by CPAP devices: time for standardization? *Sleep Med*. 2021;81(5):120-122.
25. Vidigal TA, Brasil EL, Ferreira MN, et al. Proposed management model for the use of telemonitoring of adherence to positive airway pressure equipment - position paper of the Brazilian Association of Sleep Medicine-ABMS. *Sleep Sci*. 2021;14(spec1):31-40.
26. Eiseman NA, Westover MB, Ellenbogen JM, Bianchi MT. The impact of body posture and sleep stages on sleep apnea severity in adults. *J Clin Sleep Med*. 2012;8(6):655-666A.
27. Hicks A, Cori JM, Jordan AS, et al. Mechanisms of the deep, slow-wave, sleep-related increase of upper airway muscle tone in healthy humans. *J Appl Physiol (1985)*. 2017;122(5):1304-1312.
28. Koutsourelakis I, Lamprou K, Vagiakis E, Zakyntinos S. Resolution of apnoeas in slow wave sleep. *Sleep Breath*. 2016;20(2):819-820.
29. Series F, Series I, Cormier Y. Effects of enhancing slow-wave sleep by gamma-hydroxybutyrate on obstructive sleep apnea. *Am Rev Respir Dis*. 1992;145(6):1378-1383.
30. Patel SR, Bakker JP, Stitt CJ, Aloia MS, Nouraei SM. Age and sex disparities in adherence to CPAP. *Chest*. 2021;159(1):382-389.
31. Borker PV, Carmona E, Essien UR, et al. Neighborhoods with greater prevalence of minority residents have lower continuous positive airway pressure adherence. *Am J Respir Crit Care Med*. 2021;204(3):339-346.
32. Den Teuling NG P, van den Heuvel ER, Aloia MS, Pauws SC. A latent-class heteroskedastic hurdle trajectory model: patterns of adherence in obstructive sleep apnea patients on CPAP therapy. *BMC Med Res Methodol*. 2021;21(1):269.

33. Contal O, Poncin W, Vaudan S, et al. One-year adherence to continuous positive airway pressure with telemonitoring in sleep apnea hypopnea syndrome: a randomized controlled trial. *Front Med (Lausanne)*. 2021;8(4):626361.
34. Bertelli F, Suehs CM, Mallet JP, et al. Did COVID-19 impact positive airway pressure adherence in 2020? A cross-sectional study of 8477 patients with sleep apnea. *Respir Res*. 2022;23(1):46.
35. Moritz S, Sardá A, Bartz-Beielstein T, Zaefferer M, Stork J. Comparison of different methods for univariate time series imputation in R. *arXiv:1510.03924 [cs, stat]*. 2015.
36. Huque MH, Carlin JB, Simpson JA, Lee KJ. A comparison of multiple imputation methods for missing data in longitudinal studies. *BMC Med Res Methodol*. 2018;18(1):168.
37. Ong JC, Crawford MR, Wallace DM. Sleep apnea and insomnia: emerging evidence for effective clinical management. *Chest*. 2021;159(5):2020-2028.
38. Pieh C, Bach M, Popp R, et al. Insomnia symptoms influence CPAP compliance. *Sleep Breath*. 2013;17(1):99-104.
39. Mokhlesi B, Finn LA, Hagen EW, et al. Obstructive sleep apnea during REM sleep and hypertension. results of the Wisconsin Sleep Cohort. *Am J Respir Crit Care Med*. 2014;190(10):1158-1167.
40. Grimaldi D, Beccuti G, Touma C, Van Cauter E, Mokhlesi B. Association of obstructive sleep apnea in rapid eye movement sleep with reduced glycemic control in type 2 diabetes: therapeutic implications. *Diabetes Care*. 2014;37(2):355-363.
41. Leger D, Elbaz M, Piednoir B, Carron A, Texereau J. Evaluation of the add-on NOWAPI(R) medical device for remote monitoring of compliance to continuous positive airway pressure and treatment efficacy in obstructive sleep apnea. *Biomed Eng Online*. 2016;15(2):26.
42. Midelet A, Bailly S, Tamisier R, et al. Hidden Markov model segmentation to demarcate trajectories of residual apnoea-hypopnoea index in CPAP-treated sleep apnoea patients to personalize follow-up and prevent treatment failure. *EPMA J*. 2021;12(4):535-544.
43. Bakker JP, Weaver TE, Parthasarathy S, Aloia MS. Adherence to CPAP: what should we be aiming for, and how can we get there? *Chest*. 2019;155(6):1272-1287.
44. Mehta N, Mandavia R, Patel A, et al. Patient-reported outcome measure for obstructive sleep apnea: symptoms, tiredness, alertness, mood and psychosocial questionnaire: preliminary results. *J Sleep Res*. 2020;29(2):e12960.