

Validation of breath biomarkers for obstructive sleep apnea

Nora Nowak^a, Anna Engler^b, Sira Thiel^b, Anna S. Stöberl^b, Pablo Sinues^{c,d}, Renato Zenobi^a, Malcolm Kohler^{b,§}

^a ETH Zurich, Department of Chemistry and Applied Biosciences, Vladimir-Prelog-Weg 3, 8093 Zürich, Switzerland

^b University Hospital Zurich, Department of Pulmonology, Rämistrasse 100, 8091 Zürich, Switzerland

^c University Children's Hospital Basel, Spitalstrasse 33, 4056 Basel, Switzerland

^d University of Basel, Department of Biomedical Engineering, Gewerbestrasse 14, 4123 Allschwil, Switzerland

[§]Corresponding author, **e-mail:** Malcolm.Kohler@usz.ch, **postal address:** Malcolm Kohler, University Hospital Zurich, Department of Pulmonology, Rämistrasse 100, 8091 Zürich,

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Abstract

Background and objectives: Obstructive sleep apnea (OSA) is an underdiagnosed respiratory disease with negative metabolic and cardiovascular effects. The current gold standard for diagnosing OSA is in-hospital polysomnography, a time-consuming and costly procedure, often inconvenient for the patient. Recent studies revealed evidence for the potential of breath analysis for the diagnosis of OSA based on a disease-specific metabolic pattern. However, none of these findings were validated in a larger and broader cohort, an essential step for its application in clinics.

Methods: In the present study, we validated a panel of breath biomarkers in a cohort of patients with possible OSA (N = 149). These markers were previously identified in our group by secondary electrospray ionization high-resolution mass spectrometry (SESI-HRMS).

Results: Here, we could confirm significant differences between metabolic patterns in exhaled breath from OSA patients compared to control subjects without OSA as well as the association of breath biomarker levels with disease severity. Our prediction of the diagnosis for the patients from this completely independent validation study using a classification model trained on the data from the previous study resulted in an area under the receiver operating characteristic curve of 0.66, which is comparable to questionnaire-based OSA screenings.

Conclusions: Thus, our results suggest that breath analysis by SESI-HRMS could be used to screen for OSA. Its true predictive power should be tested in combination with OSA screening questionnaires.

Clinical trial: “Mass Spectral Fingerprinting in Obstructive Sleep Apnoea”, NCT02810158, www.ClinicalTrials.gov

Keywords: obstructive sleep apnea, biomarkers, validation study, breath analysis, metabolomics, secondary electrospray ionization mass spectrometry

1. Introduction

Obstructive sleep apnea (OSA) is a highly prevalent sleep-related breathing disorder.[1] The repeated partial or complete collapse of the pharynx during sleep provokes apnea or hypopnea events, which may lead to repetitive oxygen desaturations. Frequent sleep disruptions and increased activity of the sympathetic nervous system are accompanying these apnea/hypopnea events resulting in poor sleep quality and increased daytime sleepiness.[2] Several metabolic and cardiovascular consequences, such as an increased risk for cardiovascular disease, arterial hypertension, diabetes, vascular dysfunction, as well as depression, are well known.[3–6] OSA can be effectively treated i.e. with continuous positive airway pressure (CPAP).[7–9]

The conventional diagnosis of OSA is carried out by respiratory polygraphy or even polysomnography ,[10,11] which are time-consuming, costly and inconvenient for patients. In addition, there is emerging evidence for a high night-to-night variability of OSA, posing another challenge for diagnostics.[12] Thus, for a reliable diagnosis, testing during several nights would be required. Screening for OSA is conventionally based on questionnaires, such as the Epworth Sleepiness Scale,[13] the STOP-bang,[14] Berlin[15] or NoSAS score.[16] However, the results from such questionnaires are by nature subjective.

Exhaled breath contains several hundreds of metabolites and thus provides insights into biochemical processes of the human body.[17] Many of the metabolites in breath do not originate from the lungs but are transported from blood to the airways via gas exchange in the lung. Therefore, breath metabolite levels mostly reflect systemic metabolic processes. Furthermore, consistent alterations of the molecular fingerprint of exhaled breath in patients with a certain disease may indicate disease specific metabolic changes. Such disease specific biomarkers detected in exhaled breath could be the basis for an objective and non-invasive diagnostic procedure, which is fast and easy to perform for patients.

So far, many studies with small sample sizes have obtained promising results, suggesting a great diagnostic potential of exhaled breath analysis for various diseases. However, larger validation studies are missing and, to date, exhaled breath analysis is applied in clinical routine only for very few applications, such as the evaluation of bronchial inflammation by measuring fractional exhaled nitric oxide (FeNO).[18] To achieve a more widespread clinical application of breath analysis for disease diagnosis and monitoring, the validation of preliminary findings in large cohorts of patients is essential.

The investigation of exhaled breath in OSA patients using different technical approaches revealed convincing results for diagnosing this disease and for monitoring patients' compliance to CPAP therapy.[19] In some studies, electronic sensors (e-noses) were used to recognize OSA specific patterns in exhaled breath.[20–23] In their attempts of diagnosing OSA against the gold-standard (polysomnography), areas under the receiver operating characteristic curves (AUROCs) in the range from 0.84 (no 95% CI provided) to 0.87 (95% CI 0.61–1.00) were reported, suggesting future diagnostic applicability of breath analysis. However, e-noses do not allow for compound identification and thus do not provide mechanistic insights into the disease, but merely produce a complex “signal” whose statistical evaluation can give some valuable output. Furthermore, the data from studies with e-noses were not validated in larger and broader cohorts of patients with possible OSA.

An untargeted investigation from our laboratories of an extensive spectrum of molecules in exhaled breath using secondary electrospray ionization high-resolution mass spectrometry (SESI-HRMS) revealed specific markers, allowing identification of a disease specific molecular profile of exhaled breath in patients with OSA recurrence after two weeks of CPAP therapy withdrawal.[24] In that randomized controlled trial, significant correlations between metabolite levels in breath and change in oxygen desaturation index (ODI) upon CPAP withdrawal, and significant differences between the CPAP withdrawal and treatment group were found. Further, we achieved a successful classification (AUROC = 0.87) between the withdrawal group and the group that continued the treatment. In order to transfer our promising findings into the diagnostic algorithm of OSA, this observational study aims to validate these metabolic breath profiles in a larger treatment naive cohort of patients with possible OSA.

2. Methods

2.1 Study participants

This study includes 149 participants with possible OSA in the age of 53.3 ± 13.7 years with a BMI of 30.1 ± 6.6 kg/m² (table 1). The study protocol was approved by the local ethical committee (KEK-ZH 2016-00384). The experiments were conducted in accordance with the Declaration of Helsinki and written informed consent was obtained from all participants before participation. The clinical trial was registered at ClinicalTrials.gov (NCT02810158).

All patients underwent in-hospital respiratory polygraphy (RP). Inpatient RPs were recorded by Alice 6 Diagnostic System (Philips Respironics, PA, USA), scored with validated Somnolzyer 24x7 software (Philips Respironics, PA, USA), and reviewed manually. The data obtained was evaluated according to the guidelines of the American Academy of Sleep Medicine.[13] The participants were asked to fill in Epworth Sleepiness Scale (ESS) questionnaires.

2.2 SESI-HRMS measurements

Participants were asked to refrain from eating, drinking, chewing gum, alcohol, tobacco, caffeine use or brushing their teeth at least 1 hour prior to the SESI-HRMS measurements. Exhaled breath of 149 patients was analyzed by SESI-HRMS using a commercial SESI source (SEADM, Spain) coupled to a high-resolution TripleTOF 5600+ mass spectrometer (AB Sciex, Concord, Ontario, Canada). The participants were sitting in upright position in front of the mass spectrometer and exhaled at least six times with a pressure drop of 12 mbar through a disposable mouthpiece into the heated sampling line, connected to the SESI source. A flow splitter at the front-end enabled sampling of end-tidal breath. The flow through the ion source was set to 0.2 L/min. Full scan mass spectra were recorded in positive ion mode with an accumulation time of 1 s in the range of 50-500 Da.

2.3 Data preprocessing

All mass spectral data was analyzed with MATLAB R2020a and R 4.0.0. Mass spectra obtained from exhaled breath were preprocessed as described elsewhere.[25] In short, mass spectra were interpolated, aligned, exhalation time windows were chosen and peak picking was performed on the average breath spectrum. As in the pilot study[24], breath signal intensities were normalized to the median intensity of the total ion current and then autoscaled.

2.4 Statistical analysis

Further, the features were filtered for markers that have been associated with OSA previously[24]. The mass-to-charge ratio (m/z) tolerance was set to 0.005 Da. The remaining 78 m/z features were first tested for normality in a Shapiro-Wilk's test. Since the data was not normally distributed (p-value distribution from

Shapiro-Wilk's test for normality is provided in Figure A1), we performed a correlation analysis between signal intensity and ODI as well as between signal intensity and ESS using Spearman correlation.

Moreover, we tested for differences in signal intensities between individuals without OSA and OSA patients by performing two-sided Mann-Whitney-U tests. Here, we first applied stratification criteria as they are commonly applied in the clinics: OSA ODI > 30/h or ODI > 10/h & ESS > 10 points, control ODI < 5/h or ODI < 10/h & ESS < 11 points (stratification 1). We then also tested for between-group differences with stricter stratification criteria (OSA: ODI > 30/h & ESS > 10 points, control: ODI < 10/h & ESS < 11 points, stratification 2) in order to remove individuals with ambiguous diagnosis. We also calculated $\log_2$ fold changes between the groups. To account for multiple hypothesis testing, false discovery rates (q-values) were calculated for all obtained p-values using Storey's procedure.[26]

2.5 Classification Procedure

We combined the breath intensities obtained for the 78 m/z features mentioned above of the previously reported pilot study and this validation study. We used the data of the pilot study as training set and used the MATLAB classification learner app to find the best classification algorithm. For model evaluation we used a 7-fold cross validation. We defined the OSA and control group with the above mentioned criteria of stratification 1. In order to obtain balanced group sizes we only used the before and after measurements of the 9 individuals of the placebo group, who developed OSA in the previous study.[24] A Gaussian support vector machine model performed best. We thus trained such model on the training data and predicted the validation data set obtained from this study.

2.6 Attempts of improvement of classification performance

Since both data sets were acquired on different mass spectrometers and with different generations of SESI sources, we assessed by principal component analysis (PCA) the comparability of both data sets. A slight shift between both data sets was observed. We therefore performed a batch correction based on an empirical Bayes algorithm[27] and repeated the classification procedure described above (Figure A2).

We also repeated the classification procedure with stratification criteria, which are more similar to the ones used in the pilot study. For the validation data set we defined the groups as follows: OSA: ODI > 30/h, control: ODI < 10/h. In order to get balanced group sizes in the training set, we reduced the control group to ODI < 2/h.

3. Results

3.1 Study design and patient characteristics

149 participants between 19 and 83 years with possible OSA underwent respiratory polygraphy and breath analysis by SESI-HRMS directly after the sleep study (Figure 1). Patient characteristics and results from respiratory polygraphy are shown in Table 1. Depending on the applied stratification criteria, the mean ODI in the OSA group varies between 38.3 and 46.8 events per hour. The control group had a mean ODI between 4.4 and 5.2 events per hour. The mean Epworth Sleepiness Scale score (obtained from a questionnaire estimating the extend of daytime sleepiness) in the OSA group ranged from 10 to 11 points and in the control group from 6 to 9.4 points.

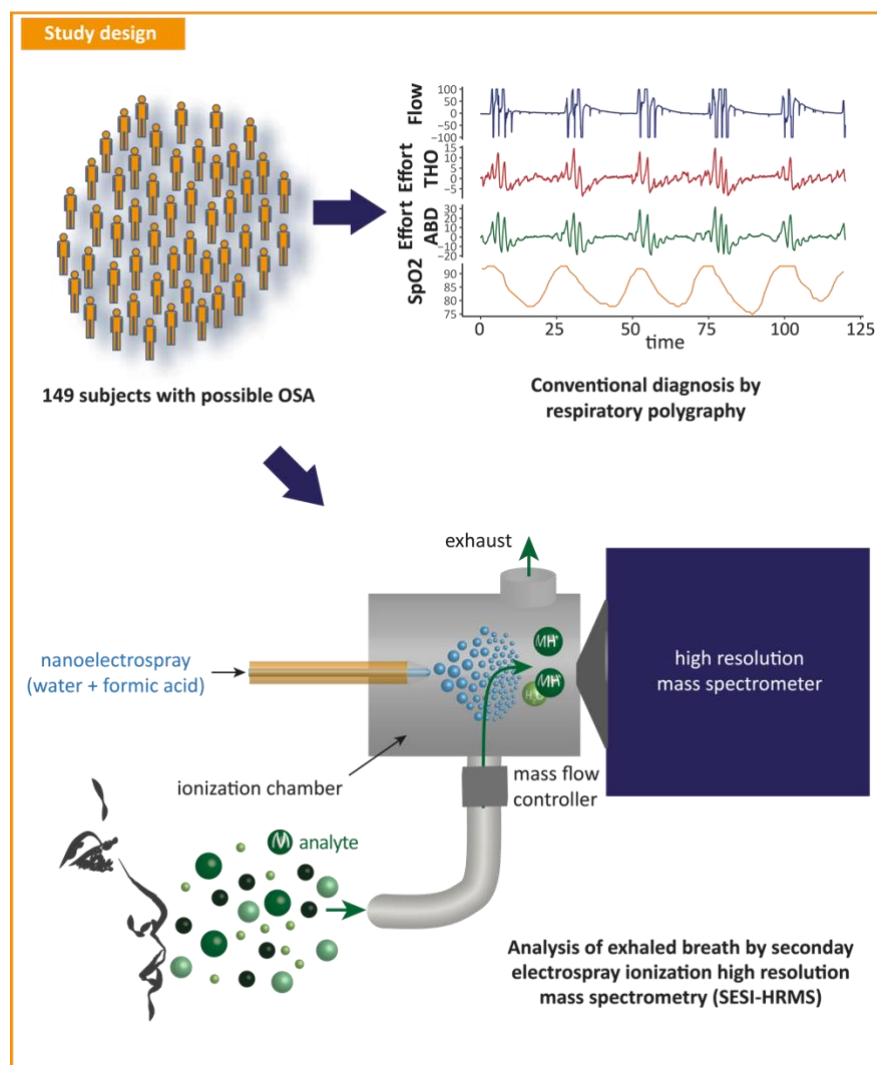


Figure 1: Study design. 149 study subjects with possible OSA underwent conventional diagnosis by respiratory polygraphy in the sleep laboratory and exhaled breath analysis by SESI-HRMS.

3.2 Metabolic patterns in exhaled breath associated with OSA

The data obtained in this validation study from SESI-HRMS measurements was pre-processed in the same way as it was done in our previous study[24], which we refer to as pilot study in the following, i.e., the signal intensities were normalized to the median of total ion current and then autoscaled. We continued our analysis in a targeted fashion focusing only on the m/z features that have previously been associated with OSA in our pilot study. However, we were not able to detect all of them, which is most likely due to technical optimization of the acquisition method that was made in the meantime. Nevertheless, 78 of the features that have been previously reported either as significantly different between the CPAP and the withdrawal group, or as correlating with the change in ODI or as predictive for OSA, were also detected in this validation study. For those 78 m/z features, we tested for significant differences between controls (without OSA) and OSA patients and for correlation with ODI and ESS. Moreover, we trained a classification model with the data from the pilot study and predicted the OSA diagnosis of the validation cohort from this study.

3.3 Significant differences in metabolic breath patterns between OSA patients and individuals without OSA

We tested for significant differences in metabolite intensities in exhaled breath between OSA patients and controls without OSA (Mann-Whitney-U test). We assigned the participants to two groups (OSA and control) based on the following criteria, commonly applied in clinics: ODI > 30/h or ODI > 10/h & ESS > 10 points (definitive OSA); and ODI < 5/h or ODI < 10/h & ESS < 11 points (definitive without relevant OSA; control) (stratification 1). All subjects in between were assigned to an “unclear” group, since no unambiguous OSA diagnosis could be stated. For 19 features we found significant ($p \leq 0.05$) differences between the two groups (Figure 2A-C, boxplots of two examples are shown in Figure 2D and E, all boxplots are provided in Figure A3). When we used stricter grouping criteria (OSA: ODI > 30/h & ESS > 10 points, control: ODI < 10/h & ESS < 11 points, stratification 2) in order to consider only patients with an unambiguous diagnosis, significance increases as shown in Figure A4 (all boxplots are given in Figure A5). All numeric results are provided in Table A1. Hence, our results from this validation study confirm our previous findings of a specific metabolic pattern in exhaled breath in OSA patients.

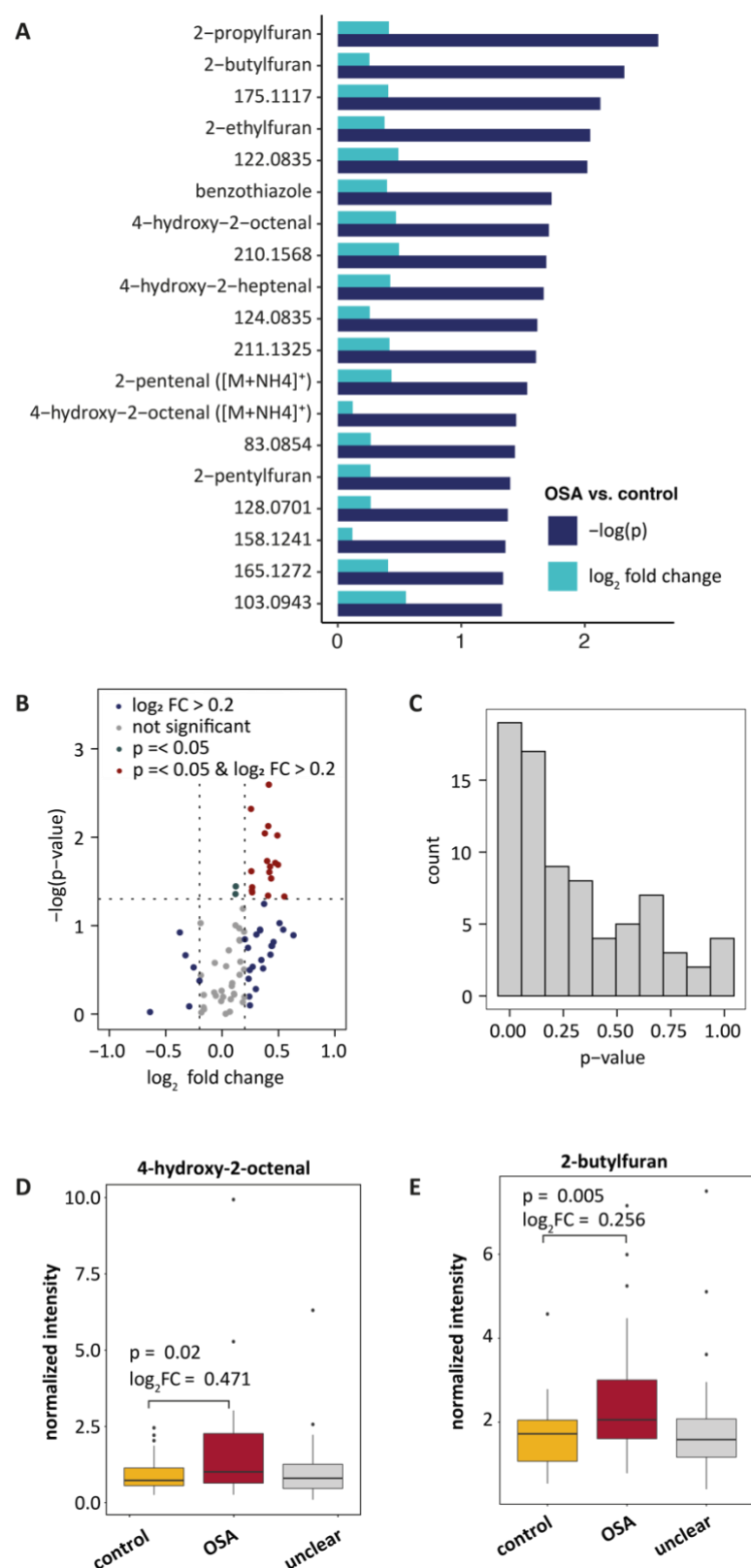


Figure 2: Significant differences in metabolic breath patterns between subjects with and without OSA (OSA: ODI > 30/h or ODI > 10/h & ESS > 10 points; control: ODI < 5/h or ODI < 10/h & ESS < 11 points; unclear: in between; stratification 1). **A** p-values and fold changes of significant features sorted by significance. Not identified features are labelled with their m/z. **B** volcano plot for all 78 metabolites. **C** p-value distribution for between-group differences from Mann-Whitney-U test. **D**, **E** Exemplary boxplots (center line: median, box limits: 25th and 75th percent quantile, whisker length: 1.5 interquartile range) of 4-hydroxy-2-octenal and 2-butylfuran. The Boxplots for all significant features are provided in Figure A1, numeric results for significant features are summarized in Table 2 and numeric results of all 78 features are given in Table A1.

3.4 Association between disease severity and breath signal intensity

To test whether the breath patterns are also correlating with the severity of OSA in this larger and more diverse cohort of patients, we performed a Spearman correlation analysis. For 21 features, we found a significant correlation between breath levels and ODI ($p \leq 0.05$). (Figure 3A, correlation plots for two examples are shown in Figure 3B, all correlation plots are provided in Figure A6) All except one show higher intensities for an increased ODI, suggesting that oxygen desaturation correlates with an enrichment of these metabolites. Amongst these metabolites correlating with the ODI are several unsaturated aldehydes as well as furanes and benzothiazole that have been identified before.[28–30] Thus, we could confirm the previously reported association between disease severity and breath signal intensity.

3.5 Association between sleepiness and breath signal intensity

For nine features we found a significant correlation between their breath intensities and the ESS ($p \leq 0.05$) (Figure 4A). Amongst them, four features furthermore correlate with ODI, such as 2-pentylfuran and 4-hydroxy-2-octenal. Correlation plots for these two examples are shown in Figure 4B (all correlation plots are provided in Figure A7).

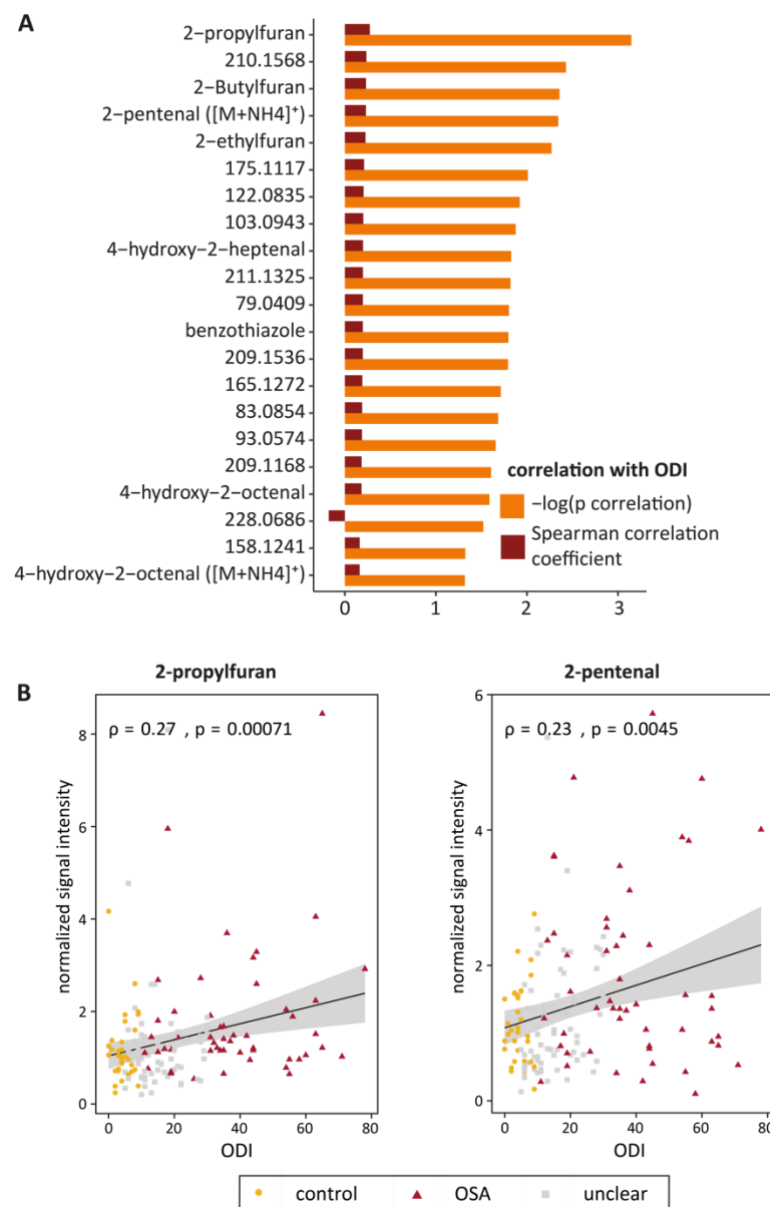


Figure 3: Correlations between metabolite levels in breath and OSA severity. **A** p-values and correlation coefficients of features from exhaled breath with significant correlations with the ODI. Not identified features are labelled with their m/z. **B** Exemplary regression lines for 2-propylfuran and 2-pentenal. Regression lines for all features with significant correlations with the ODI are provided in Figure A4. Numeric results for features with significant correlations are summarized in Table 2 and numeric results of all 78 features are given in Table A1. (OSA: ODI > 30/h or ODI > 10/h & ESS > 10 points; control: ODI < 5/h or ODI < 10/h & ESS < 11 points; unclear: in between; stratification 1)

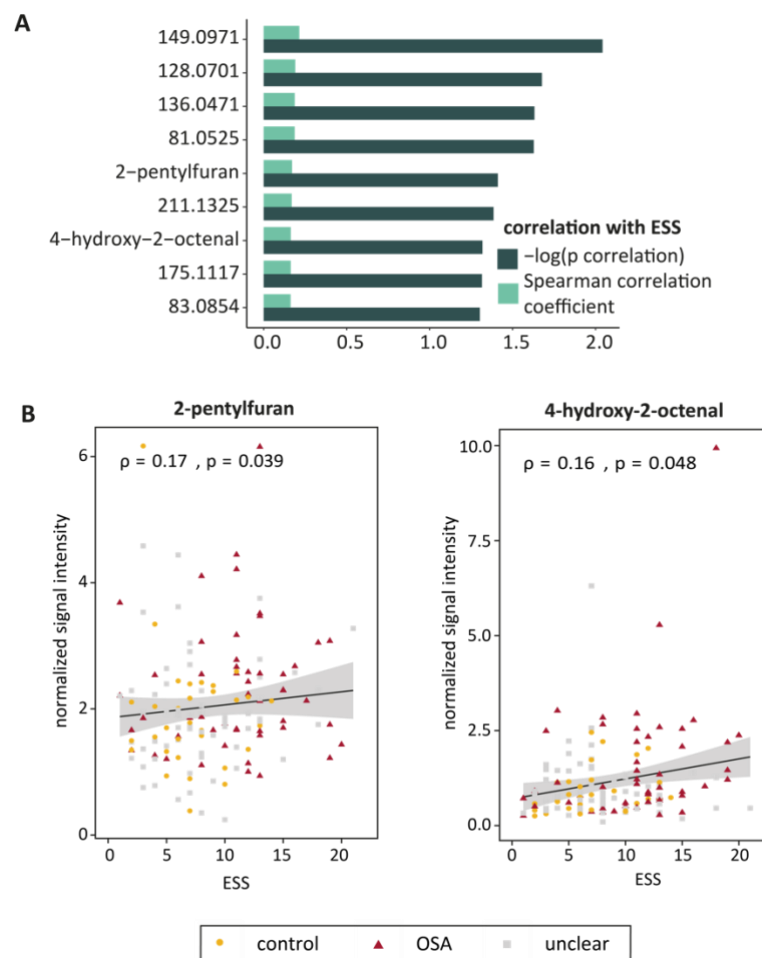


Figure 4: Correlations between metabolite levels in breath and sleepiness. **A** p-values and correlation coefficients of features from exhaled breath with significant correlations with the ESS. Not identified features are labelled with their m/z. **B** Exemplary regression lines for 2-pentylfuran and 4-hydroxy-2-octenal. Regression lines for all features with significant correlations with the ESS are provided in Figure A5. Numeric results for features with significant correlations are summarized in Table 2 and numeric results of all 78 features are given in Table A1. (OSA: ODI > 30/h or ODI > 10/h & ESS > 10 points; control: ODI < 5/h or ODI < 10/h & ESS < 11 points; unclear: in between; stratification 1)

3.6 Classification

To assess the applicability of metabolite levels measured in exhaled breath using SESI-HRMS for the clinical diagnosis of OSA, we trained a classification model with the data from our pilot study and predicted the diagnosis of OSA or control in the validation cohort measured in this study. Figure 5 shows the classification procedure schematically.

First, we grouped the patients again as described above by the clinical criteria of stratification 1. In order to obtain balanced group sizes in the training set, we used only the “before” and “after” measurements of those patients in the CPAP withdrawal group, who developed significant OSA under placebo treatment (Figure 6A). With the training data, we estimated the performance of the classification model in a 7-fold cross-validation. This resulted in an AUROC of 0.59 (Figure 6B, the confusion matrix is provided in Figure 6C). The prediction of the diagnosis for the validation data (Figure 6D) yielded in an AUROC of 0.66 (Figure 6B, confusion matrix is given in Figure 6C). The accuracy of the prediction was 63% with a sensitivity of 76% and a specificity of 42%.

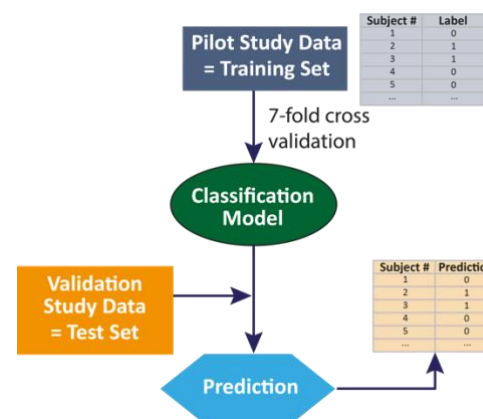


Figure 5: Classification procedure. A classification model was trained with the data from our pilot study and its performance was estimated in a 7-fold cross validation. Subsequently, the diagnosis of the patient cohort from this validation study was predicted.

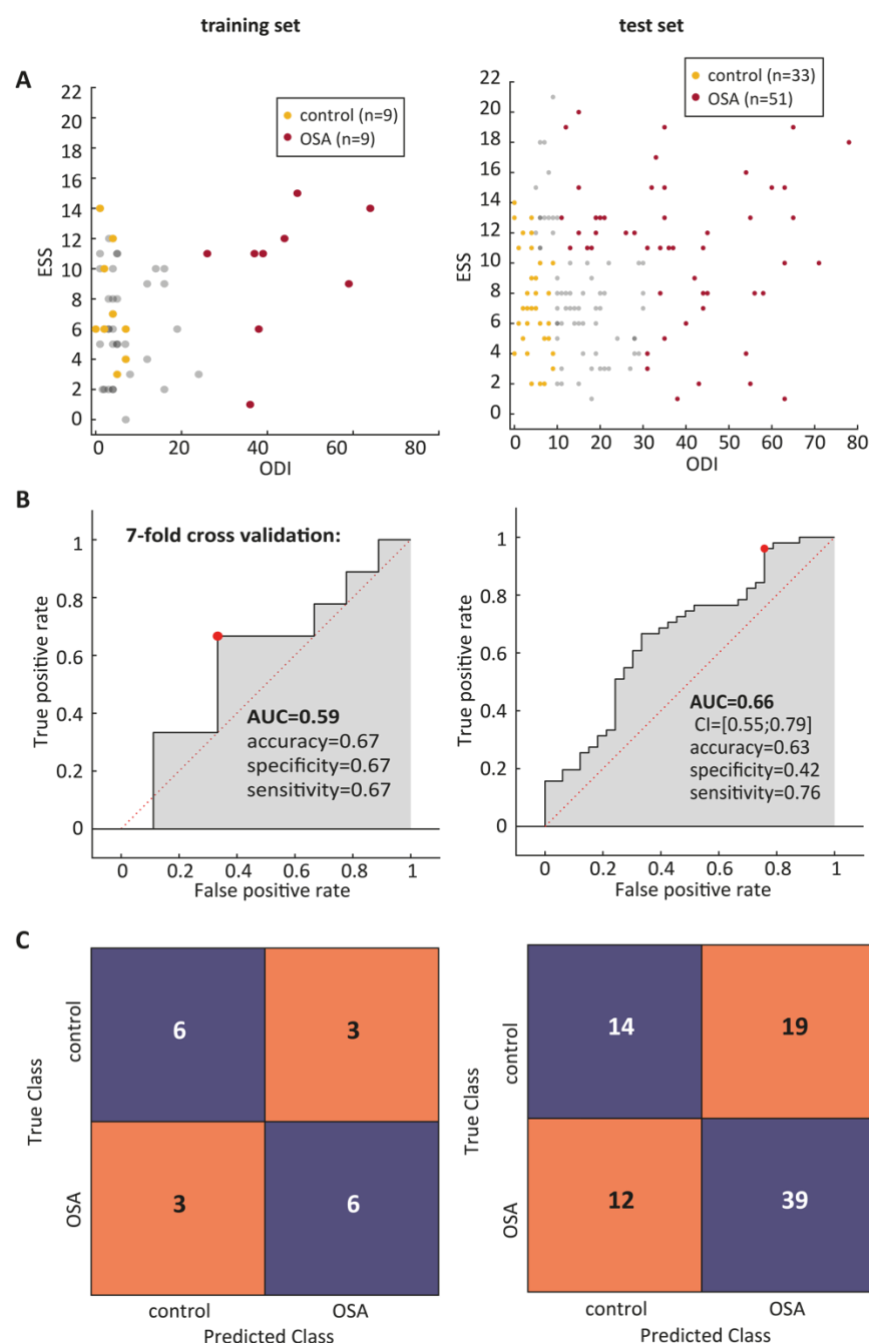


Figure 6: Classification results. **A** ESS and ODI of samples in the training set (left) and test set (right). **B** ROC curve from 7-fold cross validation of the classification model with the training set (left) and from predictions of the validation cohort. **C** corresponding confusion matrices. (OSA: ODI > 30/h or ODI > 10/h & ESS > 10 points; control: ODI < 5/h or ODI < 10/h & ESS < 11 points; stratification 1)

4. Discussion

To the best of our knowledge, this is the first report of a validation of breath biomarkers for OSA. Previous studies using e-noses, offline gas-chromatography coupled to mass spectrometry, or enzyme immunoassays to analyze exhaled breath condensate have achieved promising results regarding the distinction between OSA patients and controls without OSA from exhaled breath.[19] However, sample sizes in all these studies were

limited and none of the results has been validated in an independent cohort of patients. In this study, we could confirm in a large and independent cohort that breath intensities of many of our previously discovered potential biomarkers for OSA differ significantly between OSA patients and controls without OSA. Most of the markers are consistently increased in OSA patients. We could also confirm a correlation between breath signal intensity and disease severity (represented by the ODI) for several metabolites, supporting the association of these metabolites with apnea-related nocturnal hypoxemia. Our findings of correlations between breath signals and ESS scores indicate that not only hypoxia but also sleepiness is reflected in the metabolic breath pattern. The results from this study suggest that the 33 metabolites shown in Table 2 represent a panel of biomarkers, being robust enough concerning inter-individual variability to form a promising diagnostic tool. Inter-individual differences are the most likely reason for lower correlation coefficients between signal intensities in breath and ODI that we observed in this diverse validation cohort compared to the correlations between the signal intensities and the within-subject change of ODI upon CPAP withdrawal, which we reported previously.

It seems unlikely that there is one single biomarker, which is sufficient for diagnosing a disease like OSA, associated with complex metabolic and cardiovascular consequences. In contrast, a pattern of several biomarkers is more likely to be disease-specific. Therefore, classification algorithms based on machine learning are convenient tools for making clinical diagnoses based on biomarker patterns. Here, we achieved a classification of the validation data set with an AUROC of 0.66, 76% sensitivity and 42% specificity, when we trained the model with the data from the independent patient cohort of our previously reported study.

Since the patient cohort of this study was much more diverse compared to the one in our pilot study, a lower classification performance would be expected. However, the support vector machine (SVM) model we used, performed already worse in the cross validation with the training data ($AUC_{CV} = 0.59$) compared to the model presented in the pilot study. This is likely to be due to different stratification criteria used in both studies, since with such small training sets, few samples can have a considerable influence. To test this hypothesis, we applied different stratification criteria. When we stratified the training data only based on ODI (as it was done in the pilot study), the model performance becomes comparable with the results that were reported previously ($AUC_{CV} = 0.79$, Figure A8A-C). This result supports our choice of an SVM model. However, the prediction of the validation data, which was also stratified only based on the ODI, not only failed to improve, but even declined

slightly (AUC = 0.62, Figure A8A-C). This indicates that groups that are defined based on ODI and ESS can be distinguished better from the breath pattern with this biomarker panel than groups that are defined by ODI only. This is desirable, since a combination of ODI and ESS better reflects the clinical picture of significant OSA than ODI on its own. All performance measures of this classification procedure (classification 2) are reported in the supplementary results and are summarized in Table A2.

SESI is still a rather novel technique and its performance is constantly improving. Furthermore, in the time between the pilot study and this validation study, technical improvements have been implemented. For example, a new, more robust ion source[31], has been developed, and we adapted collision gas settings of the MS in order to prevent fragmentation, although this leads to increased cluster formation. This is most likely the reason why we did no longer detect all of the markers that were reported previously. The comparison of our data before and after batch correction between the data sets from both studies indicated that the data for the 78 potential biomarkers detected in both data sets are comparable and there is only a very small batch effect. A negligible change of the results from the classification procedure applied after batch correction confirmed this observation (classification 3, PCA plot and classification outcome are shown in Figure A2 and Figure A9, and classification results are described in supplementary results and Table A2).

Another factor that might impair the classification results is a lack of standardization of SESI-HRMS and the lack of real-time breath quality control samples, respectively, at the time when the study was conducted. Instrumental drifts are a common issue in large-scale metabolomics studies, which is overcome in offline-techniques with quality controls.[32] These samples are then used for normalization, i.e., to separate the biological variation of interest from unwanted technical variation or other confounding factors, such as exogenous influences. Since such samples are not yet available for real-time breath analysis, a higher degree of standardization of sampling and methodology is required.[33] In future studies, a reference gas mixture could be used to check the instrument performance and thereby reduce technical noise. This might improve the effectiveness of SESI-HRMS for screening for OSA or other metabolic conditions. Ideally, in a next step, calibration with standards of validated and identified biomarkers, such as the ones identified in this study, could be applied using standard addition. The standard addition procedure brings the advantage that in addition to technical fluctuations, matrix effects such as the influence of humidity on ionization efficiency, and

ion suppression effects could be eliminated. However, the biggest challenge here is the availability of gaseous standards as well as the bottleneck of identification of metabolites. One possible approach is the use of permeation tubes.[34] Another confounding factor that might compromise the results are isobaric compounds since no separation step, such as chromatography, is used in real-time SESI-HRMS. To overcome this challenge, a targeted detection of the validated and identified biomarkers could involve MS-MS quantification or a coupling to ion mobility spectrometry and thereby provide a higher molecular specificity.

To date, different scores are derived from questionnaires for an initial approach to OSA screening. Our classification performance is comparable with the performances of the STOP-bang[14] and Berlin scores[15]. The NoSAS score performs slightly better, AUROCs of 0.74 and 0.81 have been reported from two different patient cohorts. However, in terms of sensitivity our results are comparable with the NoSAS score.[16] For screening, this is most relevant, since ideally no subjects with OSA are missed. It has been previously shown that the combination of NoSAS and metabolomics data can improve predictive performance remarkably.[35] Here, real-time breath analysis could speed up diagnosis and make it even less bothersome for the patients. We think that the combination of exhaled breath analysis and the NoSAS score might provide an objective and easy-to-perform assay for screening patients with possible OSA. Only positively tested patients would then need to undergo the time-consuming, costly and inconvenient respiratory polygraphy to confirm or refute the screening result. During screening, even multiple testing would be possible, since the breath test is fast and non-invasive. This might help to overcome the problem of a considerable night-to-night variation of OSA. Further studies, looking at the combination of the NoSAS score and SESI-HRMS are needed.

In addition, exhaled breath metabolomics can provide detailed information about metabolic changes because biological compounds can be identified. Here, we could confirm the association of unsaturated aldehydes, furans and benzothiazole with OSA. These findings are strengthening the hypotheses from our pilot study of increased oxidative stress levels and altered gut microbiota in OSA patients.[24] Moreover, our findings of metabolites such as 2-pentylfuran and 4-hydroxy-2-octenal, correlating with both, ESS and ODI, suggest an association of those metabolites with the sleep deprivation going along with OSA leading to increased sleepiness.

5. Conclusions

In conclusion, we could confirm our previous findings of an OSA specific metabolic breath pattern and validate a panel of 33 biomarkers in a larger and broader cohort of patients with possible OSA. This is the first validation study for breath analysis by SESI-HRMS, bringing this technique an important step closer to its application in clinics. However, being implemented for clinical use, the added value of SESI-HRMS measurements to conventional OSA screening questionnaires, such as NoSAS, should be evaluated in further patient cohorts.

Acknowledgements

We would like to thank Thomas Gaisl, Martin Gaugg, Marloes Maathuis, Sanzio Monti, Martin Osswald and Bettina Streckenbach for valuable scientific discussion. This project is part of the Zurich Exhalomics project, a flagship project of University Medicine Zurich.

Funding

The study was supported by Uniscientia foundation and Schwyzer foundation. Pablo Sinues received funding from Fondation Botnar (Switzerland).

Declarations of interest

M. Kohler declares fees from Bayer, Novartis, Roche, Astra Zeneca, Boehringer Ingelheim, Mundipharma, GSK and Philips outside the submitted work. M. Kohler and P. Sinues are co-founders and board members of the Deep Breath Initiative (DBI), a company that provides services in the field of exhaled breath analysis.

Data and code availability

The original data and code used in this publication are made available in a curated data archive at ETH Zurich (<https://www.research-collection.ethz.ch>) under the DOI 10.3929/ethz-b-000456155.

References

- [1] Malhotra A, White DP. Obstructive sleep apnoea. *Lancet* 2002;360:237–45.
[https://doi.org/10.1016/S0140-6736\(02\)09464-3](https://doi.org/10.1016/S0140-6736(02)09464-3).
- [2] Lévy P, Kohler M, McNicholas WT, Barbé F, McEvoy RD, Somers VK, et al. Obstructive sleep apnoea

- syndrome. *Nat Rev Dis Prim* 2015;1:15015. <https://doi.org/10.1038/nrdp.2015.15>.
- [3] Schwarz EI, Engler A, Kohler M. Exhaled breath analysis in obstructive sleep apnea. *Expert Rev Respir Med* 2017;11:631–9. <https://doi.org/10.1080/17476348.2017.1338950>.
 - [4] Tregear S, Reston J, Schoelles K, Phillips B. Obstructive sleep apnea and risk of motor vehicle crash: systematic review and meta-analysis. *J Clin Sleep Med* 2009;5:573–81.
 - [5] Marin JM, Carrizo SJ, Vicente E, Agusti AG. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet* 2005;365:1046–53. [https://doi.org/10.1016/S0140-6736\(05\)71141-7](https://doi.org/10.1016/S0140-6736(05)71141-7).
 - [6] Nieto FJ, Young TB, Lind BK, Shahar E, Samet JM, Redline S, et al. Association of Sleep-Disordered Breathing, Sleep Apnea, and Hypertension in a Large Community-Based Study. *JAMA* 2000;283:1829. <https://doi.org/10.1001/jama.283.14.1829>.
 - [7] Park JG, Morgenthaler TM, Gay PC. Novel and Emerging Nonpositive Airway Pressure Therapies for Sleep Apnea. *Chest* 2013;144:1946–52. <https://doi.org/10.1378/chest.13-0273>.
 - [8] Weaver TE, Grunstein RR. Adherence to Continuous Positive Airway Pressure Therapy: The Challenge to Effective Treatment. *Proc Am Thorac Soc* 2008;5:173–8. <https://doi.org/10.1513/pats.200708-119MG>.
 - [9] McArdle N, Devereux G, Heidarnejad H, Engleman HM, Mackay TW, Douglas NJ. Long-term use of CPAP therapy for sleep apnea/hypopnea syndrome. *Am J Respir Crit Care Med* 1999;159:1108–14. <https://doi.org/10.1164/ajrccm.159.4.9807111>.
 - [10] Kushida CA, Littner MR, Morgenthaler T, Alessi CA, Bailey D, Coleman J, et al. Practice Parameters for the Indications for Polysomnography and Related Procedures: An Update for 2005. *Sleep* 2005;28:499–523. <https://doi.org/10.1093/sleep/28.4.499>.
 - [11] McNicholas WT. Diagnosis of Obstructive Sleep Apnea in Adults. *Proc Am Thorac Soc* 2008;5:154–60. <https://doi.org/10.1513/pats.200708-118MG>.
 - [12] Roeder M, Bradicich M, Schwarz EI, Thiel S, Gaisl T, Held U, et al. Night-to-night variability of respiratory events in obstructive sleep apnoea: a systematic review and meta-analysis. *Thorax* 2020;thoraxjnl-2020-214544. <https://doi.org/10.1136/thoraxjnl-2020-214544>.
 - [13] Berry RB, Budhiraja R, Gottlieb DJ, Gozal D, Iber C, Kapur VK, et al. 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 2012;8:597–619. <https://doi.org/10.5664/jcsm.2172>.
- [14] Chung F, Yegneswaran B, Liao P, Chung SA, Vairavanathan S, Islam S, et al. STOP Questionnaire. *Anesthesiology* 2008;108:812–21. <https://doi.org/10.1097/ALN.0b013e31816d83e4>.
- [15] Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire To Identify Patients at Risk for the Sleep Apnea Syndrome. *Ann Intern Med* 1999;131:485. <https://doi.org/10.7326/0003-4819-131-7-199910050-00002>.
- [16] Marti-Soler H, Hirotsu C, Marques-Vidal P, Vollenweider P, Waeber G, Preisig M, et al. The NoSAS score for screening of sleep-disordered breathing: a derivation and validation study. *Lancet Respir Med* 2016;4:742–8. [https://doi.org/10.1016/S2213-2600\(16\)30075-3](https://doi.org/10.1016/S2213-2600(16)30075-3).
- [17] de Lacy Costello B, Amann A, Al-Kateb H, Flynn C, Filipiak W, Khalid T, et al. A review of the volatiles from the healthy human body. *J Breath Res* 2014;8:014001. <https://doi.org/10.1088/1752-7155/8/1/014001>.
- [18] Pereira J, Porto-Figueira P, Cavaco C, Taunk K, Rapole S, Dhakne R, et al. Breath Analysis as a Potential and Non-Invasive Frontier in Disease Diagnosis: An Overview. *Metabolites* 2015;5:3–55. <https://doi.org/10.3390/metabo5010003>.
- [19] Schwarz EI, Engler A, Kohler M. Exhaled breath analysis in obstructive sleep apnea. *Expert Rev Respir Med* 2017;11:631–9. <https://doi.org/10.1080/17476348.2017.1338950>.
- [20] Greulich T, Hattessohl A, Grabisch A, Koepke J, Schmid S, Noeske S, et al. Detection of obstructive sleep apnoea by an electronic nose. *Eur Respir J* 2013;42:145–55. <https://doi.org/10.1183/09031936.00091712>.
- [21] Benedek P, Lázár Z, Bikov A, Kunos L, Katona G, Horváth I. Exhaled biomarker pattern is altered in children with obstructive sleep apnoea syndrome. *Int J Pediatr Otorhinolaryngol* 2013;77:1244–7. <https://doi.org/10.1016/J.IJPORL.2013.04.025>.
- [22] Dragonieri S, Porcelli F, Longobardi F, Carratù P, Aliani M, Ventura VA, et al. An electronic nose in the discrimination of obese patients with and without obstructive sleep apnoea. *J Breath Res* 2015;9:026005. <https://doi.org/10.1088/1752-7155/9/2/026005>.
- [23] Kunos L, Bikov A, Lazar Z, Korosi BZ, Benedek P, Losonczy G, et al. Evening and morning exhaled volatile compound patterns are different in obstructive sleep apnoea assessed with electronic nose. *Sleep Breath* 2015;19:247–53. <https://doi.org/10.1007/s11325-014-1003-z>.

- [24] Schwarz EI, Martinez-Lozano Sinues P, Bregy L, Gaisl T, Garcia Gomez D, Gaugg MT, et al. Effects of CPAP therapy withdrawal on exhaled breath pattern in obstructive sleep apnoea. *Thorax* 2016;71:110–7. <https://doi.org/10.1136/thoraxjnl-2015-207597>.
- [25] Gaugg MT, Nussbaumer-Ochsner Y, Bregy L, Engler A, Stebler N, Gaisl T, et al. Real-Time Breath Analysis Reveals Specific Metabolic Signatures of COPD Exacerbations. *Chest* 2019;156:269–76. <https://doi.org/10.1016/j.chest.2018.12.023>.
- [26] Storey JD. A direct approach to false discovery rates. *J R Stat Soc Ser B (Statistical Methodol)* 2002;64:479–98. <https://doi.org/10.1111/1467-9868.00346>.
- [27] Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007;8:118–27. <https://doi.org/10.1093/biostatistics/kxj037>.
- [28] García-Gómez D, Bregy L, Barrios-Collado C, Vidal-de-Miguel G, Zenobi R. Real-Time High-Resolution Tandem Mass Spectrometry Identifies Furan Derivatives in Exhaled Breath. *Anal Chem* 2015;87:6919–24. <https://doi.org/10.1021/acs.analchem.5b01509>.
- [29] García-Gómez D, Martínez-Lozano Sinues P, Barrios-Collado C, Vidal-de-Miguel G, Gaugg M, Zenobi R. Identification of 2-alkenals, 4-hydroxy-2-alkenals, and 4-hydroxy-2,6-alkadienals in exhaled breath condensate by UHPLC-HRMS and in breath by real-time HRMS. *Anal Chem* 2015;87:3087–93. <https://doi.org/10.1021/ac504796p>.
- [30] García-Gómez D, Bregy L, Nussbaumer-Ochsner Y, Gaisl T, Kohler M, Zenobi R. Detection and Quantification of Benzothiazoles in Exhaled Breath and Exhaled Breath Condensate by Real-Time Secondary Electrospray Ionization-High-Resolution Mass Spectrometry and Ultra-High Performance Liquid Chromatography. *Environ Sci Technol* 2015;49:12519–24. <https://doi.org/10.1021/acs.est.5b03809>.
- [31] Barrios-Collado C, Vidal-De-Miguel G, Martinez-Lozano Sinues P. Numerical modeling and experimental validation of a universal secondary electrospray ionization source for mass spectrometric gas analysis in real-time. *Sensors Actuators, B Chem* 2016;223:217–25. <https://doi.org/10.1016/j.snb.2015.09.073>.
- [32] Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat Protoc* 2011;6:1060–83. <https://doi.org/10.1038/nprot.2011.335>.
- [33] Hanna GB, Boshier PR, Markar SR, Romano A. Accuracy and Methodologic Challenges of Volatile

Organic Compound-Based Exhaled Breath Tests for Cancer Diagnosis. *JAMA Oncol* 2019;5:e182815.

<https://doi.org/10.1001/jamaoncol.2018.2815>.

- [34] Guillevic M, Vollmer MK, Wyss SA, Leuenberger D, Ackermann A, Pascale C, et al. Dynamic-gravimetric preparation of metrologically traceable primary calibration standards for halogenated greenhouse gases. *Atmos Meas Tech* 2018;11:3351–72. <https://doi.org/10.5194/amt-11-3351-2018>.
- [35] Lebkuchen A, Carvalho VM, Venturini G, Salgueiro JS, Freitas LS, Dellavance A, et al. Metabolomic and lipidomic profile in men with obstructive sleep apnoea: implications for diagnosis and biomarkers of cardiovascular risk. *Sci Rep* 2018;8:11270. <https://doi.org/10.1038/s41598-018-29727-6>.

Table 1: Study participant characteristics. (AHI: apnea-hypopnea-index; BMI: body mass index; ESS: Epworth Sleepiness Scale; ODI: oxygen desaturation index; FVC: expiratory forced vital capacity; FEV1: forced expiratory volume in one second)

	all participants	OSA			healthy		
		stratification 1 ODI>30 or (ODI>10 & ESS>10)	stratification 2 ODI>30 & ESS>10	stratification 3 ODI>30	stratification 1 ODI<5 or (ODI<10 & ESS<11)	stratification 2 ODI<10 & ESS<11	stratification 3 ODI<10
Number of subjects (N)	149	51	18	36	33	26	47
Age, mean (SD), years	53.3 (13.7)	55 (13.4)	51.6 (10.2)	53.6 (13)	48.5 (14.6)	49.7 (14.3)	49.3 (14.6)
male sex, n (%)	108 (72.5%)	35 (68.6%)	13 (72.2%)	26 (72.2%)	20 (60.6%)	17 (65.4%)	30 (63.8%)
BMI, mean (SD), kg/m ²	30.1 (6.6)	33 (6.5)	33.1 (6.7)	34 (6.5)	27.2 (6.1)	28 (6.4)	27.1 (5.5)
active smoker, n (%)	31 (20.8%)	17 (33.3%)	7 (38.9%)	14 (38.9%)	6 (18.2%)	5 (19.2%)	10 (21.3%)
former smoker, n (%)	53 (35.6%)	16 (31.4%)	4 (22.2%)	9 (25%)	11 (33.3%)	10 (38.5%)	16 (34%)
AHI at diagnosis, mean (SD), events per hour	19.6 (17)	34.8 (18)	42.9 (15)	42.7 (15)	5.9 (5)	5.3 (3)	6.4 (5)
ODI at diagnosis, mean (SD), events per hour	20.8 (17.5)	38.3 (17.5)	46.5 (14.7)	46.8 (13.2)	4.4 (2.7)	5 (2.6)	5.2 (2.6)
ESS at diagnosis, mean (SD), points	8.9 (4.5)	11 (4.8)	14.2 (2.8)	10 (5.2)	7.4 (3.4)	6 (2.4)	9.4 (4.5)
FEV1/FVC, mean (SD)	78.5 (7.7)	77.6 (8.7)	79.1 (7.9)	78.5 (7.6)	78 (5.5)	78 (5.6)	78.1 (5.5)
FVC, mean (SD), % predicted	101 (16.7)	97.5 (18.7)	96.1 (15)	96.2 (17.2)	108.5 (15.5)	106.3 (16.1)	105.4 (16)
FEV1, mean (SD), % predicted	99 (17.6)	95.3 (20.8)	94.9 (18.5)	94.5 (19.1)	103.9 (14.5)	102.4 (15)	101.2 (14.4)

All rights reserved. No reuse allowed without permission.

Table 2: A panel of validated OSA biomarkers. 33 of the previously detected[24] biomarkers for OSA show a significant correlation with the ODI or ESS or significant differences between OSA patients and control subjects in the validation cohort (significance level: $p < 0.05$). (unsat. aldehydes: unsaturated aldehydes)

metabolite identification		Pilot Study (Training Set)						Validation Study											
name	class	m/z	significant difference between groups	correlation with delta ODI	Pearson correlation coefficient	selected as predictor for classification	m/z	Spearman correlation coefficient ODI	p correlation ODI	q correlation ODI	Spearman correlation coefficient ESS	p correlation ESS	q correlation ESS	p between groups (stratification 1)	q between groups (stratification 1)	log2 fold change between groups (stratification 1)	p between groups (stratification 2)	q between groups (stratification 2)	log2 fold change between groups (stratification 2)
2-pentenal ([M+NH ₄] ⁺)	unsat aldehydes	102.0913	yes	no		yes	102.0912	0.23	0.005	0.017	0.07	0.430	0.34	0.029	0.087	0.43	0.003	0.004	0.75
4-hydroxy-2-heptenal	unsat. aldehydes	129.0908	yes	yes	0.48	no	129.0908	0.20	0.015	0.020	0.13	0.110	0.26	0.022	0.081	0.42	0.003	0.004	0.83
4-hydroxy-2-octenal ([M+NH ₄] ⁺)	unsat. aldehydes	143.1063	no	yes	0.42	no	143.1064	0.16	0.048	0.036	0.13	0.110	0.26	0.036	0.088	0.12	0.006	0.005	0.25
4-hydroxy-2-octenal	unsat. aldehydes	160.1329	yes	yes	0.42	no	160.1331	0.18	0.026	0.023	0.16	0.050	0.19	0.020	0.081	0.47	0.001	0.003	0.85
2-undecenal	unsat. aldehydes	169.1584	no	yes	0.38	no	169.1586	0.11	0.170	0.072	0.09	0.300	0.34	0.360	0.246	0.16	0.040	0.018	0.33
2-ethylfuran	furanes	97.0647	yes	no		yes	97.0646	0.23	0.005	0.017	0.10	0.250	0.34	0.009	0.068	0.38	0.005	0.004	0.80
2-propylfuran	furanes	111.0803	yes	yes	0.4	yes	111.0803	0.27	0.001	0.011	0.10	0.210	0.34	0.003	0.068	0.41	0.004	0.004	0.56
2-butylfuran	furanes	125.0958	yes	yes	0.44	no	125.0961	0.23	0.004	0.017	0.13	0.130	0.28	0.005	0.068	0.26	0.001	0.004	0.48
2-pentylfuran	furanes	139.1116	yes	yes	0.38	yes	139.1116	0.11	0.197	0.074	0.17	0.040	0.19	0.040	0.088	0.26	0.002	0.004	0.57
benzothiazole	thiazoles	136.0216	yes	no		no	136.0213	0.20	0.016	0.020	-0.04	0.620	0.35	0.019	0.081	0.40	0.025	0.014	0.50
4-(hexyloxy)phenol	benzenoids	195.1379	no	yes	0.38	no	195.1379	0.15	0.076	0.048	0.06	0.450	0.34	0.111	0.143	0.34	0.010	0.008	0.62
unknown	unknown	53.0391	no	no		yes	53.0374	0.10	0.237	0.079	0.08	0.370	0.34	0.314	0.223	0.19	0.032	0.015	0.48
unknown	unknown	79.0409	no	no		yes	79.0392	0.20	0.016	0.020	-0.06	0.490	0.34	0.094	0.143	0.51	0.067	0.024	0.71
unknown	unknown	81.0525	yes	no		no	81.0525	0.06	0.466	0.120	0.19	0.020	0.19	0.153	0.152	0.46	0.051	0.021	0.74
unknown	unknown	83.0854	no	no		yes	83.0853	0.19	0.021	0.022	0.16	0.050	0.19	0.037	0.088	0.27	0.030	0.015	0.48
unknown	unknown	93.0574	no	no		yes	93.0548	0.19	0.022	0.022	-0.10	0.240	0.34	0.213	0.184	0.43	0.179	0.051	0.52
unknown	unknown	103.0943	no	no		yes	103.0952	0.20	0.013	0.020	0.05	0.550	0.34	0.047	0.088	0.55	0.019	0.011	0.95
unknown	unknown	122.0835	yes	no		no	122.0806	0.21	0.012	0.020	0.03	0.680	0.37	0.010	0.068	0.49	0.005	0.004	0.80
unknown	unknown	124.0835	yes	no		no	124.0838	0.14	0.090	0.053	0.15	0.060	0.22	0.024	0.081	0.26	0.036	0.017	0.41
unknown	unknown	128.0701	no	no		yes	128.0703	0.13	0.114	0.058	0.19	0.020	0.19	0.042	0.088	0.27	0.002	0.004	0.35
unknown	unknown	136.0471	yes	no		yes	136.0511	-0.06	0.485	0.123	0.19	0.020	0.19	0.960	0.446	-0.18	0.401	0.084	-0.03
unknown	unknown	149.0971	yes	no		no	149.0959	0.11	0.176	0.072	0.21	0.010	0.19	0.126	0.144	0.30	0.019	0.011	0.38
unknown	unknown	152.0699	yes	yes	0.42	no	152.0705	0.13	0.104	0.055	0.02	0.820	0.41	0.178	0.164	0.23	0.045	0.02	0.55
unknown	unknown	158.1241	yes	no		no	158.125	0.16	0.048	0.036	0.14	0.090	0.25	0.044	0.088	0.12	0.003	0.004	0.67
unknown	unknown	165.1272	no	no		yes	165.1273	0.19	0.019	0.022	0.10	0.240	0.34	0.046	0.088	0.41	0.005	0.004	0.54
unknown	unknown	175.1117	yes	no		yes	175.1134	0.21	0.010	0.020	0.16	0.050	0.19	0.007	0.068	0.41	0.001	0.003	0.64
unknown	unknown	182.0897	no	yes	0.38	no	182.0809	0.12	0.146	0.071	0.14	0.090	0.25	0.064	0.11	0.18	0.001	0.004	0.55
unknown	unknown	207.1378	no	yes	0.38	no	207.1381	0.15	0.065	0.043	0.02	0.820	0.41	0.245	0.204	0.35	0.028	0.015	0.58
unknown	unknown	209.1168	no	yes	0.4	no	209.1173	0.18	0.025	0.023	-0.03	0.690	0.37	0.111	0.143	0.54	0.026	0.014	0.75
unknown	unknown	209.1536	yes	no		no	209.1536	0.20	0.016	0.020	0.03	0.720	0.37	0.057	0.102	0.37	0.012	0.008	0.68
unknown	unknown	210.1568	no	no		yes	210.1568	0.24	0.004	0.017	0.02	0.830	0.41	0.020	0.081	0.50	0.014	0.01	0.67
unknown	unknown	211.1325	no	yes	0.38	no	211.1328	0.20	0.015	0.020	0.17	0.040	0.19	0.025	0.081	0.42	0.002	0.004	0.67

unknown	unknown	228.0686	no	no		yes	228.0642	-0.18	0.030	0.025	-0.05	0.580	0.34	0.119	0.143	-0.37	0.119	0.04	-0.40
---------	---------	----------	----	----	--	-----	----------	-------	-------	-------	-------	-------	------	-------	-------	-------	-------	------	-------