EPMA Journal
https://doi.org/10.1007/s13167-021-00264-z

RESEARCH

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

Alphanie Midelet[1,2] · Sébastien Bailly[1,3] · Renaud Tamisier[1,3] · Jean-Christian Borel[1,4] · Sébastien Baillieul[1,3] · Ronan Le Hy[2] · Marie-Caroline Schaeffer[2] · Jean-Louis Pépin[1,3]

Received: 31 August 2021 / Accepted: 2 November 2021

Abstract

Background Continuous positive airway pressure (CPAP), the reference treatment for obstructive sleep apnoea (OSA), is used by millions of individuals worldwide with remote telemonitoring providing daily information on CPAP usage and efficacy, a currently underused resource. Here, we aimed to implement data science methods to provide tools for personalizing follow-up and preventing treatment failure.

Methods We analysed telemonitoring data from adults prescribed CPAP treatment. Our primary objective was to use Hidden Markov models (HMMs) to identify the underlying state of treatment efficacy and enable early detection of deterioration. Secondary goals were to identify clusters of rAHI trajectories which need distinct therapeutic strategies.

Results From telemonitoring records of 2860 CPAP-treated patients (age: 66.31 ± 12.92 years, 69.9% male), HMM estimated three states differing in variability within a given state and probability of shifting from one state to another. The daily inferred state informs on the need for a personalized action, while the sequence of states is a predictive indicator of treatment failure. Six clusters of rAHI trajectories were identified ranging from well-controlled patients (cluster 0: 669 (23%); mean rAHI 0.58 ± 0.59 events/h) to the most unstable (cluster 5: 470 (16%); mean rAHI 9.62 ± 5.62 events/h). CPAP adherence was 30 min higher in cluster 0 compared to clusters 4 and 5 (P value < 0.01).

Conclusion This new approach based on HMM might constitute the backbone for deployment of patient-centred CPAP management improving the personalized interpretation of telemonitoring data, identifying individuals for targeted therapy and preventing treatment failure or abandonment.

Keywords Apnoea-hypopnoea index · Continuous positive airway pressure: Patient-centred chronic pathology management · Targeted therapy · Predictive preventive personalized medicine · Telemonitoring

Alphanie Midelet and Sébastien Bailly contributed equally to this work.

✉ Jean-Louis Pépin
JPepin@chu-grenoble.fr

1 HP2 Laboratory, Inserm U1042, Grenoble Alpes University, 38000 Grenoble, France

2 Probayes, Montbonnot-Saint-Martin, France

3 EFCR Laboratory, Grenoble Alpes University Hospital, 38000 Grenoble, France

4 AGIR à dom. HomeCare Charity, 38240 Meylan, France

Abbreviations

3P-Medicine	Predictive, preventive and personalized medicine
AHI	Apnoea-hypopnoea index
BMI	Body mass index
CPAP	Continuous positive airway pressure
HMM	Hidden Markov model
IQR	Interquartile range
OSAS	Obstructive sleep apnoea syndrome
rAHI	Residual apnoea-hypopnoea index
SD	Standard deviation

Published online: 25 November 2021

Introduction

Obstructive sleep apnoea (OSA) is one of the most prevalent chronic diseases with nearly one billion people worldwide suffering from mild to severe forms [1]. OSA prevalence is likely to continue to increase owing to the parallel epidemics of obesity and diabetes. It is now widely accepted that the presence of OSA contributes to several poor health outcomes, including neurocognitive impairment, the accumulation and progression of metabolic and cardiovascular diseases and early mortality [2].

The reference first-line treatment for OSA is continuous positive airway pressure (CPAP) which is used nightly by millions of individuals worldwide. In the USA alone, the cost of treating OSA in 2015 reached approximately $US 6 billion, primarily for CPAP and oral appliance therapies [3]. Long-term CPAP treatment requires considerable human resources, and caregivers face difficulties in juggling a challenging workload so as to ensure good treatment adherence and personalize treatment when patients experience difficulties during their CPAP therapy. Sleep apnoea represents the perfect setting for developing and validating predictive, preventive and personalized solutions for the management of an essentially ambulatory chronic disease with millions of patients being regularly monitored [4–11]. A major step forward has been the recent availability of CPAP devices incorporating telemonitoring, providing daily information on CPAP usage and efficacy and helping to reduce the occurrence of abnormal events during sleep. The overall goal of remote monitoring platforms is to simplify follow-up pathways, particularly by reducing the burden on physicians, with the potential for reducing costs and inequalities in access to care.

One major concern during long-term CPAP treatment is to ensure the reduction and the normalization of respiratory events during sleep so as to restore normal oxygenation and improve sleep quality and continuity. Patients exhibiting an elevated residual apnoea-hypopnoea index (rAHI) have been reported to be at higher risk of therapy termination [7]. A challenge is thus the rapid notification of high rAHI through telemonitoring to facilitate early intervention to prevent therapy discontinuation and personalize accessories and ventilator modalities [5].

This extraordinary promise of CPAP telemonitoring and the avalanche of related data are expected to facilitate personalized care and multi-disease management in OSA patients, but this is currently hampered by the paucity of data mining in the field. Recent technological advances in CPAP telemonitoring data generation infrastructures have not been yet complemented by appropriate data analysis methodologies. rAHI alert thresholds are uniquely based on a monthly mean value or a moving average giving a clinical picture that massively under-uses the richness of information contained in the CPAP telemonitoring data. There is a need to implement state-of-the-art data science to describe the heterogeneity and diversity of raw CPAP telemonitoring data allowing a better understanding of disease progression and improving health processes.

The goal of the current study was to develop data mining methods to characterize and cluster the rAHI of CPAP-treated patients. We aimed to provide both visualization/cohort generation tools and informative novel rAHI indices to better characterize the CPAP telemonitoring trajectories of homogeneous subgroups of patients to personalize therapeutic strategies according to the group. In analogy with the sleep stage "hypnogram", we introduce the telemonitogram: a diagram that presents the stages of treatment efficacy as a function of time to facilitate early detection of any deterioration in treatment efficacy and prevent treatment failure.

Methods

Dataset and choice of datamining method

We analysed a CPAP telemonitoring database (registered and ethically approved by the French C.C.T.I.R.S: N°15.925bis and ethics regulations MR003 N° 1996650v0) to model and cluster patients' rAHI trajectories. The dataset included 2860 adults prescribed CPAP-treatment in fixed or automatic mode by the Grenoble Alpes University Hospital sleep laboratory between 2015 and 2020. All data were fully anonymized and only limited information (age and sex) regarding the patients' clinical condition was available.

The CPAP telemonitoring database contained one time series of nightly rAHI corresponding to the number of respiratory events occurring under CPAP treatment divided by the duration of device nightly usage for each patient.

Patients with more than 90 consecutive days of CPAP treatment were considered for the study. Data from nights with less than 2 h of CPAP use were removed from the analysis. A preliminary exploration of the data was carried out to select the longest period with less than 30% of missing rAHI data.

A set of N telemonitoring rAHI time series T_i corresponding to N distinct individuals was available, where each time series was of length l_i (with i an integer in $\{1, 2, .. , N-1, N\}$). The mean absolute value, variability and evolution of the time series, respectively, characterize the average efficiency of CPAP treatment, the stability of CPAP efficacy over time and the evolving nature of response to treatment potentially explained by changes in medical conditions or technical interventions.

The following data specificities restricted the choice of a clustering method: (i) variable-length time series, (ii)

long duration for 25% of the time series (more than 910 data points), (iii) missing values in the time series and (iv) importance of various features including absolute and relative shifts as well as variance.

For a benchmark of time series clustering methods, we referred to the work of Javed et al. [12] who concluded that clustering methods giving acceptable accuracy depend on the characteristics of the dataset. Our dataset was not suitable for the use of classical clustering methods, such as partitional (k-means, k-medoids), density-based (density peaks) or hierarchical (agglomerative, divisive) methods, on the raw time series. Indeed, these methods become intractable with long time series, are not suited to variable-length time series and would require the imputation of missing values. Extensive processing of the data to apply these methods would introduce biases and reduce the data volume. Indeed, filtering, imputing missing values, normalizing magnitude and smoothing were modelling choices whose impact could not be measured by any objective metric since we were working in an unsupervised setting. Besides, considering only fixed-length time series would imply choosing a reasonable length reached by most of the series, and truncating series longer than this threshold, thus reducing the volume of records analysed. Instead, we wanted to keep as close as possible to our real-world data and take advantage of its large volume.

Residual AHI time series are non-deterministic and discrete: they show some random aspects that prevent their behaviour from being described explicitly. The rAHI depends on numerous factors, including unknown or unmeasurable factors, and the physiological nature of their generation is inherently random. Thus, deterministic modelling, such as piece-wise linear regression, would not give a good fit and we instead assumed that the time series were the manifestations of some stochastic (random) processes.

To account for these specificities, Hidden Markov models (HMMs) were employed in a probabilistic model-based approach extracting features from time series of variable length and capturing statistics relative to the magnitude, dispersion and dynamic of the series [13]. HMMs are currently used in medical applications for the analysis of sequential data in various different fields, for example, for biological sequence (proteins or genome) segmentation and alignment [14, 15], and physiological signal segmentation, such as sleep stage classification or human activity recognition [16].

Data analysis

Hidden Markov Model learning and state predictions

Hidden Markov models (HMMs) are generative statistical models used for the modelling of stochastic time-varying processes [17]. A HMM is defined by a vector of initial probabilities, a transition matrix and the emission densities of the hidden states. HMMs are described in the supplementary material.

Model training was performed in the Python programming language using the *pomegranate* package [18]. The HMM parameters were trained from data using an unsupervised approach, the Baum-Welch algorithm [19]. In this study, it was decided to train a homogeneous, or stationary, HMM, in which we assume that the parameters are independent of time t. The transition matrix and state emission densities are thus invariant. We constrained the model to only log-normal state emission distributions and fixed the number of states at 3 (states A, B, and C). In the supplementary material (p 1-2), we describe the method employed to choose and validate the number of states and the emission distribution type. Model selection criteria suggest that a HMM with more than 3 states might better fit the data, but when there are over 3 states, the emission distributions are not clinically interpretable as the difference between distinct states is not clinically significant.

Once the model had been defined, we inferred the most likely sequence of states for each sequence of observations. The sequence of states provides a segmentation of the time series.

Features extraction and cluster number

We trained the HMM to extract information from the signal characterizing the plateaus' levels, its associated variance and duration. In our HMM, the plateaus are represented by the hidden states. The time spent in each state combined with the frequency of transitions between states describes the volatility, or stability, of the rAHI of a given patient.

In the supplementary material (p 3), we describe how the transition matrix specific to each patient summarizes his/her stability and constitutes a set of features extracted from the time series which can be used to compute K-means clustering. The K-means algorithm is an unsupervised method of data partitioning which consists in dividing the dataset into k homogeneous groups. Among the clustering methods the K-means algorithm was the best at separating groups (see supplemental material).

The main challenges when applying unsupervised K-means without any a priori knowledge as to the number of groups are choosing the right number of clusters and assessing the meaningfulness of the partition. In this context, we did not know, a priori, how many distinct types of CPAP telemonitoring trajectories existed; we were ignorant of how many groups we would be able to identify in the dataset. The supplementary material reports details on the method used to fix the number of clusters at 6 (p 4), based on both objective metrics and the clinical relevance of the clusters.

Statistics for evaluation of the relevance of each cluster

To assess the significance of the partition, we computed descriptive statistics regarding several characteristics of the clusters: mean and standard deviation of rAHI, leaks, and nightly usage by the patients in each cluster. Normality among groups was assessed graphically with quantile-quantile plots. Missing adherence values were considered as null values. Comparisons between clusters were performed using one-way ANOVA. Corrections for multiple tests were performed using a conservative Bonferroni correction to compare the mean values for each pair of clusters. In addition, a sensitivity analysis was undertaken to assess the stability of clusters over time and is presented in the supplementary material.

Results

Identification of three states from a Hidden Markov Model (HMM)

From the telemonitoring records of 2860 CPAP-treated patients (age, 66.31 ± 12.92 years; male, 69.9%), HMM estimated the emission distribution of the three distinct states (Figures 1 and 2). The HMM allowed to assess rAHI over time in two complementary domains: variability inside a given state and probability of shifting from one state to another.

The mean and variance of rAHI (log normal emission densities) were related, and the three states corresponded to three different levels of logarithmic mean rAHI and standard deviation: A, 0.43 ± 0.28; B, 1.08 ± 0.30; and C, 2.05 ± 0.62 events per hour. We labelled the three states according to the magnitude of the rAHI mean level and variability: low, medium or high.

Clinical relevance of HMM classification

HMM segmentation provides a visualization dashboard for routine follow-up of CPAP-treated patients. Figure 3 depicts how HMM classification and segmentation allows to visualize rAHI time series using a different colour for each state and transitions during long term CPAP-treatment. We propose the concept of the "telemonitogram" by analogy with the sleep stage "hypnogram".

HMM segmentation allows to separate homogeneous groups of patients exhibiting similar telemonitoring trajectories so as to propose personalized follow-up. Six clusters of telemonitoring trajectories were identified. The first four clusters gathered rather stable patients. The first cluster grouped 669 (23%) stable CPAP-treated patients with very well-controlled OSAS in terms of their rAHI, characterized by a mean rAHI of 0.58 ± 0.59 events per hour. The second, third and fourth clusters assembled 303 (11%), 321 (11%)

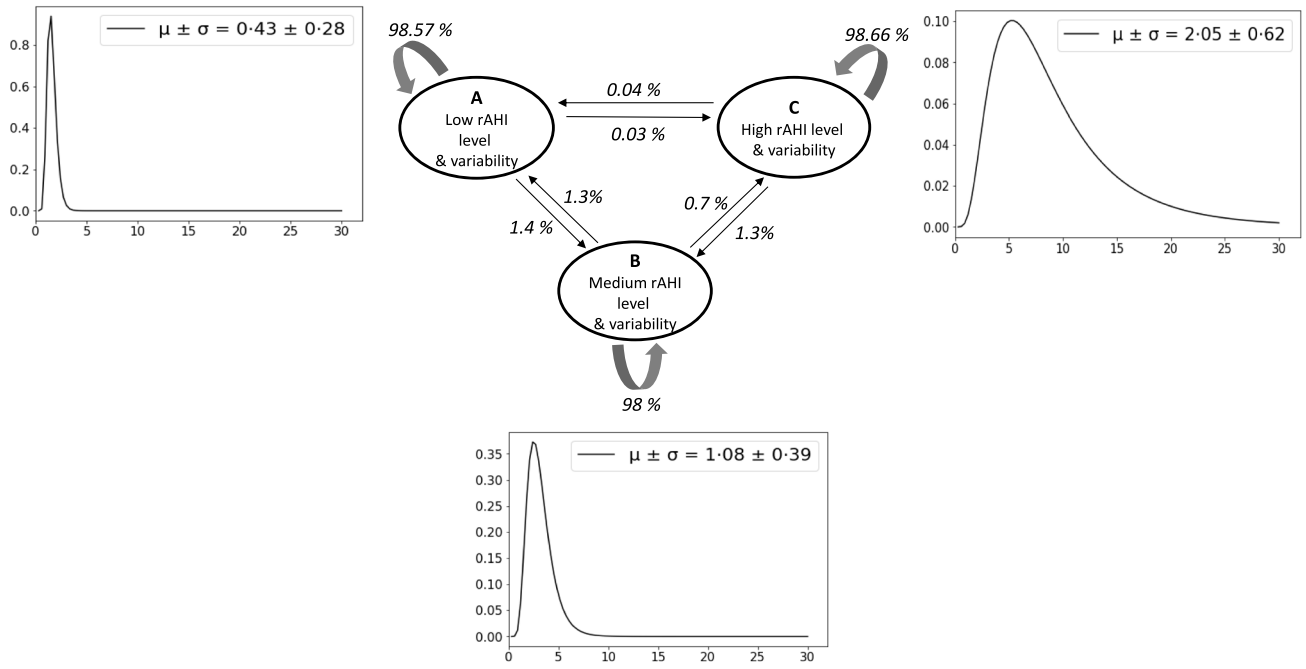

Fig 1 The three HMM states with transition probabilities (in %) and plots of the log-normal emission densities. For example, a patient in state A at day X has only a 0.03% risk of being in state C the following day

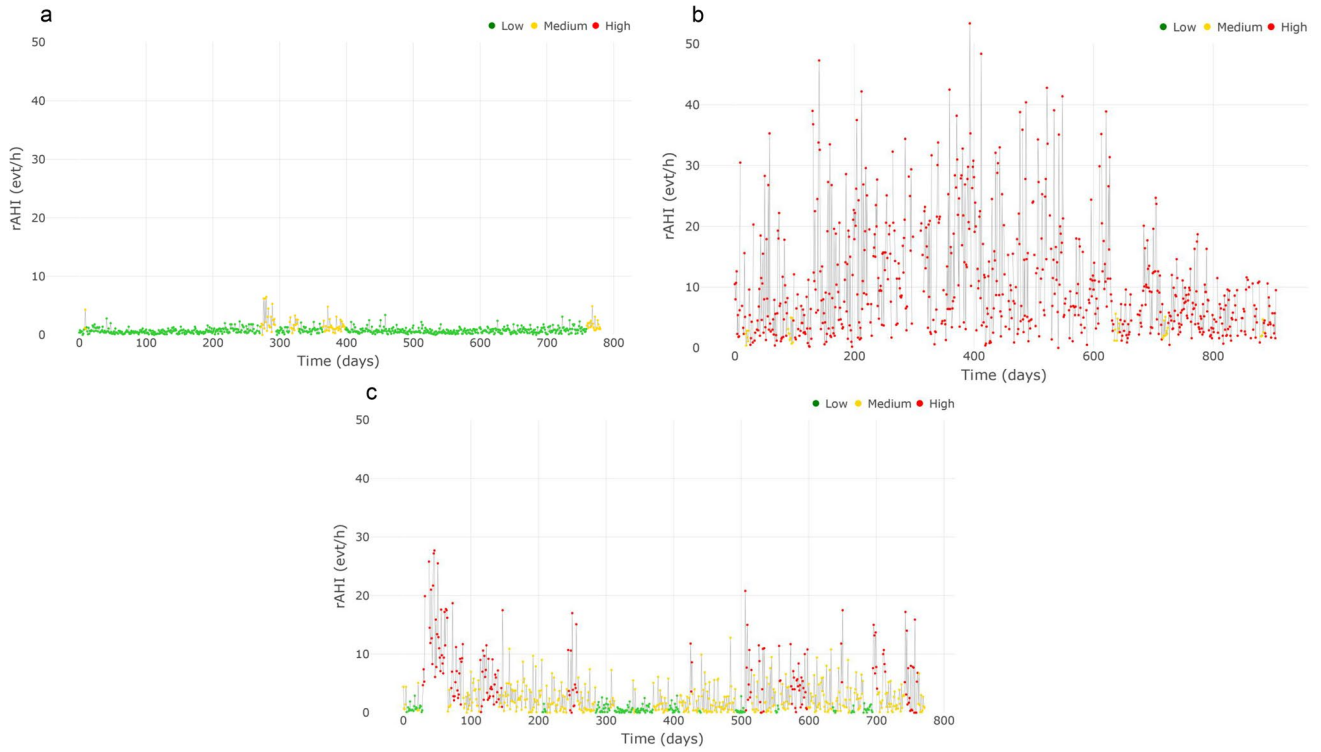

Fig 2 **a** This patient spent most of his CPAP treatment in state A and only transitioned a few times from state A to state B demonstrating a well-controlled and stable reduction of rAHI. **b** This patient spent most of his CPAP treatment in state C, with high absolute rAHI values and variability, and transitioned very few times to states A or B demonstrating an insufficient reduction in rAHI. **c** This last patient spent one-third of his CPAP treatment in each of the three states demonstrating unstable and sometimes insufficient response to CPAP treatment

and 715 (25%) patients, respectively, with mean rAHI of 1.31 ± 1.45, 1.64 ± 1.54 and 2.38 ± 1.54 events per hour, respectively. In contrast, the last two clusters grouped 382 (13%) and 470 (16%) patients exhibiting the most unstable rAHIs, with mean rAHI of 4.9 ± 4.05 and 9.62 ± 5.62 events per hour, respectively. Figure 4 depicts time series of representative patients across the different clusters. Figure 5A is a heatmap showing the burden of state-to-state transitions among clusters of patients on CPAP telemonitoring. Figure 5B presents the relationship between mean rAHI and variability across clusters.

As a comparison, clustering was performed on classical features extracted—descriptive statistics such as mean, standard deviation and trend. The internal validity measures of the resulting partitions indicated that the instances of a same cluster were not necessarily closely related, and the distinct clusters were not well-separated from each other. In addition, the partitions proved less stable and less interpretable than with the selected HMM-then-K-means approach.

CPAP telemonitoring clusters were related to outcomes (Table 1). There was a significant difference in mean rAHI between all the pairs of clusters (P value < 0.01). Mean CPAP adherence was significantly 30 min longer in cluster 0 compared to clusters 4 and 5 (385 versus 347 and 353 min, respectively, P value < 0.01). The mean number of leaks were also significantly greater in cluster 5 compared to clusters 0, 1, 2 and 3 (P value < 0.01).

Discussion

We developed a Hidden Markov model (HMM) approach to extract latent state features revealing, summarizing and visualizing the information contained in rAHI time series generated night after night by CPAP telemonitoring. This probabilistic sequence classification provides the opportunity of going far beyond the mean rAHI usually considered by clinicians and home-care providers to guide the personalized care of CPAP-treated patients. We used these extracted features to propose clinically relevant clusters of rAHI trajectories and demonstrated that the clusters were linked to important outcomes including treatment adherence and leaks.

Hidden Markov models are statistical models that can be used to describe the evolution of observed data (such as, in our case, the daily measurements of rAHI) as a series of outputs generated by one of several (hidden) internal states. HMM have been successfully applied to solve various biological sequence problems particularly in the genomics

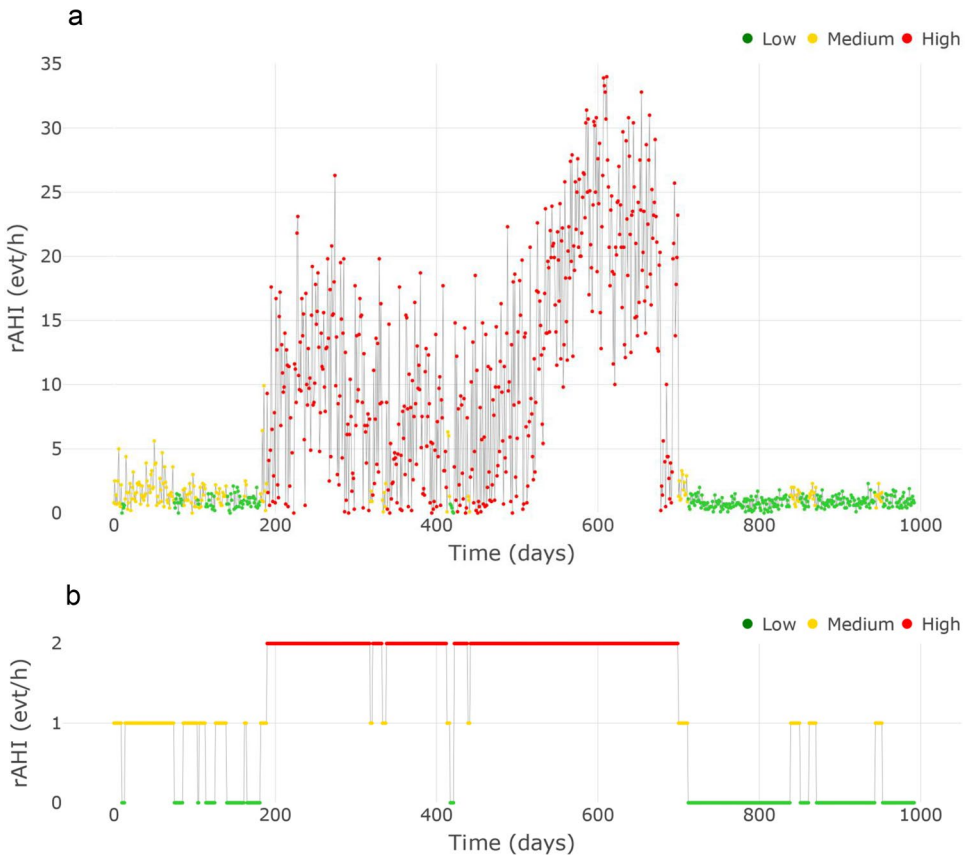

Fig 3 **a** CPAP rAHI telemonitoring data over almost 1000 days coloured according to HMM classification in an exemplary patient. **b** Graphical summary of the states established overtime by the HMM (green, state A; yellow, state B; red, state C): We call this representation a "telemonitogram" by analogy with the sleep stage "hypnogram"

and proteomics field [14, 15]. Feature extraction with the HMM approach has the specific advantage of operating on variable-length time series, which is typically the context encountered during CPAP telemonitoring. Moreover, our feature extraction method reduces the impact of outliers that can hamper other analytical methodologies. Displaying the rAHI time series using HMM allowed us to summarize rAHI as sequences of states and to identify significant transitions so as to prevent further worsening. This is more informative than current dashboards and allows easier interpretation and translation to targeted therapy by caregivers and will also provide opportunities for automation reducing the burden of workflows concerning millions of individuals.

In comparison with current rudimentary reporting of a mean rAHI, we identified features of rAHI variability related to adherence to CPAP. Previous studies with big data analysis have already suggested that a mean rAHI above 5 events/h is associated with a higher risk of CPAP termination [7]. We provide new knowledge by demonstrating that a mean rAHI generally accepted as being in the normal range but associated with high variability is a risk of lower adherence.

Our study had several strengths including a large sample size of 2860 CPAP-treated patients coming from the real-life practice of a regional home-care provider in France. Our study population was unselected and unbiased and representative of a real-life CPAP-treated OSA population. This implies that our findings can probably be widely generalized. Nevertheless, external validation using data from other CPAP databases remains desirable. CPAP-telemonitoring data came from patients with different lengths of time since CPAP initiation. An important additional study will be to prospectively implement HMM analysis from CPAP initiation. The nature of the dataset means that there was a lack of granular data regarding the patients' medical history and data regarding material settings (interface type and pressure settings) was not of sufficient quality to be analysed in this study. Also, the different HMM states might correspond to different sleep apnoea phenotypes and might be linked to cardiovascular comorbidities, but this remains to be explored in further studies [20–22].

Next steps and perspectives in the framework of 3P medicine

CPAP telemonitoring is beginning and expected to have an impact at three distinct levels for sleep and respiratory clinicians [23], predominantly via rapid, personalized, accurate interpretation of CPAP adherence, residual events and leaks; for health systems, by improving workflow and the potential for early detection of acute events [24, 25]; and

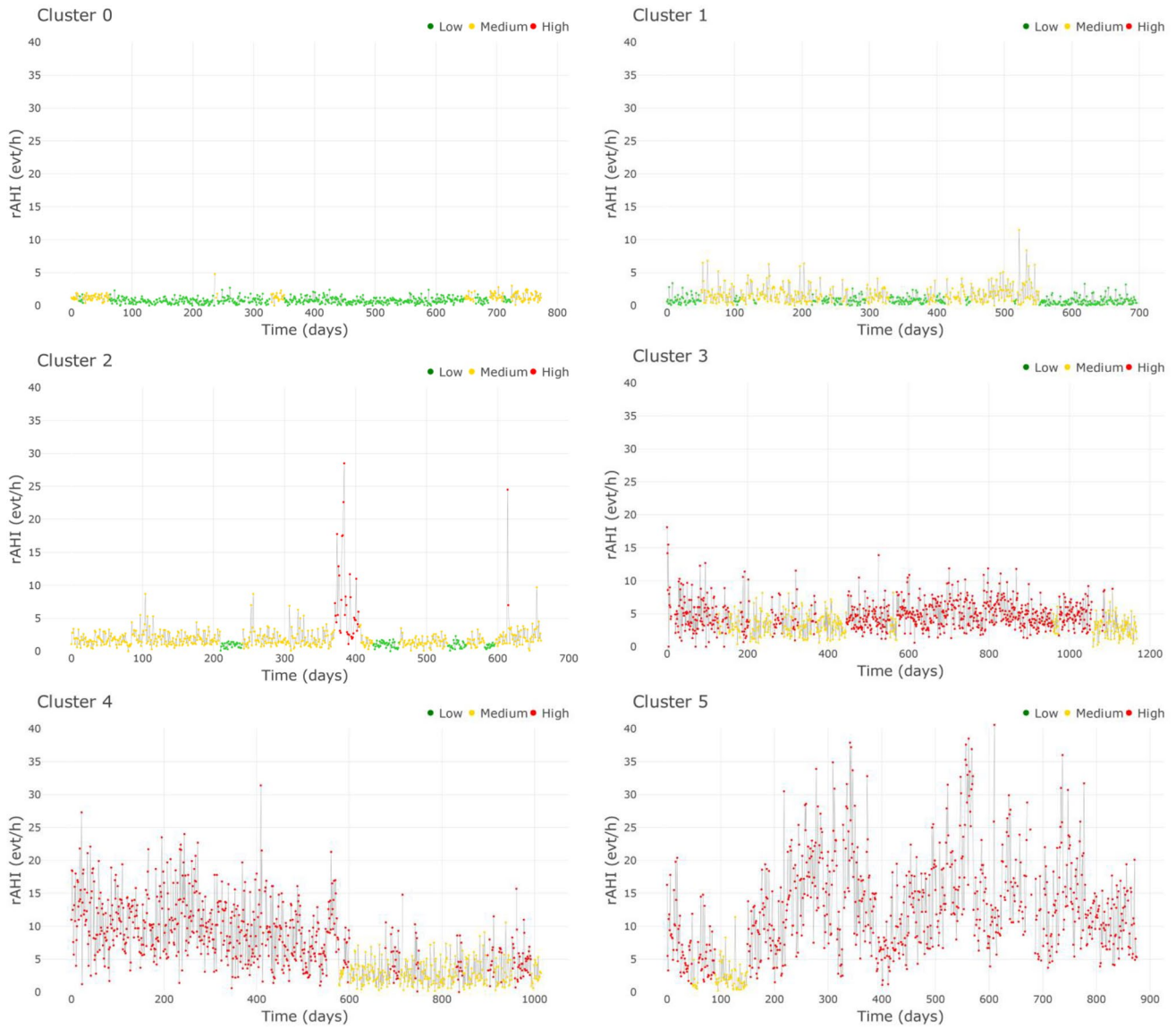

Fig 4 Telemonitoring rAHI time series of representative patients across the different clusters

for patients, by enabling them to process their own data to promote better health [26–29].

Our findings are important and open new avenues for several reasons. In the context of 3P medicine, our subgrouping may help CPAP providers and physicians to focus on difficult-to-treat patients and prompt them to investigate the reasons for CPAP treatment inefficacy or variability in efficacy over time. The number of technical and medical interventions might be cluster dependent. A next step will be to deploy the classifier to predict the affiliation of a given patient to a cluster by continuously updating with the most recent CPAP data. Alerts could be raised when a patient is assigned to a less stable cluster than previously and used to dynamically adapt their treatment follow-up.

The residual AHI is a key parameter of CPAP treatment efficiency but paints an incomplete picture. HMMs merit to be developed for other indices reflecting quality of CPAP care, namely, CPAP usage, leaks and positive airway pressure (PAP) settings. These parameters are inter-related, and there are needs to provide an integrative view of the complete landscape. The HMM states, including their variability, are probably related to underlying comorbidities, especially cardiovascular diseases and OSA characteristics at diagnosis, whereas transitions might reflect interface changes [30], leaks or acute clinical events. Oro-nasal masks are associated with a higher residual AHI and a switch from a nasal pillow or nasal mask might be a triggering event for transition from one state to another, which are currently poorly detected in

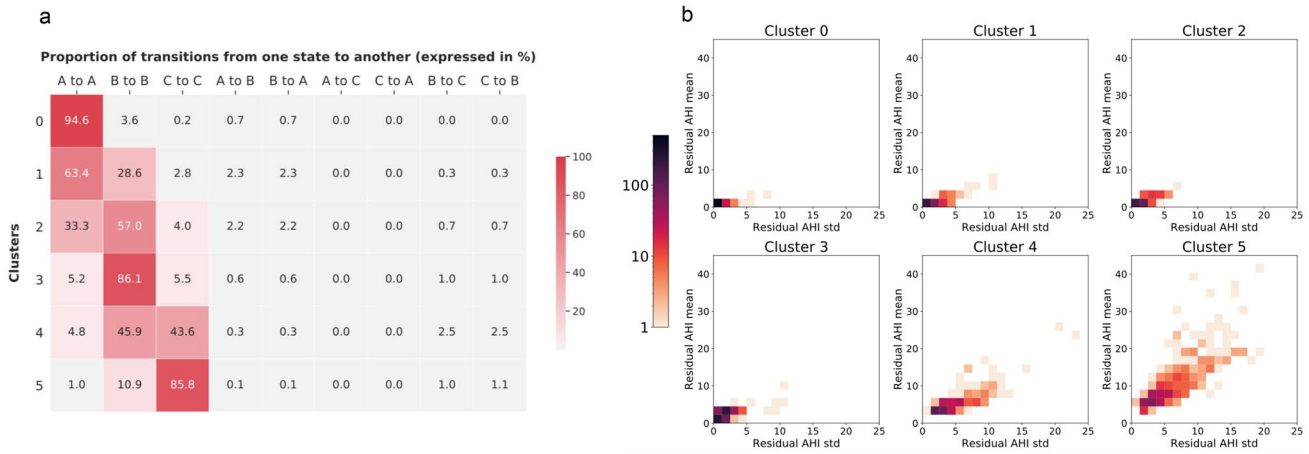

Fig 5 **a** Heatmap of the clusters' centres showing the burden of state transitions among clusters of CPAP telemonitoring. **b** Relationship between mean rAHI level (mean on y-axis) and variability (standard deviation on x-axis) across clusters

Table 1 Descriptive statistics of the 6 clusters. Results are presented as mean ± SD

Cluster	N (%)	Mean age (years)	Sex (M/F)	Mean rAHI (evt/h)	Mean adherence (min)	Mean leaks (L/min)
0	669 (23%)	64·0 ± 11·9	410/259 (61/39 %)	0·58 ± 0·59	385 ± 124	38 ± 5
1	303 (11%)	64·6 ± 13·2	204/99 (67/33 %)	1·31 ± 1·45	364 ± 139	40 ± 7
2	321 (11%)	66·0 ± 12·9	229/92 (71/29 %)	1·64 ± 1·54	364 ± 128	38 ± 6
3	715 (25%)	64·2 ± 12·5	505/210 (71/29 %)	2·38 ± 1·54	365 ± 129	38 ± 5
4	382 (13%)	70·0 ± 12·9	281/101 (74/26 %)	4·9 ± 4·05	347 ± 140	41 ± 9
5	470 (16%)	71·2 ± 13·0	370/100 (79/21 %)	9·67 ± 5·62	353 ± 132	43 ± 9

routine CPAP care [31]. Also, there is potential in the field to leverage CPAP telemonitoring as an early warning system for exacerbations of common chronic diseases such as heart failure or atrial fibrillation [24, 25, 31]. Underlying cardiovascular disease might explain membership of a high variability cluster and decompensation of these chronic diseases certainly led to transition to highly variable clusters with higher rAHIs [32]. To confirm these hypotheses, future studies will explore well-annotated databases with clinical documentation regarding comorbidities, medications and information on technical interventions including mask changes.

Providing subgrouping with an HMM approach will allow physicians to personalize CPAP-treated patients follow-up management pathways and early detection of treatment failure. Dissemination of this innovation additionally requires randomized controlled trials to compare management driven by HMM with current standard care in terms of burden of interventions, costs and improvement in patient reported outcomes.

Conclusion

We propose a new analysis based on Hidden Markov Models that might constitute the backbone for deployment and dissemination of digital health solutions in respiratory medicine improving the interpretation of telemonitoring data from CPAP-treated patients. This method allows to visualize and reveal interesting novel features for guiding personalized and preventive CPAP care and for the development of predictive tools.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13167-021-00264-z.

Acknowledgment We thank Alison Foote (Grenoble Alpes University Hospital, France) for editing the manuscript.

Author contribution AM verified the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. AM, SB, JLP, RLH and MCS conceived and designed the study. AM analysed the data. AM and JCB were involved in data

collection. JCB, RT and SB provided strategic guidance and oversight. AM, SB, JLP, RLH and MCS drafted the manuscript with input from all authors. The final version of the article has been approved by all the authors.

Funding JLP, SB, and RT are supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02) and 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). AM is supported by Probayes and MIAI 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).

Availability of data and material The CPAP telemonitoring database will be available for justified research purposes on request to JCB after publication of the present article and subject to a written agreement.

Code availability NA.

Declarations

Ethics approval The CPAP telemonitoring database was registered and ethically approved by the French C.C.T.I.R.S: N°15.925bis and ethics regulations MR003 N° 1996650v0.

Consent to participate Before beginning CPAP telemonitoring, each patient gave written informed consent that their anonymized data could be used for research.

Conflict of interest JLP reports grants from Air Liquide Foundation, grants and personal fees from Agiradom, grants and personal fees from AstraZeneca, grants from Fisher and Paykel, grants from Mutualia, grants and personal fees from Philips, grants and personal fees from Resmed, grants from Vitalaire, grants from Boehringer Ingelheim, grants from Jazz Pharmaceuticals, grants from Night Balance and grants from Sefam, outside the submitted work. JCB is an employee of AGiR-à-dom, a French home-care provider. RT reports other from AGiR-à-dom and grants from Resmed, outside the submitted work. AM, RLH and MCS report other from AGiR-à-dom, and SB has no conflict of interest directly related to this work.

References

1. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MSM, Morrell MJ, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. Lancet Respir Med. 2019;7(8):687–98.
2. Lévy P, Kohler M, McNicholas WT, Barbé F, McEvoy RD, Somers VK, et al. Obstructive sleep apnoea syndrome. Nat Rev Dis Primers. 2015;1:15015.
3. Frost & Sullivan. Hidden health crisis costing America billions. Underdiagnosing and undertreating obstructive sleep apnea draining healthcare systems. Frost & Sullivan, Mountain View, CA, USA, American Academy of Sleep Medicine; 2016.
4. Cistulli PA, Armitstead J, Pepin J-L, Woehrle H, Nunez CM, Benjafield A, et al. Short-term CPAP adherence in obstructive sleep apnea: a big data analysis using real world data. Sleep Med. 2019;59:114–6.
5. Pépin J-LD, Woehrle H, Liu D, Shao S, Armitstead JP, Cistulli PA, et al. Adherence to positive airway therapy after switching from CPAP to ASV: A Big Data Analysis. J Clin Sleep Med. 2018;14:57–63.
6. Pépin JL, Tamisier R, Hwang D, Mereddy S, Parthasarathy S. Does remote monitoring change OSA management and CPAP adherence? Respirology. 2017;22(8):1508–17.
7. Liu D, Armitstead J, Benjafield A, Shao S, Malhotra A, Cistulli PA, et al. Trajectories of emergent central sleep apnea during CPAP therapy. Chest. 2017;152:751–60.
8. Drager LF, Malhotra A, Yan Y, Pépin J-L, Armitstead JP, Woehrle H, et al. Adherence with positive airway pressure therapy for obstructive sleep apnea in developing versus developed countries: a big data study. J Clin Sleep Med. 2021;17(4):703–9. https://doi.org/10.5664/jcsm.9008.
9. Pack AI. Application of personalized, predictive, preventative, and participatory (P4) medicine to obstructive sleep apnea a roadmap for improving care? Ann Am Thorac Soc. 2016;13(9):1456–67. https://doi.org/10.1513/annalsats.201604-235ps .
10. Bonsignore MR, Baiamonte P, Mazzuca E, Castrogiovanni A, Marrone O. Obstructive sleep apnea and comorbidities: a dangerous liaison. Multidiscip Respir Med. 2019;14:8. https://doi.org/10.1186/s40248-019-0172-9.
11. Sutherland K, Kairaitis K, Yee BJ, Cistulli PA. From CPAP to tailored therapy for obstructive sleep apnoea. Multidiscip Respir Med. 2018;13:44. https://doi.org/10.1186/s40248-018-0157-0 .
12. Javed A, Lee BS, Rizzo DM. A benchmark study on time series clustering. Mach Learn Appl. 2020;1:100001.
13. Sotoodeh M, Ho JC. Improving length of stay prediction using a hidden Markov model. AMIA Jt Summits Transl Sci Proc. 2019;2019:425–34.
14. Krogh A, Brown M, Mian IS, Sjölander K, Haussler D. Hidden Markov Models in computational biology. J Mol Biol. 1994;235:1501–31.
15 Hidden JK. Markov models for protein domain homology identification and analysis. In: Machida K, Liu B, editors. SH2 Domains. Methods in Molecular Biology, vol. 1555. New York: Humana Press; 2017. https://doi.org/10.1007/978-1-4939-6762-9_3 .
16. Attal F, Mohammed S, Dedabrishvili M, Chamroukhi F, Oukhellou L, Amirat Y. Physical human activity recognition using wearable sensors. Sensors. 2015;15(12):31314–38.
17. Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE. 1989;77(2):257–86.
18. Schreiber J. Pomegranate: fast and flexible probabilistic modeling in python. arXiv:171100137 [cs, stat]. 2018; http://arxiv.org/abs/1711.00137. Accessed May 2020
19. Baum LE, Petrie T, Soules G, Weiss N. A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. Ann Math Stat. 1970;41(1):164–71.
20. Bailly S, Grote L, Hedner J, Schiza S, McNicholas WT, Basoglu OK, et al. Clusters of sleep apnoea phenotypes: a large pan-European study from the European Sleep Apnoea Database (ESADA). Respirology. 2021;26(4):378–87.
21. Bailly S, Destors M, Grillet Y, Richard P, Stach B, Vivodtzev I, et al. Obstructive sleep apnea: a cluster analysis at time of diagnosis. PLoS One. 2016;11(6):e0157318.
22. Bailly S, Daabek N, Jullian-Desayes I, Joyeux-Faure M, Sapène M, Grillet Y, et al. Partial failure of CPAP treatment for sleep apnoea: Analysis of the French national sleep database. Respirology. 2020;25(1):104–11.
23. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25:44–56.
24. Schmickl CN, Heckman E, Owens RL, Thomas RJ. The Respiratory signature: a novel concept to leverage continuous positive airway pressure therapy as an early warning system for exacerbations of common diseases such as heart failure. J Clin Sleep Med. 2019;15:923–7.

25. Light M, Orr JE, Malhotra A, Owens RL. Continuous positive airway pressure device detects atrial fibrillation induced central sleep apnoea. Lancet. 2018;392(10142):160.
26. Pépin J-L, Baillieul S, Tamisier R. Reshaping sleep apnea care: time for value-based strategies. Ann Am Thorac Soc. 2019;16:1501–3.
27. 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. Chest. 2018;153(4):843–50.
28. Bhattacharjee R, Benjafield AV, Armitstead J, Cistulli PA, Nunez CM, Pepin JD, et al. Adherence in children using positive airway pressure therapy: a big-data analysis. Lancet Digit Health. 2020;2(2):e94–101.
29. Keenan BT, Schwab RJ. Using the remote monitoring framework to promote adherence to continuous positive airway pressure. Sleep Med Clin. 2020;16(1):85–99. https://doi.org/10.1016/j.jsmc.2020.11.001 .
30. Borel JC, Tamisier R, Dias-Domingos S, Sapene M, Martin F, Stach B, et al. Type of mask may impact on continuous positive airway pressure adherence in apneic patients. PLoS One. 2013;8(5):e64382.
31. Andrade RGS, Viana FM, Nascimento JA, Drager LF, Moffa A, Brunoni AR, et al. Nasal vs oronasal cpap for osa treatment. Chest. 2018;153(3):665–74.
32. Prigent A, Serandour A, Luraine R, Poineuf JS, Bosseau C, Pépin J. Interrelated atrial fibrillation and leaks triggering and maintaining central sleep apnoea and periodic breathing in a CPAP -treated patient. Respir Case Rep. 2020;8(8):e00666. https://doi.org/10.1002/rcr2.666 .