



# Utilisation of machine learning to predict surgical candidates for the treatment of childhood upper airway obstruction

Xiao Liu<sup>1</sup> · Yvonne Pamula<sup>2</sup> · Sarah Immanuel<sup>3,4</sup> · Declan Kennedy<sup>2,5</sup> · James Martin<sup>2</sup> · Mathias Baumert<sup>1</sup> 

Received: 16 March 2021 / Revised: 24 May 2021 / Accepted: 21 June 2021 / Published online: 17 July 2021  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

## Abstract

**Objective** To investigate the effect of adenotonsillectomy on OSAS symptoms based on a data-driven approach and thereby identify criteria that may help avoid unnecessary surgery in children with OSAS.

**Methods** In 323 children enrolled in the Childhood Adenotonsillectomy Trial, randomised to undergo either early adenotonsillectomy (eAT; N = 165) or a strategy of watchful waiting with supportive care (WWSC; N = 158), the apnea-hypopnea index, heart period pattern dynamics, and thoraco-abdominal asynchrony measurements from overnight polysomnography (PSG) were measured. Using machine learning, all children were classified into one of two different clusters based on those features. The cluster transitions between follow-up and baseline PSG were investigated for each to predict those children who recovered spontaneously, following surgery and those who did not benefit from surgery.

**Results** The two clusters showed significant differences in OSAS symptoms, where children assigned in cluster A had fewer physiological and neurophysiological symptoms than cluster B. Whilst the majority of children were assigned to cluster A, those children who underwent surgery were more likely to stay in cluster A after seven months. Those children who were in cluster B at baseline PSG were more likely to have their symptoms reversed via surgery. Children who were assigned to cluster B at both baseline and 7 months after surgery had significantly higher end-tidal carbon dioxide at baseline. Children who spontaneously changed from cluster B to A presented highly problematic ratings in behaviour and emotional regulation at baseline.

**Conclusions** Data-driven analysis demonstrated that AT helps to reverse and to prevent the worsening of the pathophysiological symptoms in children with OSAS. Multiple pathophysiological markers used with machine learning can capture more comprehensive information on childhood OSAS. Children with mild physiological and neurophysiological symptoms could avoid AT, and children who have UAO symptoms post AT may have sleep-related hypoventilation disease which requires further investigation. Furthermore, the findings may help surgeons more accurately predict children on whom they should perform AT.

**Keywords** Sleep apnea · Children · Adenotonsillectomy · Machine learning · Data-driven

## Statement of significance

This study shows that machine learning can help stratify children with obstructive sleep apnea syndrome. We identified previously unreported baseline differences in children who reversed cardiorespiratory symptoms spontaneously and therefore, could avoid surgery. The combination of OAH13, N2 event-free heart period pattern dynamics and N3 event-free thoraco-abdominal asynchrony measurements may help identify children with mild-to-moderate symptoms who do not require surgery.

✉ Xiao Liu  
xiao.liu.au@outlook.com

✉ Mathias Baumert  
mathias.baumert@adelaide.edu.au

Extended author information available on the last page of the article

## Introduction

Between 3 and 15% of children are reported as having upper airway obstruction (UAO) during sleep. [28] Although a broad spectrum of UAO from primary snoring to obstructive sleep apnea syndrome (OSAS) exists, the majority of children are at the milder end of the range. Children with UAO have demonstrated impaired neurocognitive and behavioural function, [12, 18, 29] and increasing evidence suggests they also have altered cardiovascular function [10]. OSAS is considered a key driver of changes in the cardiovascular system [20–22] that may lead to cardiovascular disease later in life. [1, 31, 32] Early detection and treatment of UAO in

childhood may, therefore, reduce cardiovascular morbidity in adulthood.

Contrary to adults, OSAS in otherwise normal healthy children stems most commonly from enlarged tonsils and adenoids. The first-line treatment is, therefore, adenotonsillectomy (AT), which reduces upper airway resistance by removing the enlarged tissue. Although AT has demonstrated positive health outcomes in children with significant OSAS, residual symptoms often exist in some children, and related parental concerns also persist postsurgery [4, 6, 8, 13, 37].

In contrast, the benefits of AT in children with milder UAO remains mostly untested. The landmark Childhood Adenotonsillectomy Trial (CHAT) reported that around 46% of the children had their apnea–hypopnea index (AHI) spontaneously normalise without having AT treatment. [30] The AHI metric is obtained from overnight polysomnography (PSG) and is the current clinical measure of UAO severity and a major determinant in the decision to treat UAO. However, concerns have been raised about the limitations of the AHI as it over-simplifies the spectrum and severity of UAO and correlates poorly with numerous health endpoints [9, 11, 33]. This raises the question of how to identify individuals for whom surgery will be beneficial and those who may recover without surgery, thereby reducing the health care costs and risks of performing AT. [4, 6, 8]

Children with mild UAO generally have a low AHI, and it is currently not known if they would benefit from AT as the AHI only measures the frequency of discrete respiratory events but not necessarily the severity. Furthermore, other markers of UAO, such as increased respiratory load and other abnormal breathing patterns are not quantified in routine PSG. [11] Even though some children enrolled in the CHAT had spontaneously normalised AHI at follow-up without surgical intervention, our previous study found that those children had a relatively lower quality of life after seven months compared to those whose AHI normalised via AT. [25] This illustrates one of the numerous limitations of relying on the AHI as the sole measure of UAO severity.

Our previous studies have found that those children whose AHI-normalised spontaneously in the CHAT had relatively lower cardiac autonomic activation and inspiratory load at baseline than children who's AHI did not resolve spontaneously [24, 25]. Similar findings were shown in studies on normal children in comparison with children with sleep-disordered breathing. [14, 26, 27] Further analysis of the CHAT suggested that the Paediatric Sleep Questionnaire (PSQ) and snoring score could be used to identify children who do not need AT. [7]

To overcome the limitations of the current treatment criteria for childhood UAO, assessment should comprise multiple physiological variables rather than relying on the AHI alone. Using a data-driven approach, children

with similar pathophysiological profiles should cluster into the same group and possibly help identify which children have fewer pathophysiological symptoms thereby possibly not requiring surgery. The emergence of large data sets in the medical field in recent years has seen machine learning increasingly being utilised to aid diagnosis, prognosis, and treatment decisions across a range of clinical disciplines including sleep disorders. [5, 19, 23, 39] This study aimed to investigate the effects of AT for OSAS using a data-driven approach, on the currently used clinical diagnostic marker AHI and our previously identified non-invasive biomarkers which contain additional information. By separating children into two clusters based on their cardiorespiratory characteristics, and analysing the transition of children between clusters from baseline to follow-up, we sought to stratify the intervention outcomes.

## Method

### Study samples

Details of the CHAT protocol have been previously published [35]. All data are publicly available at <https://sleepdata.org/datasets/chat>. In brief, children between 5.0–9.9 years of age with PSG-confirmed OSAS (obstructive apnea–hypopnea index [AHI]  $\geq 2$  events/h or an obstructive apnea index [OAI]  $\geq 1$  events/h), a history of snoring and considered to be surgical candidates for AT were recruited from paediatric sleep centres/sleep laboratories, paediatric otolaryngology clinics, general paediatric clinics and the general community from six clinical centres. Exclusion criteria included comorbidities, medications for psychiatric or behavioural disorders, recurrent tonsillitis, extreme obesity (body mass index  $> 2.99$  for age group and sex-z-score) and severe OSAS (AHI  $\geq 30$  events/h, OAI  $\geq 20$  events/h or oxyhemoglobin saturation  $< 90\%$  for  $> 2\%$  of total sleep time). The study was approved by the Institutional Review Board of each institution. Informed consent was obtained from caregivers, and assent from children  $\geq$  seven years of age. The study was registered at Clinicaltrials.gov (#NCT00560859).

### CHAT interventions

Children were randomly assigned to either early adenotonsillectomy (eAT; surgery within 4 weeks after randomisation) or a strategy of watchful watching with supportive care (WWSC) with a reassessment of all the study variables at approximately seven months. Complete bilateral tonsillectomy and removal of obstructing adenoid tissue were performed using standard surgical techniques.

## Overnight polysomnography

Each child underwent in-laboratory baseline and follow-up PSG carried out by study-certified technicians, following the American Academy of Sleep Medicine paediatric guidelines for both acquisition and scoring. [3] The PSGs were scored centrally by registered sleep technicians. Overnight PSG was repeated approximately 7 months after randomisation. [34, 35]

## Data-driven analysis

The goal of our data-driven analysis (Fig. 1) was to assign all children to one of two clusters (named A and B) based on three variables that we have previously identified as potentially useful descriptors of OSAS symptoms: (1) OAH3 (as derived from the CHAT study), (2) heart period (HP) pattern dynamics in N2 and REM sleep, and (3) thoraco-abdominal asynchrony (TAA, log-transformed) during N3 sleep, (the latter two variables were calculated using original methods [24, 25]. OAH3 represents the number of apneas and hypopneas per hour of sleep associated with a  $> 3\%$  oxygen desaturation. The OAH3 index was transformed due to its nonlinearity using box-cox transformation with lambda at 0.2019. Heart period patterns quantify the degree of cardiac autonomic modulation related to OSAS and were demonstrated to help to identify children who spontaneously normalised their AHI without surgery at baseline. [24] We chose event-free N2 sleep for analysis because it provides more reliable and stable recordings than REM sleep. To measure inspiratory-effort changes due to upper airway obstruction we considered N3 sleep epochs free of respiratory events. [25]

All features were normalised to zero mean and unit variance before further analysis. We used the CHAT follow-up PSG, including children from both arms (WWSC and eAT), for training the classifier, whereas baseline PSG was used as the predicated dataset. All training data Cluster quality was tested using the Silhouette metric. Results are shown in the supplement.

## Cluster definition

The characteristics of the two clusters were determined with the CHAT follow-up dataset using the K-means clustering method, which is a frequently utilised unsupervised machine learning technique [17]. To form two clusters, two points are randomly chosen as the initial cluster centroid. Subsequently, each data point is assigned to its closest cluster, based on the shortest space representation distance to the cluster centroid compared the other cluster centroids. The cluster centroid is recalculated every time a data point was

assigned to a cluster. This process is repeated until the centroid of each cluster no longer changes. We used the squared Euclidean distance to represent space, and the clustering was repeated 20 times to obtain the best clustering results, where a new set of initial centroids was used each time.

## Classification using Linear discriminant analysis

Once the two reference clusters were defined on the follow-up dataset, linear discriminant analysis (LDA) was used to create a classifier model. The LDA is a supervised machine learning method frequently used to reduce the feature space dimension, maximising the difference in means between classes whilst minimising the variance of each class to separate two or more classes by finding a linear combination of predictive features. The LDA model was used to validate the follow-up dataset and classify the baseline dataset.

## Definition of cluster transition classes

Four transition classes were defined according to the cluster change from baseline to follow-up study, which are baseline cluster A to Follow-up cluster B ( $A \rightarrow B$ ), baseline cluster B to follow-up cluster A ( $B \rightarrow A$ ), baseline cluster A to follow-up cluster A ( $A \rightarrow A$ ), and baseline cluster B to follow-up cluster B ( $B \rightarrow B$ ). The cluster transitions were examined to identify children would not need AT surgery.

## Statistical analysis

Anthropometric data were compared by using t-tests and  $X^2$  tests as appropriate. One-way analysis of covariance (ANCOVA) was carried out to investigate the effect of clustering on critical physiological and neurophysiological measures on all available data. ANCOVA was also carried out to investigate the effect of cluster transition classes on physiological and neurophysiological measures for the baseline WW and eAT groups respectively, followed by a Bonferroni test based on Student's t statistic for posthoc comparisons. The measures included in the analysis were: obstructive apnea–hypopnea index ( $\geq 3\%$  desaturation, OAH3), central apnea index all desaturations (CAI), percentage of time  $< 90\%$  oxygen saturation (T90), percentage of total sleep time where end-tidal carbon dioxide  $> 50$  mmHg ( $\text{EtCO}_2 < 50$ ), event-free symbolic HP patterns, event-free TAA, attention and executive function scores, previously reported measures of behaviour, OSAS symptom indicators, paediatric sleep questionnaire scale, global quality of life, and intellectual functioning (DAS-II GCA). Anthropometric variables that were likely to confound statistical analysis (BMI z-score, age, gender, and race) were included in the statistical model as covariates.

## Results

### Subject demographics

A total of 323 children of the original CHAT study who underwent both baseline and follow-up studies and who had all three physiological discriminators and PSG that met the technical requirements were included in this analysis. The dataset comprised 165 children who underwent eAT and 158 children in the WWSC group (Fig. 2). Baseline anthropomorphic characteristics are summarised in Table 1. No significant differences in demographic profiles were observed between both groups. Overall, the mean age of the participants at baseline was 6.6 years and 48% were male. Approximately half (54.5%) of the samples were African American, and 34.4% were obese. Around 4.6% of children were treated with Montelukast and approximately 21.1% received nasal glucocorticoids for rhinitis or asthma at the time of the baseline PSG. At follow up, 83% of subjects in the eAT group no longer had AHI-defined OSAS, i.e. values of  $AHI \leq 2$  and  $OAI \leq 1$ , whilst in the WWSC group 40.5% of children had spontaneous normalisation of AHI scores. Approximately 6.7% of children in the eAT

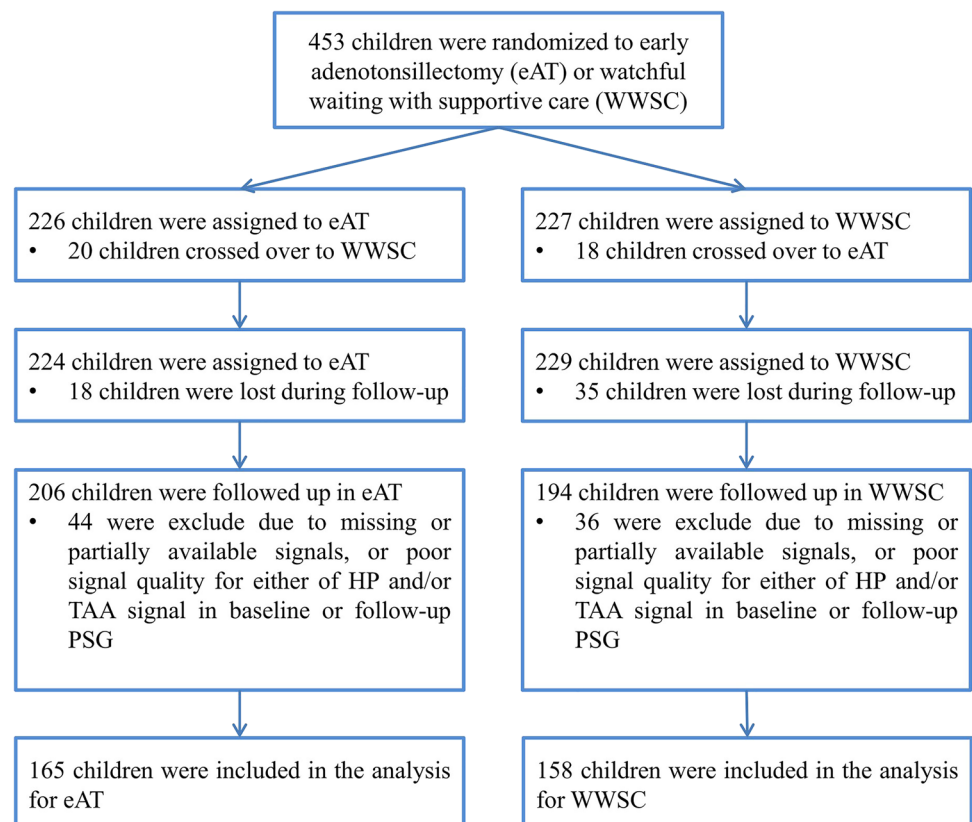
arm and 8.7% in the WWSC arm were on Montelukast and 23.6% (eAT) and 26.6% (WWSC) were on nasal glucocorticoids at the time of the follow-up PSG, representing a small but statistically non-significant increase compared to the baseline sample.

### Data-driven analysis

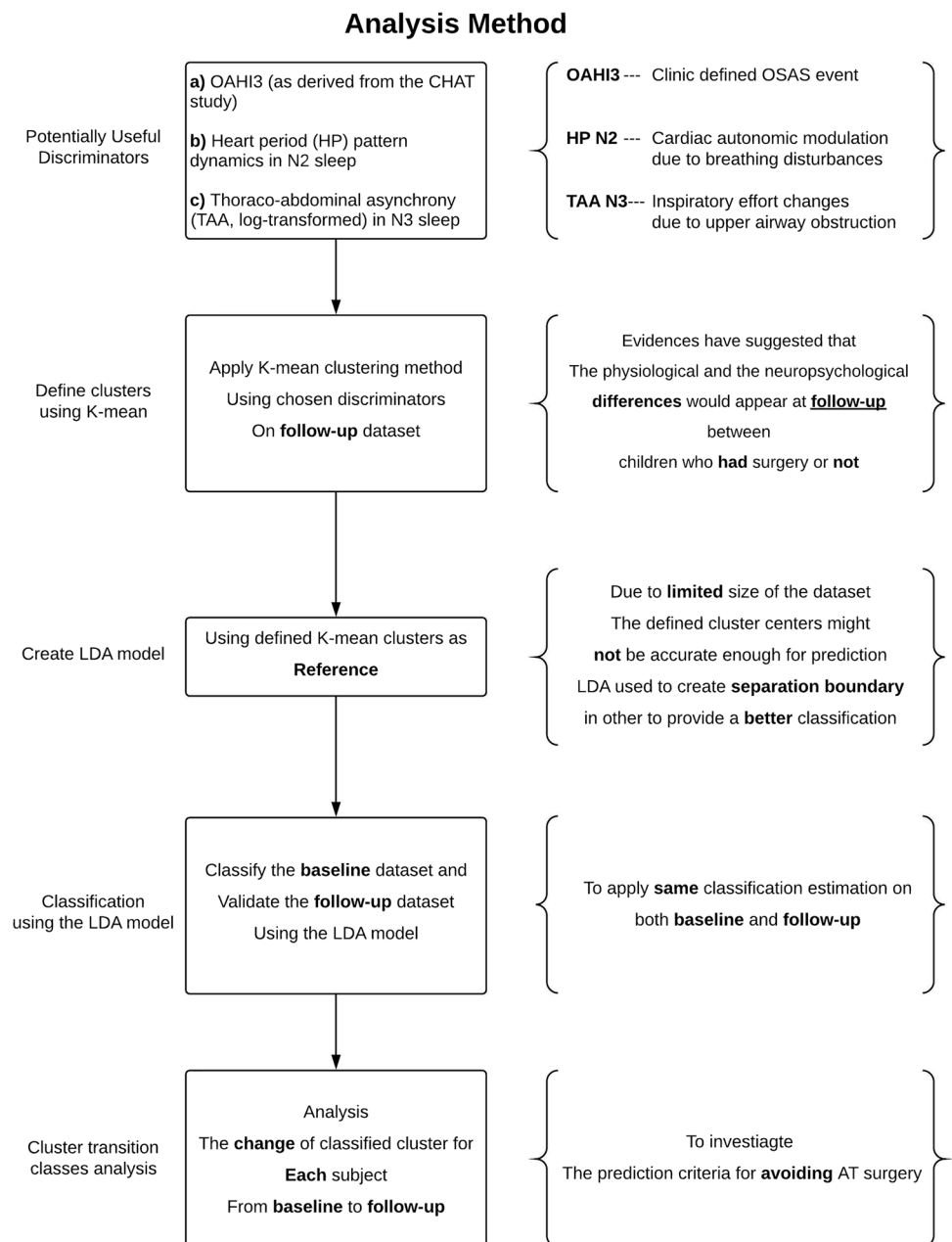
#### Patient clustering using the K-means clustering method

From the follow-up dataset (Table 2), 199 out of 323 children were assigned to cluster A, which has its centroid at OAH13, HP patterns and TAA values of  $-0.484$ ,  $-0.444$ , and  $-0.385$ , respectively. The remaining 124 children were assigned to cluster B whose centroid of OAH13, HP patterns and TAA dimensions was located at  $0.763$ ,  $0.705$ , and  $0.603$ , respectively. Considering the positive effect of surgery, 132 children (80%) in the intervention arm were assigned to cluster A and 33 children (20%) assigned to cluster B. Of the children in the WW arm, around 42.4% (67 children) were part of cluster A and the remaining 57.6% (91 children) were part of cluster B.

**Fig. 1** Data-driven analysis processes flowchart presented vertically; each process module was presented with the key description and purpose horizontally



**Fig. 2** Summary of Childhood Adenotonsillectomy Trial study participants included in the data driving analysis. eAT: early adenotonsillectomy; WWSC: watchful waiting with supportive care; PSG: polysomnography; TAA: thoraco-abdominal asynchrony; HP: Heart rate patterns



### Classification of children using linear discriminant analysis of follow-up PSG

Considering the clustering results of the previous section, linear discriminant analysis was employed to create a function that separates both clusters. Defining the coefficients of classification equation as.

$$Y = -2.2259 \cdot \text{OAH13} - 2.8297 \cdot \text{HP} - 2.3290 \cdot \text{TAA} + 1.3966,$$

where values of  $Y \leq 0$ , result in a given child classified cluster A (otherwise cluster B), yields classification results shown in Fig. 3.

The accuracy of the classification results for the follow-up dataset is:

$$\text{Accuracy} = \frac{TA + TB}{TA + FA + TB + FB} = \frac{199 + 122}{199 + 2 + 122 + 0} = 99.38\%$$

where TA is the true classification of cluster A, FA is the false classification of cluster A, TB is the true classification of cluster B and FB is the false classification of cluster B. The results obtained for the follow-up study with linear discriminant analysis (Table 3) was in agreement with the K-means clustering results.

Eighty-one per cent of children who underwent surgery (134 out of 165 children) were classified into cluster



**Table 1** Subject characteristics at baseline and 7 month follow-up grouped according to study arm

Characteristics	Early Adenotonsillectomy (N = 165)		Watchful Waiting (N = 158)	
	Baseline	Follow-up	Baseline	Follow-up
Age <sup>^</sup> (years)	6.6 ± 1.1	7.2 ± 1.5	6.6 ± 1.4	7.1 ± 1.4
Male sex- N (%)	74 (44.8%)		81 (51.3%)	
Race—N (%) <sup>†</sup>				
African American	86 (52.1%)		90 (57.0%)	
Caucasian	60 (36.4%)		53 (33.5%)	
Other	19 (11.5%)		15 (9.5%)	
BMI z score <sup>^</sup>	0.91 ± 1.35	1.19 ± 1.21	0.88 ± 1.23	1.03 ± 1.26
Weight Class—N (%) <sup>‡</sup>				
Overweight (BMI ≥ 85th percentile)—N (%)	85 (51.5%)	93 (56.4%)	74 (46.8%)	85 (53.8%)
Obese (BMI ≥ 95th percentile)—N (%)	58 (35.2%)	69 (41.8%)	53 (33.5%)	58 (36.7%)
Montelukast—N (%)	6 (3.6%)	11 (6.7%)	9 (5.7%)	14 (8.7%)
Glucocorticoids—N (%)	34 (20.6%)	39 (23.6%)	34 (21.5%)	42 (26.6%)

<sup>^</sup> Data are presented as mean ± SD<sup>†</sup> Race reported by caregivers<sup>‡</sup> Overweight was defined as a body-mass index in the 85th percentile or higher, obese as a BMI in the 95th percentile or higher

A and 19% (31 out of 165) into cluster B (Tables 2 and 3). Furthermore, comparing linear discriminant analysis results with the current clinical marker of OSAS using AHI normalisation, about 84% (168 out of 201) of normalised children were in cluster A and 73% (89 out of 122) of children were considered as not normalised were in cluster B.

Classification of the baseline PSG using linear discriminant analysis.

The linear discriminant analysis was applied to the baseline dataset using the classification equation created by using follow-up dataset (Figs. 3 and 4). Around 2/3 (208 out of 323) of children were classified as cluster A, whilst only about 1/3 (115 out of 323) were classified as cluster B (Table 3). Similar classified results were obtained in both watchful waiting and surgery groups, 106 children are considered as cluster A in WW group, and 102 children in the eAT group. For cluster B, 52 out of 158 were in the WW group, which is over 42% less than the follow-up result, and 63 out of 165 children were in the eAT group.

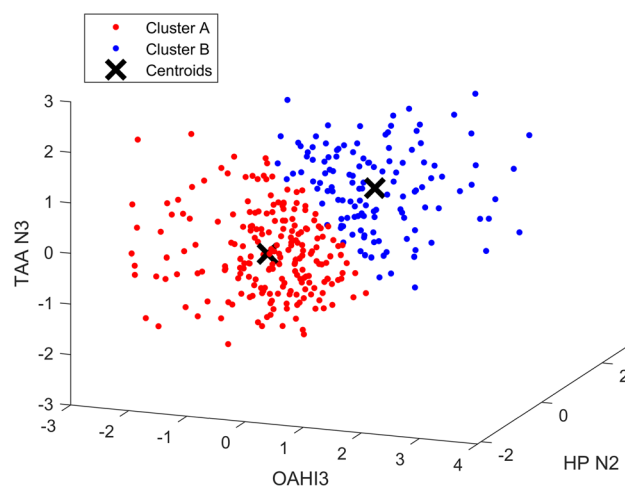
**Table 2** Cross-tabulation of classification result from K-means clustering vs. factor study arm at follow-up

Cluster	Study arm		Total
	WW	eAT	
A	67	132	199
B	91	33	124
Total	158	165	323

### Effect of the clusters on physiological and neurophysiological measurements

To explore the physiological and the neuropsychological differences between the two clusters, one-way ANCOVA was applied to baseline and follow-up data for the two clusters, respectively (Tables 4 and 5). Of the 646 PSGs recorded, 514 sleep studies contained complete sets of physiological and neuropsychological measurements; 256 at baseline and 258 at follow-up.

At baseline (Table 4), significant differences occurred between clusters in OAH13, RPCTCO2G50, PCTLT90,

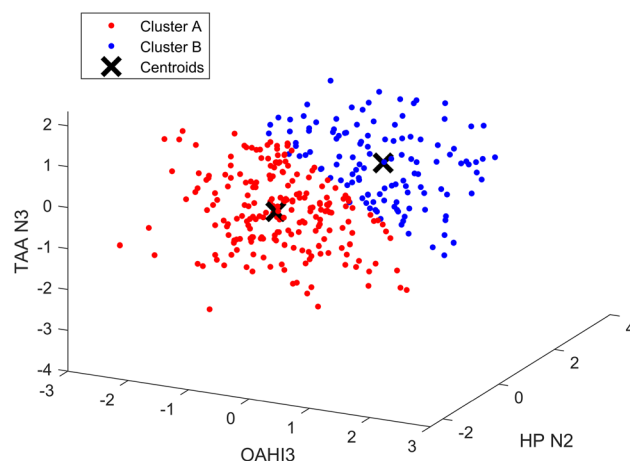
**Fig. 3** Clusters separation by linear discriminant analysis for the follow-up dataset

**Table 3** Cross-tabulation of classification results obtained with linear discriminant analysis for study arm at baseline and follow-up, and normalisation at follow-up, respectively

Cluster	Baseline		Follow-up			
	Study arm		Study arm		Normalisation	
	WW	eAT	WW	eAT	Not normalised	Normalised
A	106	102	67	134	33	168
B	52	63	91	31	89	33

symbolic heart rate patterns and TAA in all sleep stages (N2, N3, and REM). Considering that the clusters were defined by OAH13, TAA, and symbolic heart rate patterns, differences in these variables are to be expected. Of the covariates included in the models, age was associated with symbolic heart rate pattern in N3 sleep; male gender showed a significant effect on TAA in N3 sleep; generalised intellectual functioning was significantly influenced by race and BMIZ; moreover, BMIZ significantly affected OAH13, PCTLT90, the total score of the paediatric sleep questionnaire, and TAA in N2 and N3 sleep.

Similar significant differences between clusters were found at follow-up (Table 4), except for TAA in REM sleep. Additionally, questionnaire-based OSAS symptom indicators and the total score of the paediatric sleep questionnaire showed significant differences between clusters, where cluster B has higher scores than cluster A. Of the covariates included in the models, age was associated with symbolic heart rate pattern in all sleep stages (N2, N3, and REM) and TAA in REM sleep; both male gender and race showed a significant effect on the total score of the paediatric sleep questionnaire; male gender also had an impact on the central apnoea index (CAIOP); generalised intellectual functioning was significantly influenced by race and BMIZ.

**Fig. 4** Clusters separation by linear discriminant analysis for the baseline dataset

## Cluster transition analysis

### Cluster transition and the effect of surgery

Between baseline to follow-up, 62 children transitioned from A → B, 55 children from B → A, 146 children from A → A and 60 children from B → B (Table 5). Considering both treatment arms individually, 47 children in the WW arm and 15 in the eAT arm transitioned from A → B, whilst 8 children in the WW arm and 47 children in the eAT arm transitioned from B → A. A further 59 children in the WW arm and 87 children in the eAT arm remained in cluster A (A → A), whilst 44 children in the WW arm and 16 children in the eAT arm remained in cluster B (B → B).

### Relationship between cluster transition classes and physiological and neurophysiological variables at baseline in the WW arm

Considering the baseline characteristics in the WW arm (Table 6), 118 out of 158 children had complete data on neurophysiological measurements (37 transitioning from A → B, five from B → A, 45 from A → A, and 31 from B → B). Significant difference between all classes were observed in OAH13, OMAH13, the extent of oxygen desaturation, Conners' GI Restless—Impulsive T-Score, Conners' GI Emotional Liability T-Score, the Behaviour Rating Inventory of Executive Function (BRIEF), symbolic heart rate pattern in all sleep stages (N2, N3, and REM), TAA in NREM sleep stages (N2 and N3). None of the covariates included in the statistical models was significantly associated with any of the above measurements.

Post-hoc comparisons showed that children transitioning from B → A had significantly higher Conners' GI Restless—Impulsive T-Score compared to other classes (B → A vs. A → A:  $p=0.002$ ; B → A vs. B → B:  $p=0.010$ ; B → A vs. A → B:  $p=0.008$ ), Conners' GI Emotional Liability T-Score (B → A vs. A → A:  $p=0.0004$ ; B → A vs. B → B:  $p=0.0005$ ; B → A vs. A → B:  $p=0.0005$ ), Behaviour Rating Inventory of Executive Function (BRIEF) (B → A vs. A → A:  $p=0.0078$ ; B → A vs. B → B:  $p=0.011$ ; B → A vs. A → B:  $p=0.015$ ), TAA in sleep stage N2 (B → A vs. A → A:  $p<0.0001$ ; B → A vs. B → B:  $p=0.038$ ; B → A vs. A → B:  $p=0.0005$ ) and TAA

**Table 4** Comparison of physiological and neuropsychological measurements between clusters considering all baseline and follow-up data

Measurements	Baseline			Follow-up		
	Cluster A	Cluster B	p-value	Cluster A	Cluster B	p-value
	(N = 169)	(N = 87)		(N = 165)	(N = 93)	
OAH13	3.46 ± 3.06	9.42 ± 6.50	< 0.00001	0.79 ± 1.01	6.68 ± 9.57	< 0.00001
CAI	0.91 ± 0.88	1.19 ± 1.82	0.1272	0.87 ± 0.93	0.81 ± 0.97	0.7547
EtCO <sub>2</sub> > 50	5.82 ± 14.75	16.35 ± 22.89	0.00003	6.09 ± 13.35	13.82 ± 22.32	0.0007
T90	0.03 ± 0.20	0.20 ± 0.43	0.00003	0.00 ± 0.02	0.17 ± 0.64	0.0012
attention and executive function score	100.9 ± 15.27	103.9 ± 14.78	0.0790	109.5 ± 15.17	107.3 ± 16.26	0.2085
CI_Restless_T score	53.09 ± 11.07	53.47 ± 11.09	0.6499	51.96 ± 11.23	51.31 ± 9.44	0.6226
CI_Emotional_T score	49.98 ± 11.30	48.55 ± 10.18	0.3897	48.58 ± 10.37	46.60 ± 7.53	0.1199
Global Executive Composite T score	50.11 ± 10.74	48.92 ± 10.30	0.4843	48.71 ± 11.84	47.74 ± 10.15	0.5018
The total score of the Paediatric Sleep Questionnaire	0.49 ± 0.18	0.51 ± 0.18	0.2845	0.29 ± 0.21	0.42 ± 0.22	< 0.00001
Paediatric Quality of Life Inventory Parent Total Scale Score	77.99 ± 15.49	79.03 ± 15.91	0.7770	82.25 ± 14.64	81.00 ± 15.28	0.5490
Total score of the OSA-18	2.97 ± 1.00	3.11 ± 1.13	0.2915	2.05 ± 1.02	2.58 ± 1.15	0.0002
Differential Ability Scales II	94.93 ± 10.81	97.16 ± 10.61	0.1401	97.22 ± 10.66	97.83 ± 12.74	0.4877
TAA (N2)	2.91 ± 0.63	3.53 ± 0.63	< 0.00001	2.74 ± 0.64	3.37 ± 0.70	< 0.00001
TAA (N3)	2.88 ± 0.70	3.68 ± 0.70	< 0.00001	2.73 ± 0.73	3.59 ± 0.79	< 0.00001
TAA (REM)	3.48 ± 0.68	3.76 ± 0.57	0.0014	3.35 ± 0.63	3.47 ± 0.68	0.1468
HP (N2)	16.11 ± 5.89	23.50 ± 5.47	< 0.00001	12.89 ± 5.08	20.03 ± 4.79	< 0.00001
HP (N3)	14.10 ± 6.26	21.10 ± 6.19	< 0.00001	10.96 ± 5.22	18.33 ± 5.40	< 0.00001
HP (REM)	14.90 ± 4.47	19.03 ± 5.09	< 0.00001	13.09 ± 3.92	17.10 ± 4.04	< 0.00001

Data are presented as mean ± SD. All *p*-values have been obtained using one-way ANCOVA adjusted for likely confounding factors of age (5 to 10 years of age), race (black, white and other), BMI z-score and gender. *OAH13* – obstructive apnea hypopnea (> = 3% desaturation) index; *CAI* – central apnea index all desaturations; *EtCO<sub>2</sub> > 50* – percentage of total sleep time where end-tidal carbon dioxide > 50 mm Hg; *T90* – Percentage of time < 90% oxygen saturation; *CI* – Connor inventory

in sleep stage N3 (B → A vs. A → A: *p* < 0.0001; B → A vs. A → B: *p* = 0.0004).

Children who remained in cluster A had significant lower OAH13 (*p* < 0.0001), TAA in sleep stage N2 (*p* = 0.0025), and N3 (*p* < 0.0001), and heart rate pattern in all three sleep stages (N2: *p* < 0.0001; N3: *p* < 0.0001; REM: *p* = 0.0004) compared to children who remained in B.

### Relationship between cluster transition classes and physiological and neurophysiological variables at baseline in the eAT arm

Considering the baseline characteristics in the eAT arm (Table 7), 138 out of 165 children had data on

neurophysiological measurements (10 children transitioned from A → B, 39 children from B → A, 77 children from A → A and 12 children from B → B). Significant differences between all classes were observed in OAH13, OMAH13, the extent of oxygen desaturation, peak end-tidal carbon dioxide, symbolic heart rate pattern and TAA in all sleep stages (N2, N3, and REM). Of the covariates included in the models, age was associated with TAA in N3 stage, symbolic heart rate pattern in N2, N3, and REM sleep; TAA in N3 stage was significantly associated with race; and BMI z-score affected OAH13, the extent of oxygen desaturation, and TAA in N2 and N3 sleep stages.

Post-hoc comparisons showed that children who remained in cluster B had significant higher values of peak end-tidal carbon dioxide compared to all other children (B → B vs A → A: *p* < 0.0001; B → B vs B → A: *p* = 0.0067; B → B vs A → B: *p* = 0.0086). Additionally, those children had significantly higher OAH13 values (*p* < 0.022), TAA in sleep stage N2 (*p* = 0.0008) and N3 (*p* < 0.0038), and monotonous heart rate pattern in all three sleep stages (N2: *p* < 0.0001; N3: *p* < 0.0007; REM: *p* = 0.0023) compared to children who remained in the A cluster.

**Table 5** Cross-tabulation of cluster transitions from baseline to follow-up vs study arm

Transition	Study arm		Total
	WW	eAT	
A → B	47	15	62
B → A	8	47	55
A → A	59	87	146
B → B	44	16	60
Total	158	165	323



**Table 6** Comparison of physiological and neurophysiological measurements at baseline for all transition classes in the watchful-waiting arm

Measurements	A → B (N = 37)	B → A (N = 5)	A → A (N = 45)	B → B (N = 31)	p-value
OAH13	3.67 ± 2.41	5.80 ± 2.04	3.51 ± 3.48	10.13 ± 6.81	< 0.00001
CAI	0.96 ± 0.66	0.88 ± 0.64	0.70 ± 0.69	1.10 ± 0.84	0.1041
EtCO <sub>2</sub> > 50	7.68 ± 14.67	14.56 ± 31.55	4.17 ± 14.68	12.26 ± 20.03	0.2171
T90	0.01 ± 0.02	0.00 ± 0.00	0.04 ± 0.25	0.16 ± 0.37	0.0497
attention and executive function score	100.0 ± 15.36	98.40 ± 15.92	100.7 ± 15.42	103.2 ± 12.33	0.9713
CI_Restless_T_score	52.03 ± 10.08	67.00 ± 14.37	51.04 ± 9.10	52.10 ± 9.58	0.0053
CI_Emoional_T_score	46.81 ± 11.30	64.20 ± 14.74	46.98 ± 7.32	46.29 ± 7.02	0.0007
Global Executive Composite T score	48.32 ± 8.78	62.80 ± 13.37	48.07 ± 10.27	47.58 ± 9.76	0.0128
The total score of the Paediatric Sleep Questionnaire	0.50 ± 0.15	0.65 ± 0.19	0.47 ± 0.18	0.48 ± 0.19	0.1990
Paediatric Quality of Life Inventory Parent Total Scale Score	78.27 ± 13.87	70.88 ± 20.98	79.44 ± 14.81	80.48 ± 15.51	0.6703
Total score of the OSA-18	3.01 ± 0.98	3.61 ± 1.53	2.85 ± 1.01	3.10 ± 1.18	0.3815
The Differential Ability Scales II (DAS), a measure of generalised intellectual functioning	93.27 ± 9.60	89.60 ± 8.26	93.84 ± 10.58	97.55 ± 12.69	0.4692
TAA (N2)	3.02 ± 0.62	4.24 ± 0.77	2.86 ± 0.69	3.41 ± 0.59	< 0.00001
TAA (N3)	2.97 ± 0.68	4.33 ± 0.81	2.85 ± 0.75	3.63 ± 0.66	< 0.00001
TAA (REM)	3.51 ± 0.68	4.08 ± 0.52	3.45 ± 0.62	3.51 ± 0.64	0.2541
HP (N2)	18.20 ± 5.51	21.50 ± 5.31	15.16 ± 6.20	23.48 ± 4.53	< 0.00001
HP (N3)	16.53 ± 5.65	18.46 ± 5.92	12.90 ± 6.30	21.23 ± 5.54	< 0.00001
HP (REM)	16.22 ± 4.25	18.73 ± 6.23	13.91 ± 4.30	18.51 ± 5.02	0.0003

Data are presented as mean ± SD. All *p*-values have been obtained using one-way ANCOVA adjusted for likely confounding factors of age (5 to 10 years of age), race (black, white and other), BMI z-score and gender. *OAH13* – obstructive apnea hypopnea (> = 3% desaturation) index; *CAI* – central apnea index all desaturations; *EtCO<sub>2</sub> > 50* – percentage of total sleep time where end-tidal carbon dioxide > 50 mm Hg; *T90* – Percentage of time < 90% oxygen saturation; *CI* – Connor inventory

**Table 7** Comparison of physiological and neurophysiological measurements at baseline for all transition classes in the eAT arm

Measurements	A → B (N = 10)	B → A (N = 39)	A → A (N = 77)	B → B (N = 12)	p-value
OAH13	4.17 ± 3.70	9.78 ± 7.23	3.24 ± 3.03	7.88 ± 3.39	< 0.00001
CAI	0.66 ± 0.49	1.30 ± 2.58	1.05 ± 1.07	1.21 ± 0.91	0.7443
EtCO <sub>2</sub> > 50	8.84 ± 18.45	14.15 ± 20.70	5.49 ± 14.46	34.80 ± 26.92	0.00002
T90	0.00 ± 0.00	0.30 ± 0.53	0.03 ± 0.22	0.06 ± 0.12	0.0012
attention and executive function score	100.5 ± 14.73	105.3 ± 15.33	101.5 ± 15.48	103.8 ± 19.18	0.3546
CI_Restless_T_score	51.90 ± 11.33	53.51 ± 11.39	54.95 ± 12.37	51.25 ± 9.70	0.5702
CI_Emoional_T_score	51.80 ± 14.97	48.97 ± 11.03	53.01 ± 12.01	46.50 ± 6.72	0.1262
Global Executive Composite T score	49.70 ± 14.33	48.18 ± 9.81	52.22 ± 11.15	49.00 ± 8.79	0.1758
The total score of the Paediatric Sleep Questionnaire	0.44 ± 0.15	0.52 ± 0.18	0.51 ± 0.19	0.51 ± 0.13	0.4790
Paediatric Quality of Life Inventory Parent Total Scale Score	78.91 ± 19.82	78.96 ± 16.39	76.90 ± 16.22	78.92 ± 14.10	0.8869
Total score of the OSA-18	2.93 ± 1.13	3.07 ± 1.14	3.02 ± 1.01	3.08 ± 0.81	0.9724
Differential Ability Scales II	95.20 ± 14.98	97.95 ± 8.78	96.34 ± 10.92	96.75 ± 10.97	0.8161
TAA (N2)	2.99 ± 0.40	3.52 ± 0.57	2.88 ± 0.63	3.55 ± 0.71	< 0.00001
TAA (N3)	2.79 ± 0.39	3.69 ± 0.68	2.87 ± 0.71	3.54 ± 0.72	< 0.00001
TAA (REM)	3.43 ± 0.75	3.87 ± 0.49	3.48 ± 0.71	3.91 ± 0.48	0.0040
HP (N2)	19.09 ± 5.10	23.46 ± 6.00	15.27 ± 5.69	24.51 ± 6.36	< 0.00001
HP (N3)	17.81 ± 5.80	21.22 ± 6.49	13.16 ± 6.17	21.51 ± 7.35	< 0.00001
HP (REM)	15.81 ± 3.33	19.15 ± 4.53	14.72 ± 4.68	20.05 ± 6.85	< 0.00001

Data are presented as mean ± SD. All *p*-values have been obtained using one-way ANCOVA adjusted for likely confounding factors of age (5 to 10 years of age), race (black, white and other), BMI z-score and gender. *OAH13* – obstructive apnea hypopnea (> = 3% desaturation) index; *CAI* – central apnea index all desaturations; *EtCO<sub>2</sub> > 50* – percentage of total sleep time where end-tidal carbon dioxide > 50 mm Hg; *T90* – Percentage of time < 90% oxygen saturation; *CI* – Connor inventory

Children who transitioned from B → A had a significant higher values in OAH13 ( $p < 0.0001$ ), OMAH13 ( $p < 0.0001$ ), the extent of oxygen desaturation ( $p = 0.0011$ ), TAA in all three sleep stages (N2:  $p < 0.0001$ ; N3:  $p < 0.0001$ ; REM:  $p = 0.010$ ), and heart rate pattern in all three sleep stages (N2:  $p < 0.0001$ ; N3:  $p < 0.0001$ ; REM:  $p = 0.0001$ ) compared to children who transitioned from A → B.

## Discussion

Our data-driven analysis of the CHAT dataset confirms that children with OSAS can benefit from AT. By separating the entirety of follow-up study into 2 clusters (A, B) based on their AHI and respiratory effort and ANS activation, significant differences in OSAS clinical indicators and sleep quality measures were observed, where children in cluster A showed fewer physiological and neurophysiological symptoms compared to those in cluster B. Additionally, we found that the majority of children were assigned to cluster A at baseline, despite being diagnosed with OSAS based on their AHI. Analysis of cluster transitions from baseline to follow-up PSG demonstrated that surgery (Table 5) could reverse symptoms (47 out of 165 children) or preventing symptoms (87 out of 165 children) from worsening. Additionally, children who had residual UAO despite AT in terms of physiological and neurophysiological symptoms had significantly higher peak end-tidal carbon dioxide at baseline. Children who spontaneously reversed PSG indices of UAO (i.e., AHI) had worse ratings on behaviour and emotion scores.

Machine-learning can reveal new patterns in large, complex datasets and uncover hidden relationships by analysing non-linear associations amongst multiple variables. A study has suggested that the spectrum of children with sleep-disordered breathing can be classified into six unique classes using machine learning [36]. In our study, each follow-up PSG was assigned to one of two clusters, solely based on the three features. The three chosen features, i.e. OAH1, N2 sleep stage heart period dynamics and N3 sleep stage respiratory event-free TAA, may represent pathophysiological domains in children critically affected by OSAS. As these three features were calculated from entire recordings that included respiratory events as well from segments that excluded respiratory events, they capture complementary information about OSAS. [24–26] Statistical comparison of all other clinical variables between the two clusters enabled a more comprehensive characterisation of both groups of children and helped to define their pathophysiological profile.

Previous studies of the CHAT data have demonstrated the benefit of AT for children with OSAS. Children who had surgery demonstrated fewer pathophysiological symptoms

compared to their baseline PSG and the follow-up PSG of children in the WW arm considering AHI, [30] heart rate, heart rate variability, [2, 24] respiratory rate, [15] respiratory effort. [25] Therefore, the follow-up PSG data provides an ideal training set for machine learning to create our classification model, which may overcome limitations of using the AHI as the only dimension of OSAS.

Our study showed that children in cluster A had fewer discrete OAH events during sleep, showed less respiratory effort and fewer autonomic activations during the respiratory-event free sleep period compared with children in cluster B. These findings may suggest children in cluster A have a milder form of OSAS despite being clinically diagnosed with OSAS based on their AHI and thus, may be more likely to show spontaneous resolution of AHI. [10, 14, 26].

Our previous studies reported fewer pathophysiological symptoms at baseline PSG in children who spontaneously normalised their AHI [24, 25]. This further questions the validity of the AHI as the current gold standard for defining clinically significant OSAS in children, which may over-diagnose OSAS in paediatric populations due to the low cut-off that is typically applied. Since our data-driven analysis included the AHI, we observed a positive correlation between AHI normalisation and data classification at follow-up (Table 3), where the majority of AHI-normalised children were assigned to cluster A. In contrast, Cluster B contained most of the children whose AHI did not normalise. However, approximately 20% of the children categorised into the other clusters reflect the additional information gained by heart rate patterns and TAA analysis. Furthermore, almost 2/3 of children were classed as A using baseline data (Table 3), demonstrating comparably fewer pathophysiological symptoms despite meeting the diagnostic (AHI) criteria for OSAS.

Considering the number of children classified into each cluster at baseline and follow-up, we found a similar ratio (2:1) of children in cluster A vs. cluster B in the surgery and watchful waiting group at baseline (Table 3), reflecting the randomisation of the trial. However, follow-up results (Table 3) showed an increased number of children who underwent surgery classified into cluster A, whilst most of the children in cluster B did not undergo surgery, confirming the beneficial effect of AT. [10, 24–26].

By analysing the cluster transitions for each child separately for both study arms (Table 5), we found that over half of the children who underwent surgery were assigned to cluster A at baseline and follow-up, whilst only 1/3 of children in the watchful waiting group were assigned to, and remained in cluster A. Considering the higher percentage for children who underwent surgery, surgery appears to improve the likelihood of children staying in cluster A. The presence of children of the watchful waiting arm in cluster A suggests that their condition was mild and

did not worsen during watchful waiting. By contrast, our results suggest it is more likely for children in the watchful waiting arm, who were assigned to cluster B, to remain in B even after seven months. Importantly, 16 children who had surgery and remained in cluster B; our posthoc analysis showed that those children had extremely high peak end-tidal carbon dioxide measure during sleep at baseline PSG. This suggests the presence of sleep-related hypoventilation even in the absence of discrete respiratory events and implies these children still had significant UAO despite AT.

It was almost six times more likely for children in the surgery group who were in cluster B at baseline to transition to cluster A at follow-up than for children in the WW arm. Children in the WW arm were > 3 times more likely to transition from cluster A to B than those who underwent surgery. This further demonstrates that surgery can help reduce OSAS related physiological symptoms. Moreover, those symptoms are more likely getting worse if children remain without treatment. Interestingly, five children in the WW arm transitioned from cluster B to cluster A after seven months. These children had higher pathophysiological symptoms at baseline PSG, and significantly more problematic behaviour and emotional issues than others, which is in contrast with the literature [16, 38].

Our study has several limitations. Children enrolled in CHAT were within the mild-to-moderate spectrum of OSAS, and the follow-up duration was relatively short. Children were free of significant comorbidities, which may limit the generalisability of our findings. Distinct clusters may form more clearly if more severe cases of OSAS are considered; possibly more than two clusters could be considered. The size of the data set prevented us from validating our classifier with unseen children. This may have introduced a bias. Since the trial did not include an arm of healthy children, we cannot verify if children in cluster A had characteristics similar to healthy children. We excluded several children from the original trial due to poor signal quality. Some of the transition subclasses had very few children, which may have impacted the reliability of some results.

In conclusion, data-driven analysis shows that AT has a beneficial effect on children with OSAS by reversing or preventing the worsening of the syndrome. Multidomain analysis of PSG markers and machine learning may yield a more comprehensive picture of OSAS than AHI alone. Our findings indicate that children could avoid AT if their physiological and neurophysiological symptoms are mild. UAO symptoms persist post AT in children who may suffer sleep-related hypoventilation at baseline. These findings may help predict children with UAO would benefit from AT.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11325-021-02425-w>.

**Acknowledgements** We would like to thank Michael Rueschman, Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Boston, MA, USA for support with handling and interpreting the CHAT dataset. Xiao Liu, Sarah Immanuel, and Mathias Baumert had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Yvonne Pamula, James Martin, and Declan Kennedy contributed substantially to the interpretation and the writing of the manuscript.

**Funding** The Childhood Adenotonsillectomy Trial (CHAT) was supported by the National Institutes of Health (HL083075, HL083129, UL1-RR-024134, UL1 RR024989). The National Sleep Research Resource was supported by the National Heart, Lung, and Blood Institute (R24 HL114473, 75N92019R002).

## Declarations

**Ethical approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee of each participating institution and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

**Conflict of interest** All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent/licencing arrangements), or nonfinancial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

## References

1. Baumert M, Pamula Y, Martin J et al (2016) The effect of adenotonsillectomy for childhood sleep apnoea on cardiorespiratory control. *ERJ Open Research* 2:00003–02016
2. Baumert M, Walther T, Hopfe J, Stepan H, Faber R, Voss A (2002) Joint symbolic dynamic analysis of beat-to-beat interactions of heart rate and systolic blood pressure in normal pregnancy. *Med Biol Eng Compu* 40:241–245
3. Berry RB, Budhiraja R, Gottlieb DJ et al (2012) Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med* 8:597–619
4. Bhattacharjee R, Kheirandish-Gozal L, Spruyt K et al (2010) Adenotonsillectomy outcomes in treatment of obstructive sleep apnea in children: a multicenter retrospective study. *Am J Respir Crit Care Med* 182:676–683
5. Bozkurt S, Bostanci A, Turhan M (2017) Can statistical machine learning algorithms help for classification of obstructive sleep apnea severity to optimal utilization of polysomnography resources? *Methods Inf Med* 56:308–318
6. Brietzke SE, Gallagher D (2006) The effectiveness of tonsillectomy and adenoidectomy in the treatment of pediatric obstructive sleep apnea/hypopnea syndrome: a meta-analysis. *Otolaryngology-Head and Neck Surgery* 134:979–984

7. Chervin RD, Ellenberg SS, Hou X et al (2015) Prognosis for spontaneous resolution of OSA in children. *Chest* 148:1204–1213
8. Costa DJ, Mitchell R (2009) Adenotonsillectomy for obstructive sleep apnea in obese children: a meta-analysis. *Otolaryngology-Head and Neck Surgery* 140:455–460
9. Di-Tullio F, Ernst G, Robaina G et al (2018) Ambulatory positional obstructive sleep apnea syndrome. *Sleep Science* 11:8
10. El-Hamad F, Immanuel S, Liu X et al (2017) Altered Nocturnal Cardiovascular Control in Children With Sleep-Disordered Breathing. *Sleep* 40:zsx127
11. Guilleminault C, Huang Y-S, Chin W-C, Okorie C (2018) The nocturnal-polysomnogram and “non-hypoxic sleep-disordered-breathing” in children. *Sleep medicine*
12. Guilleminault C, Lee J, Chan A (2005) Pediatric obstructive sleep apnea syndrome. *Arch Pediatr Adolesc Med* 159:775–785
13. Guilleminault C, Li K, Khramtsov A, Palombini L, Pelayo R (2004) Breathing patterns in prepubertal children with sleep-related breathing disorders. *Arch Pediatr Adolesc Med* 158:153–161
14. Immanuel S, Kohler M, Martin J et al (2014) Increased thoraco-abdominal asynchrony during breathing periods free of discretely scored obstructive events in children with upper airway obstruction. *Sleep Breath* 19:65–71
15. Immanuel SA, Pamula Y, Kohler M et al (2012) Respiratory timing and variability during sleep in children with sleep-disordered breathing. *J Appl Physiol* 113:1635–1642
16. Jackman AR, Biggs SN, Walter LM et al (2012) Sleep-disordered breathing in preschool children is associated with behavioral, but not cognitive, impairments. *Sleep Med* 13:621–631
17. Jain AK (2010) Data clustering: 50 years beyond K-means. *Pattern Recogn Lett* 31:651–666
18. Katz ES, D’ambrosio CM (2010) Pediatric Obstructive Sleep Apnea Syndrome. *Clinics in Chest Medicine* 31:221–34
19. Kim T, Kim J-W, Lee K (2018) Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques. *Biomed Eng Online* 17:16
20. Kontos A, Lushington K, Martin J et al (2017) Relationship between vascular resistance and sympathetic nerve fiber density in arterial vessels in children with sleep disordered breathing. *Journal of the American Heart Association*, 6: e006137
21. Kontos A, Van Den Heuvel C, Pamula Y et al (2015) Delayed brachial artery dilation response and increased resting blood flow velocity in young children with mild sleep-disordered breathing. *Sleep Med* 16:1451–1456
22. Kontos A, Willoughby S, Van Den Heuvel C et al (2018) Ascending aortic blood flow velocity is increased in children with primary snoring/mild sleep-disordered breathing and associated with an increase in CD8+ T cells expressing TNF $\alpha$  and IFN $\gamma$ . *Heart Vessels* 33:537–548
23. Laing EE, Möller-Levet CS, Dijk D-J, Archer SN (2018) Identifying and validating blood mRNA biomarkers for acute and chronic insufficient sleep in humans: a machine learning approach. *Sleep*, 42
24. Liu X, Immanuel S, Kennedy D, Martin J, Pamula Y, Baumert M (2018a) Effect of adenotonsillectomy for childhood obstructive sleep apnea on nocturnal heart rate patterns. *Sleep*, 41: zsy171-zsy71
25. Liu X, Immanuel S, Pamula Y, Kennedy D, Martin J, Baumert M (2017) Adenotonsillectomy for childhood obstructive sleep apnoea reduces thoraco-abdominal asynchrony but spontaneous apnoea–hypopnoea index normalisation does not. *European Respiratory Journal*, 49
26. Liu X, Immanuel S, Pamula Y et al (2018b) Pulse wave amplitude and heart period variability in children with upper airway obstruction. *Sleep Medicine*
27. Liu X, Pamula Y, Kohler M, Baumert M (2019) A Method for Estimating Pulse Wave Amplitude Variability in children with Sleep Disordered Breathing. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC): 2289–92
28. Lumeng JC, Chervin RD (2008) Epidemiology of pediatric obstructive sleep apnea. *Proc Am Thorac Soc* 5:242–252
29. Marcus CL, Brooks LJ, Ward SD et al (2012) Diagnosis and management of childhood obstructive sleep apnea syndrome. *Pediatrics* 130:e714–e755
30. Marcus CL, Moore RH, Rosen CL et al (2013) A Randomized Trial of Adenotonsillectomy for Childhood Sleep Apnea. *N Engl J Med* 368:2366–2376
31. Ng D, Chan C, Chow A, Chow P, Kwok K (2005) Childhood sleep-disordered breathing and its implications for cardiac and vascular diseases. *J Paediatr Child Health* 41:640–646
32. Nisbet LC, Yiallourou SR, Walter LM, Horne RS (2014) Blood pressure regulation, autonomic control and sleep disordered breathing in children. *Sleep Med Rev* 18:179–189
33. Rapoport DM (2016) POINT: Is the apnea-hypopnea index the best way to quantify the severity of sleep-disordered breathing? Yes. *Chest* 149:14–16
34. Redline S, Amin R, Beebe D et al. "Childhood Adenotonsillectomy Study." National Sleep Research Resource. Web. <http://sleepdata.org/datasets/chat>.
35. Redline S, Amin R, Beebe D et al (2011) The Childhood Adenotonsillectomy Trial (CHAT): Rationale, Design, and Challenges of a Randomized Controlled Trial Evaluating a Standard Surgical Procedure in a Pediatric Population. *Sleep* 34:1509–1517
36. Spruyt K, Verleye G, Gozal D (2010) Unbiased Categorical Classification of Pediatric Sleep Disordered Breathing. *Sleep* 33:1341–1347
37. Suen JS, Arnold JE, Brooks LJ (1995) Adenotonsillectomy for treatment of obstructive sleep apnea in children. *Archives of Otolaryngology-Head & Neck Surgery* 121:525–530
38. Tamanyan K, Walter LM, Weichard A et al (2018) Age effects on cerebral oxygenation and behavior in children with sleep-disordered breathing. *Am J Respir Crit Care Med* 197:1468–1477
39. Zhang Z, Mayer G, Dauvilliers Y et al (2018) Exploring the clinical features of narcolepsy type 1 versus narcolepsy type 2 from European Narcolepsy Network database with machine learning. *Sci Rep* 8:10628

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Xiao Liu<sup>1</sup> · Yvonne Pamula<sup>2</sup> · Sarah Immanuel<sup>3,4</sup> · Declan Kennedy<sup>2,5</sup> · James Martin<sup>2</sup> · Mathias Baumert<sup>1</sup> 

<sup>1</sup> School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia

<sup>2</sup> Department of Respiratory and Sleep Medicine, Women's and Children's Hospital, Adelaide, Australia

<sup>3</sup> Centre for Artificial Intelligence Research and Optimisation, Torrens University, Adelaide, Australia

<sup>4</sup> College of Medicine and Public Health, Flinders University, Adelaide, Australia

<sup>5</sup> Children's Research Centre, School of Paediatrics and Reproductive Health, The University of Adelaide, Adelaide, Australia