Talanta 256 (2023) 124299

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Determination of global chemical patterns in exhaled breath for the discrimination of lung damage in postCOVID patients using olfactory technology

Blanca Nohemí Zamora-Mendoza [a], Hannia Sandoval-Flores [a], Maribel Rodríguez-Aguilar [b], Carlos Jiménez-González [a], Luz Eugenia Alcántara-Quintana [c], Alejandra Abigail Berumen- Rodríguez [a], Rogelio Flores-Ramírez [c,*]

[a] Faculty of Medicine-Center for Applied Research on Environment and Health (CIAAS), Autonomous University of San Luis Potosí, Avenida Sierra Leona No. 550, CP 78210, Colonia Lomas Segunda Sección, San Luis Potosí, Mexico
[b] Department of Pharmacy, Health Sciences Division. University of Quintana Roo, Quintana Roo, Mexico
[c] CONACYT Research Fellow, Coordination for Innovation and Application of Science and Technology (CIACYT), Autonomous University of San Luis Potosí, Avenida Sierra Leona No. 550, CP 78210, Colonia Lomas Segunda Sección, San Luis Potosí, Mexico

ARTICLE INFO

Handling Editor: Prof. J.-M. Kauffmann

Keywords:
postCOVID screening
Electronic nose
Volatile organic compounds
Exhaled breath
Pulmonary sequelae

ABSTRACT

The objective of this work was to evaluate the use of an electronic nose and chemometric analysis to discriminate global patterns of volatile organic compounds (VOCs) in breath of postCOVID syndrome patients with pulmonary sequelae. A cross-sectional study was performed in two groups, the group 1 were subjects recovered from COVID-19 without lung damage and the group 2 were subjects recovered from COVID-19 with impaired lung function. The VOCs analysis was executed using a Cyranose 320 electronic nose with 32 sensors, applying principal component analysis (PCA), Partial Least Square-Discriminant Analysis, random forest, canonical discriminant analysis (CAP) and the diagnostic power of the test was evaluated using the ROC (Receiver Operating Characteristic) curve. A total of 228 participants were obtained, for the postCOVID group there are 157 and 71 for the control group, the chemometric analysis results indicate in the PCA an 84% explanation of the variability between the groups, the PLS-DA indicates an observable separation between the groups and 10 sensors related to this separation, by random forest, a classification error was obtained for the control group of 0.090 and for the postCOVID group of 0.088 correct classification. The CAP model showed 83.8% of correct classification and the external validation of the model showed 80.1% of correct classification. Sensitivity and specificity reached 88.9% (73.9%–96.9%) and 96.9% (83.7%–99.9%) respectively. It is considered that this technology can be used to establish the starting point in the evaluation of lung damage in postCOVID patients with pulmonary sequelae.

1. Introduction

The World Health Organization refers to the postCOVID condition as a set of prolonged symptoms lasting at least 2 months that cannot be explained by an alternative diagnosis, among the most common symptoms are fatigue, dyspnea, mental fog problems, persistent cough, chest pain, slurred speech, muscle aches, loss of smell, depression or anxiety, and fever, among many others [1]. Such symptoms may be permanent and/or fluctuating in frequency and intensity over time, remaining in some cases for more than twelve months [2,3], estimating a global pooled prevalence of postCOVID in 43% of the population [4].

This disease affects multiple organs and there is great concern in the sequelae on the respiratory system [5]. Although the pathophysiology of the persistent symptoms is unknown, they have been attributed to the high viral load where immune cells receive a high stimulus for the expression of inflammatory mediators, generating cytokine storms which in turn cause damage to the alveolar structures and their function, therefore, complications may occur in the acute phase of the infection related to intubation, prolonged rest, malnutrition, among others [6].

Subsequently, after the acute phase, the lungs continue in a pro-inflammatory and pro-thrombotic state leading to a fragile and vulnerable epithelium, due to the expression of immune mediators for virus

* Corresponding author. Avenida Sierra Leona No. 550, CP 78210, Colonia Lomas Segunda Sección, San Luis Potosí, Mexico.
E-mail address: rfloresra@conacyt.mx (R. Flores-Ramírez).

https://doi.org/10.1016/j.talanta.2023.124299
Received 7 November 2022; Received in revised form 18 January 2023; Accepted 20 January 2023
Available online 20 January 2023

clearance and repair of damage in the lung parenchyma, which in turn compromises the local and/or systemic oxygenation status, generating the release of damage-associated molecular patterns (DAMPs) and a constant stimulus of the repair-damage cycle, among other mechanisms, resulting in thickening of the epithelial basement membrane, remodeling, diffuse alveolar damage, and pulmonary fibrosis during the post-COVID phase [7–11]. In this way, it is important to evaluate the pulmonary effects in post-COVID due to the probable pulmonary interstitial fibrosis as it has negative effects on the quality of life of people [12].

Clinical follow-up of postCOVID patients is assessed by pulmonary function and/or imaging tests. The initial step is chest radiography; however, in patients with residual radiographic or functional damage, high-resolution chest CT is indicated not only for further characterization of the anatomical regions of the lung parenchyma, but also to establish a baseline for the follow-up and as one of the current methods for pulmonary fibrosis diagnosis reaching a classification by its anatomical pattern [13]. Although these images are key to diagnosis, they do not reveal functionality, only anatomy. Hence, techniques to evaluate functionality have acquired relevance, among the most studied in the postCOVID era is the evaluation of carbon monoxide diffusing capacity (DLCO), which has been described as one of the most important pulmonary evaluation methods in the postCOVID era since it allows the identification of conditions in alveolar gas exchange associated with lung parenchymal damage caused by SARS-CoV 2 , which on a larger scale can be reflected in the alteration of pulmonary flows and volumes through forced spirometry [14,15]. However, a high demand, low availability, and accessibility to such diagnostic methods as well as lack of follow-up, highlight the urgency of care protocols for people with postCOVID, particularly in pulmonary health.

In this regard, our group has demonstrated, in different studies, the usefulness of evaluating volatile organic compounds (VOCs) in exhaled breath as a screening method for chronic diseases such as lung cancer, breast cancer, preeclampsia, diabetes and chronic obstructive pulmonary disease, among others [16–21]. This biological matrix consists of a mixture of organic compounds that come from cellular metabolism and have low solubility in the blood, which means that they can be expelled during exhalation and thus be determined by analytical techniques [22].

These chemical patterns represent normal physiological or specific pathophysiological conditions. In a recent study by our working group, we identified a global chemical pattern of volatile organic compounds in exhaled breath capable of discriminating between patients with COVID-19 and controls using electrochemical nanosensors, the automated learning model presented a sensitivity of 100% and a specificity of 97.6% [23], In addition, a pilot study was applied in which patients recovered from infection but with pulmonary sequelae were followed up, identifying VOC patterns that allow to discriminate between people with infection, people recovered with pulmonary sequelae and people recovered without pulmonary sequelae [24]. This represents a practical screening approach capable of quickly identifying patients suspected of having it and providing useful epidemiological information to guide community health strategies in the context of COVID-19.

An important point of this methodology is that we have demonstrated in different studies the assessment of lung damage associated with various pathologies such as COPD and lung cancer [18–20]. In this way, the use of olfactory technology consisting of functionalized nanosensors in the adsorption of VOCs can be a useful method to assess lung damage in patients during postCOVID which is fast, easy to apply, it can reach a high number of people and at a low cost. The objective of this study was to develop a methodology to evaluate global chemical patterns in exhaled breath for the discrimination and classification of patients with persistent lung damage caused by COVID.

2. Material and methods

2.1. Study design

This study was approved by the research ethics committee of the Hospital General de México "Dr. Eduardo Liceaga" with protocol number 023/058OHGM/SLP/22, in compliance with the International Ethical Guidelines for Biomedical Research Involving Human Subjects of the Council for International Organizations of Medical Sciences (CIOMS). The participation of the study subjects was voluntary, obtaining the signature of an informed consent letter.

The study design consisted of an analytical cross-sectional study, the participants were recruited from health care centers in Mexico who attended the postCOVID follow-up evaluation, they were randomly recruited and divided into two working groups for the construction of the predictive model: group 1) postCOVID patients without impaired pulmonary function: subjects recovered from COVID-19 without lung damage and group 2) postCOVID patients with impaired pulmonary function: subjects recovered from COVID-19 with impaired lung function. For the sample size calculation the conservative values already proposed by Cohen [25] were used to detect a standardized difference between the means of the groups equivalent to an $r = 0.5$ with a power of 0.8, resulting in 80 cases per group,

The study inclusion criteria for the postCOVID patients without impaired pulmonary function were people between 18 and 70 years old, of both genders, with a diagnosis of COVID-19 confirmed by RT-qPCR, with an evolution time of 4 weeks from the onset of symptoms, with normal pulmonary function test results by pre- and post-bronchodilator forced spirometry or, with no history of respiratory symptoms or clinical data suggestive of ventilatory abnormalities on pulmonary examination. The proposed exclusion criteria were: pregnant patients, patients with confirmed or suspected pulmonary infection different from COVID-19 (influenza, tuberculosis or other infectious disease), patients with a history of thoracic or ocular surgery in the last 3 months, cognitive and/or psychomotor disability that limits breath sampling, as well as spirometric maneuver, rib cage malformation, subjects who did not present themselves in the necessary conditions for sampling and spirometry, subjects who at the time of the evaluation showed signs and symptoms of acute infectious disease and/or reported contact with a person confirmed or suspected of having COVID-19 in the last seven days, furthermore, data from subjects who failed to complete the pulmonary function tests, obtained an insufficient breath sample and/or had incomplete clinical information were also eliminated.

For the postCOVID patients with impaired pulmonary function, the same exclusion and inclusion criteria were established as for the control group, except that they obtained an abnormal spirometry result, with or without respiratory signs and symptoms.

All participants had an RT-qPCR- or negative antigen test to exclude reinfection at the moment of breath sample collection and pulmonary evaluation.

2.2. Pulmonary function evaluation and health questionnaire

The pulmonary function evaluation of the two groups was performed by forced spirometry (pre- and post-bronchodilator). Subjects, were classified as impaired lung function when the Forced Expiratory Volume/Forced Vital Capacity (FEV_1/FVC) ratio was greater than 80% and the FVC was below 80% of the predicted value, suggesting a spirometric pattern of lung restriction, characteristic of COVID-19 impairment. Likewise, an obstructive pattern was considered when the FEV_1/FVC ratio was below 80% with or without a significant post-bronchodilator response in FEV_1. The severity of the obstruction and severity were classified according to the parameters established by international guidelines. Forced spirometry tests were performed with a certified portable spirometer; EasyOne®, the evaluations were performed by health personnel certified by the National Institute for Occupational

Safety and Health (NIOSH) and the National Institute of Respiratory Diseases (INER), in addition, the maneuvers complied with the quality criteria standardized by the American Thoracic Society/European Respiratory Society ATS/ERS [26], following the biosafety guidelines established for the prevention of SARS-CoV 2 transmission in pulmonary function laboratories. Likewise, the cut-off points result interpretation were those established for the Mexican population. Moreover, other sociodemographic variables were collected such as: age, sex, and clinical data such as: respiratory symptoms, supplemental oxygen use, hospitalization, pulse oximetry, comorbidities, pathological history, pharmacological treatment, evolution time, number of previous contagions, vaccination against SARS-CoV 2 . In addition to the above, evaluation of thoracoabdominal mobility and oxygen saturation during a 6-min walk, supported the clinical correlation of lung damage.

2.3. Exhaled breath sampling

Exhaled breath sampling was performed based on previous studies by our research group [24]. The participant was asked to relax and take three deep inhalations, then exhale through a straw into a breath collection bag, which consisted of a metalized polypropylene bag of approximately 250 mL with a hermetic seal, this material was purged with ultra-pure nitrogen. The participant criteria for sample collection were the following conditions: i) minimum fasting of 4 h, ii) no smoking before the study (minimum of 2 h), iii) no oral hygiene and iv) before taking any medication, the oximetry and spirometry parameters will be taken at the time of the study. Samples were preserved and transported cold for analysis.

2.4. Olfactory technology

A Cyranose 320 (Sensigent ®, California, US) electronic nose was used to determine the chemical pattern of the breath of the study groups. This technology has 32 chemical nanosensors that have different VOC adsorption properties, producing different degrees of response due to their polymeric composition (polyvinyl butyral, polyvinyl acetate, polystyrene, and polyethylene oxide) and the conductive nanoparticles (carbon nanotubes) of which they are composed.

For sample processing, each sample was incubated at 37 °C for 5 min before analysis. The electronic nose configuration consisted of a constant flow rate of 120 mL/min for 40 s of baseline recording with ultra-pure nitrogen with a sample analysis period of 45 s, subsequently the flow rate was increased to 180 mL/min of ultra-pure nitrogen for sample line purge and air input, with a substrate temperature of 45 °C. During the analysis, the instrument recorded the increase in electrical resistance of each sensor as a result of the adsorption of volatile compounds on the sensors and the responses of methanol and undecane standards at a concentration of 1 ppm were recorded (Tables 1 and 2 supplementary material).

2.5. Multivariate statistical analysis and automated learning methods

Different multivariate statistical analyses were performed using the increase in the resistance of the 32 sensors obtained from the fractional difference: $\Delta R/Ro = (Rmax-Ro)/Ro$ where R is the maximum system response of each sensor, and Ro is the reference reading of each sensor (ultra-pure nitrogen).

Subsequently, for all analyses, a normalization of the sum was performed to reduce the environmental effect by dividing the response of each sensor by the sum of the absolute values of the response of each sensor: $(\Delta R/Ro)i = (\Delta R/Ro)i/\sum |\Delta R/Ro|j$. Additionally, auto-scaling was performed to remove the effects of the magnitude of the sensor responses by subtracting the average of the samples from the individual response of each sample and dividing by the standard deviation of the samples.

The data were analyzed by chemometric methods, Principal Component Analysis (PCA) was used as an unsupervised methodology, this analysis expresses a set of variables in a set of linear combinations of uncorrelated factors with each other and the variability of the data, in this way, it allows to represent the original data in a space of lower dimension than the original space, and it limits to the maximum the loss of information [27]. Subsequently, a Partial Least Square-Discriminant Analysis (PLS-DA) was performed, this is a supervised method that uses the multiple linear regression technique to find the maximum covariance direction between a data set (X) and a class membership (Y), as a supervised method, PLS-DA can perform both classification and feature selection by obtaining the variable importance in projection (VIP), this is obtained from a weighted sum of squares of the PLS loadings, the weights are based on the amount of variance Y explained in each dimension and important features are selected [28].

Additionally, the supervised random forest method was used, which uses a set of classification trees, each of which grows by randomly selecting features from a bootstrap sample on each branch. Class prediction is based on the majority vote of the ensemble. During tree construction, approximately one third of the instances are left out of the bootstrap sample. These data are used as a test sample to obtain an unbiased estimate of the classification error (OOB). The importance of the variable is assessed by measuring the increase in OOB error when permuted.

Finally, Canonical Analysis of Principal Coordinates (CAP) was used to order the matrices, and to further determine the level of misclassification between sampling regions, the leave-one-out method was applied

Table 1
Sociodemographic and clinical characteristics.

Parameters	postCOVID patients with impaired pulmonary function (n = 157)	postCOVID patients without impaired pulmonary function (n = 71)
Age (years)	53 ± 6	38 + 4
Female (%)	37	53
Male (%)	62	46
BMI	26 ± 3	21 + 4
FVC (% predicted)	75 ± 14 *	96 + 6 *
FEV_1 (% predicted)	71 ± 17*	97 + 7 *
Chest mobility (cm)	2 ± 1	3 + 1 *
Abdominal mobility (cm)	5 ± 2	7 + 1
Hospitalizated, (%)	18	2
Hypoxic episode (%)	97	2
Days of oxygen use	48 + 6	3 + 3
Smoking, n (%)	58	29
Asthma	30	3
Comorbidity	73	46
Fatigue (%)	74	38
Dyspnea (%)	68	5
Cough (%)	35	3
Headache (%)	43	21
Myalgia (%)	57	24
Arthralgia (%)	22	15
Brain fog	55	39

BMI: Body Mass index, FVC: Forced Vital Capacity, FEV_1: Forced Expiratory Volume in 1 s.

Table 2
Comparison of sensitivity and specificity according to the classification method performed.

Classification method	Sensitivity (CI 95%)	Specificity (CI 95%)
Random forest	81.4 (69.1–85.4%)	88.1 (78.2–95.3%)
Canonical Analysis of Principal Coordinates	88.9 (73.9–96.9%)	96.9 (83.7–99.9%)

to the variables in the canonical space (using a K-fold of n = 228) to predict group associations and thus obtain the overall classification success rates, using a value of m = 29 [23].

The performance of the CAP model was evaluated by ROC using the PC1 axis values obtained from the training model, with a 95% confidence interval analysis, and the threshold value with the highest specificity/sensitivity ratio was selected. An external validation was conducted by selecting 70% of the population for group definition and the other 30% were randomly selected to validate the model [23].

Descriptive statistical parameters were calculated, statistical analysis was carried out using the statistical software Graphpad Prism 5.0. For chemometric analyses, the statistical programs MetaboAnalyst 5.0 (https://dev.metaboanalyst.ca/MetaboAnalyst/home.xhtml) and PRIMER v7® statistical software with PERMANOVA ad in were used.

3. Results

Table 1 shows the characteristics of the study participants. A total of 228 participants were obtained, 171 for the post-COVID group and 71 for the control group. The group 2 had an average age of 53.4 years, 62.3% were men. Significant differences were found between the values of FVC and FEV 1, thoracic mobility and abdominal mobility of the group 2 versus group 1. The most common symptoms in the group 2 were cough (35.6%), headache (43.5%), cognitive damage (55.4%), myalgia (57%), dyspnea (68.3%) and fatigue (74.2%). The time range of symptoms varied between 1 and 12 months and 97% presented hypoxia during the acute episode in a range between 45 and 85%.

In the group 2, the spirometry values indicated that 68.5% presented an obstructive pattern, 26.3% a suggestive pattern of restriction and 5.2% a normal pattern, while in the control group only 4.2% presented an obstructive pattern and 95.8% a normal pattern, additionally, presented a higher frequency of cardiopulmonary comorbidities in 57% compared to the group 1 with 38.5%. Furthermore, through linear regression analysis, it was identified that the variables that showed the greatest contribution to the forced vital capacity less than 80% of the predicted value were: exposure time to risk factors for lung damage (p, 0.01) within which smoking and exposure to wood smoke were the most frequent (35%) and on the other hand the postCOVID functional status (p, 0.00) and hypoxia (p, 0.02). It is worth noting that the degree of perceived dyspnea showed no association with the restrictive spirometric pattern, making evident a possible silent hypoxia that has remained in some subjects in this study.

Chemometric analyses were performed to establish the differences between groups, Fig. 1 shows the PCA results, a good separation between groups can be visualized which indicates VOC patterns in the exhaled breath, this analysis explains up to 84% of the variability with 3 PCs (PC_1 = 67.1%, PC_2 = 10.4%, PC_3 = 6.5%).

Fig. 2 shows the supervised PLS-DA analysis, showing discrimination between the groups, this analysis allows to establish the significant variables of importance for the groups, with respect to the postCOVID patients with impaired pulmonary function the sensors in order of importance are S17, S26, S32, S12, S9, S25, S22, S21, S10 and S24.

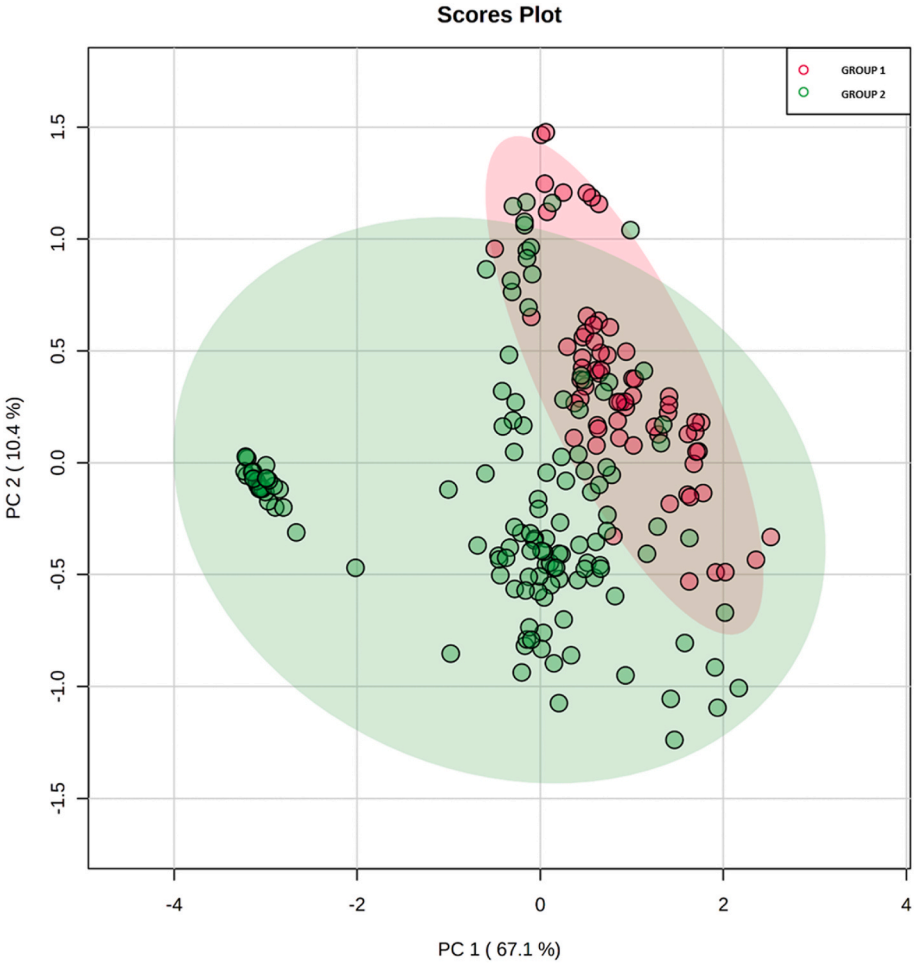

Fig. 1. Principal Component Analysis plot of the study groups. Red circle: Group 1 postCOVID patients without impaired pulmonary function, and green circle: Group 2 postCOVID patients with impaired pulmonary function. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

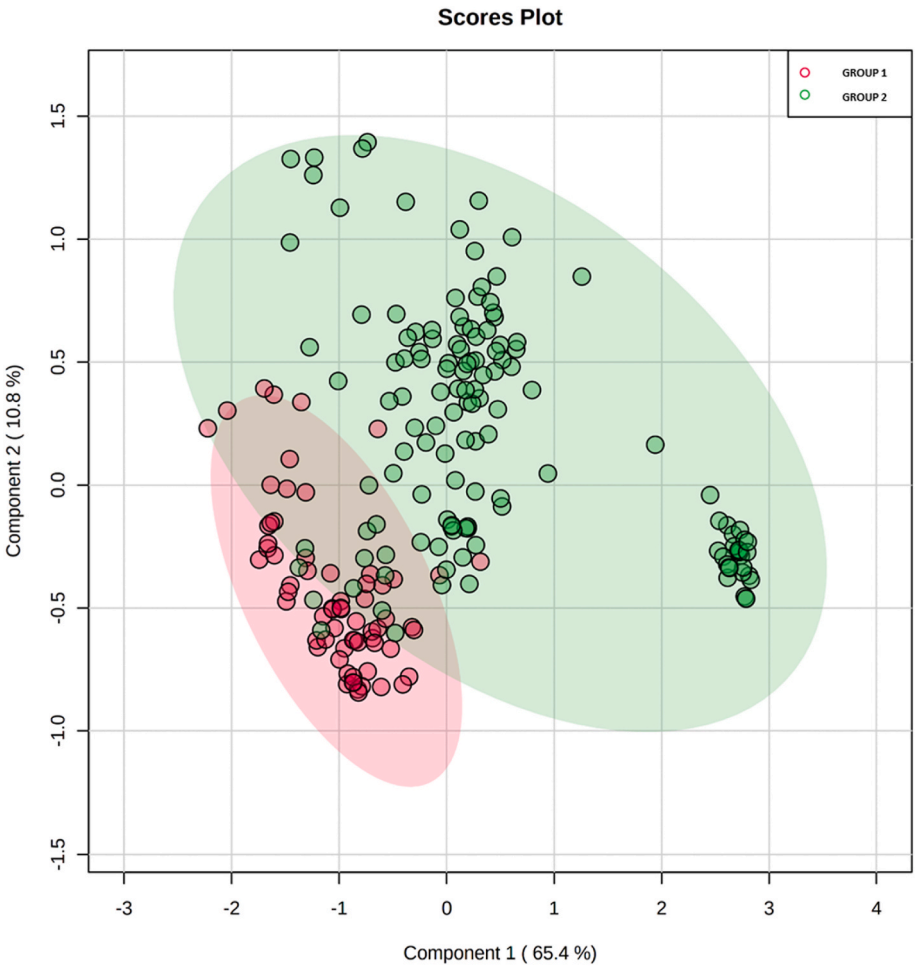

Fig. 2. Partial Least Squares - Discriminant Analysis (PLS-DA) plot of the study groups. Red circle: Group 1 postCOVID patients without impaired pulmonary function, and green circle: Group 2 postCOVID patients with impaired pulmonary function. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

With respect to random forest classification, we obtained a classification error (ODB) for the group 1 of 0.090 and for thegroup 2 of 0.088 correct classification.

The CAP analysis of postCOVID and controls shows a separation of VOCs on the CAP 1 axis with a correlation coefficient of 0.89 and 83.8% of correct classification (Fig. 3), the external validation of the CAP model obtains 80.1% of correct classification.

Furthermore, with the values generated in the CAP 1 score, the cut-off point of −0.029 was established for group 1 vs group 2, which provided 88.9% sensitivity (confidence intervals at 95%: 73.9–96.9%) and 96.9% specificity (confidence intervals at 95%: 83.7–99.9%) (Fig. 4).

Table 3 shows the correlation values between the CAP 1 value and the sensors. For group 1, 20 sensors are positively associated, while for group 2, 5 sensors are negatively associated.

4. Discussion

This investigation demonstrated the potential use of the electronic nose for the discrimination of people with pulmonary sequelae due to COVID and people who presented COVID-19 and no sequelae, both groups well characterized by spirometry.

Since the beginning of the COVID 19 pandemic, several research groups focused on the use of the electronic nose as a potential method for the identification of symptomatic and asymptomatic individuals with COVID 19 in exhaled breath [29–32], most studies have been based on the hyperinflammatory response caused by the SARS-CoV 2 virus in organs [33], particularly in the lung, which alters the volatile and semivolatile metabolite patterns identified in this biological matrix [24]. During the course of the pandemic, more has been learned about this virus, and while the acute phase is important due to the severity of death, patients, who manage to overcome the disease, may present prolonged heterogeneous symptoms which require identification for personalized treatment, as well as continued support [34].

One of the most serious complications in postCOVID syndrome is observed in the respiratory tract; the development of pulmonary fibrosis has been described [35]. Similarly, this pulmonary damage can cause changes in metabolite patterns, e.g., Cui et al. conducted a follow-up study of patients with COVID 19 for 1 year of serum metabolites, an increase in sphingosine-1-phosphate was reported, which is associated with inflammatory factors [36], additionally it plays an important role in tissue fibrosis [37]. On the other hand, normally, the diagnosis of this type of lung damage is based on clinical, radiological, and pathological findings [38]. However, they present problems of interpretation due to lack of specialists, health inequity due to lack of availability of equipment in developing countries, among others [39]. Thereby, the metabolic changes present during postCOVID, represent an important approach to diagnosis and it can be added as respiratory function tests, since one rubric is “how it looks” and another is “how it works” [40].

Two pilot studies have been conducted up to the time of the scientific literature review [24,41]. The study proposed by Nidheesh et al. discussed the discrimination ability in the exhaled breath of 24 postCOVID patients, 24 healthy subjects and 12 subjects with asthma, the

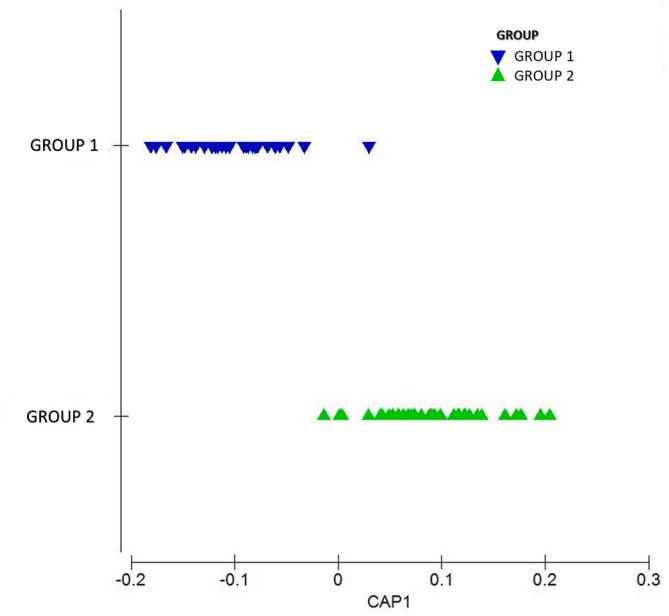

Fig. 3. CAP model of patients with and without lung function impairment. Blue triangle: Group 1 postCOVID patients without impaired pulmonary function, and green triangle: Group 2 postCOVID patients with impaired pulmonary function. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

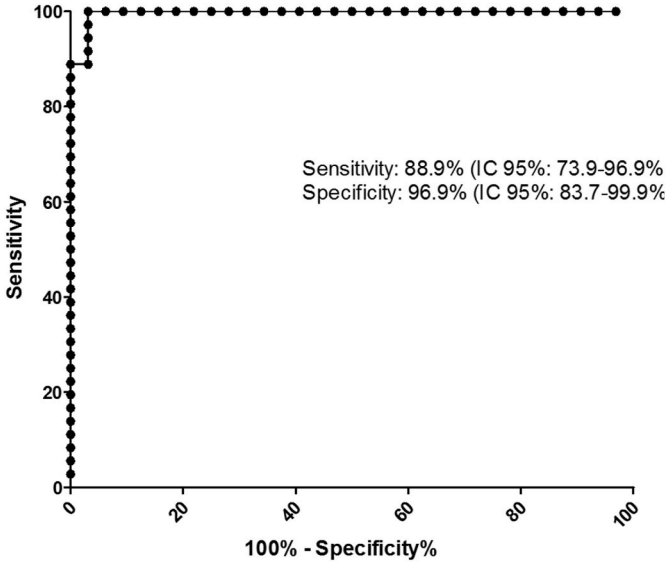

Fig. 4. ROC curve for screening study groups using the CAP1 axis.

preliminary result was that through a supervised algorithm using kNN the prediction of the model reaches 100% sensitivity and specificity [41], although, this result is encouraging, the authors do not show how the classification of lung damage in postCOVID patients was performed. On the other hand, they point out that the healthy subjects did not report having suffered from COVID 19, this condition may generate biases at the time of analysis, we consider relevant a clinical assessment with standardized pulmonary evaluation methods, given that the differences may be due to the conditions of the control group, however, the pilot study presents an important guideline on the potential use of the electronic nose.

In our analysis, interesting results were obtained in the spirometry patterns, we demonstrated that there is discrimination between the controls with normal patterns with respect to the restrictive, mixed, normal and obstructive patterns of postCOVID patients, indeed, we can observe in Supplementary Fig. 1, the clustering of patterns in the post-COVID group, indicating that the separation is due to the condition of lung damage associated with postCOVID, this was also proved with patients with asthma, in Supplementary Fig. 2, we observe that patients with postCOVID and asthma differ from controls with asthma and achieve a good classification based on the condition of lung damage associated with postCOVID.

Although the number of samples is considerably high to meet the objective of the study, one of the weaknesses of the study was the lack of evaluation of other variables for pulmonary function that have been strongly associated with postCOVID syndrome, e.g., in our study we evidenced a silent hypoxia that has remained in some subjects in this study, which is why the evaluation of pulmonary diffusing capacity of carbon monoxide (DL_{CO}) has become one of the most important pulmonary function tests for evaluation in postCOVID [42], this requires an increase in the number of sample and the use of highly specialized equipment. On the other hand, it would be important to identify specific metabolites by techniques such as mass spectrometry, which would allow to understand altered biochemical pathways in postCOVID and to increase the sensitivity of the test by using specific sensors for these compounds. However, one of the strengths of our study was the characterization of patients by spirometry, which allows to categorize lung damage and compare it with the results obtained by the electronic nose.

Another strength has been the evaluation of clinical variables that support the semiology of lung damage in the postCOVID patient, being recognized as a heterogeneous pathological entity in which sometimes the patient does not refer symptoms typically correlated with lung damage, which is why during the clinical evaluation, it could not be properly diagnosed. This fact highlights the importance of this tool as an auxiliary in the differential diagnosis of lung damage.

Therefore, in this study we were able to establish a model for discrimination of volatile organic compound profiles in exhaled breath between postCOVID patients with and without lung damage by an electronic nose and chemometric strategies. We obtained a test that has a good classification, the analysis is less than 3 min, compact, portable and it has the potential to be used on a massive scale, especially in less favored regions where there is a lack of equipment or adequate medical infrastructure for the identification of pulmonary sequelae.

Last but not least, it is critical to identify risk factors for lasting lung damage and to define at-risk patients who should receive specialized respiratory follow-up after infection. This technology will support physicians and researchers to establish rapid mitigation strategies for post-pandemic scenarios and to develop therapeutic strategies and pulmonary rehabilitation.

Moreover, we believe that this technology can be used to establish the starting point for lung damage evaluation in postCOVID patients, and due to the simplicity of the technique it can support strategies for patient identification and monitoring with pulmonary sequelae during rehabilitation. Further research efforts are required for the implementation of this technology in clinical management and support in the post-pandemic scenario.

Author contribution

BNZM: Conceptualization, Analytical methods, writing; HSF: Conceptualization and Analytical methods, Writing and Editing; MRA: Conceptualization, Sampling; CJG: Supervision and Validation, LEAQ: Sampling, biosafety Supervision; AABR: Sampling and Supervision; RFR: Conceptualization, Analytical methods, Writings, Editing and Funding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

Table 3
Spearman correlation analysis between CAP 1 scores and sensors per group.

	Group 1			Group 2		
Sensor	Spearman r	P value (two-tailed)	Significative	Spearman r	P value (two-tailed)	Significative
S1	0.213	0.084	–	−0.132	0.153	–
S2	0.233	0.059	–	−0.154	0.096	–
S3	0.267	0.030	[a]	−0.137	0.138	–
S4	0.258	0.036	[a]	−0.167	0.071	–
S5	0.284	0.020	[a]	−0.0665	0.475	–
S6	0.273	0.026	[a]	−0.164	0.075	–
S7	0.205	0.098	–	−0.194	0.035	[a]
S8	0.172	0.166	–	−0.196	0.034	[a]
S9	0.291	0.017	[a]	−0.134	0.147	–
S10	0.199	0.108	–	−0.088	0.344	–
S11	0.295	0.016	[a]	−0.092	0.320	–
S12	0.402	0.0008	[a]	−0.142	0.125	–
S13	0.267	0.029	[a]	−0.105	0.259	–
S14	0.295	0.016	[a]	−0.202	0.028	[a]
S15	0.326	0.007	[a]	−0.109	0.240	–
S16	0.274	0.026	[a]	−0.187	0.043	[a]
S17	0.210	0.089	–	−0.145	0.117	–
S18	0.337	0.005	[a]	−0.142	0.126	–
S19	0.319	0.009	[a]	−0.152	0.100	–
S20	0.334	0.006	[a]	−0.067	0.472	–
S21	0.0370	0.767	–	−0.142	0.126	–
S22	0.187	0.132	–	−0.139	0.133	–
S23	0.280	0.022	[a]	−0.113	0.222	–
S24	0.0389	0.756	–	−0.235	0.010	[a]
S25	0.269	0.028	[a]	−0.168	0.069	–
S26	0.218	0.078	–	−0.122	0.186	–
S27	0.273	0.026	[a]	−0.161	0.081	–
S28	0.276	0.024	[a]	−0.121	0.193	–
S29	0.335	0.005	[a]	−0.102	0.271	–
S30	0.135	0.278	–	0.010	0.912	–
S31	0.307	0.012	[a]	0.0351	0.707	–
S32	0.102	0.411	–	−0.162	0.08	–

[a] statistically significant Group 1: postCOVID patients without impaired pulmonary function, Group 2: postCOVID patients with impaired pulmonary function.

the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

COPOCYT. Fideicomiso 23871 Multas Electorales Convocatoria 2021–01, Development and evaluation of a screening system based on exhalation metabolomics for the detection and classification of post-COVID lung damage using automated learning methods. We thank Lorena Díaz and Berenice Balderas for their support in sample analysis and Labinova for the temporary use of equipment for the project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.talanta.2023.124299.

References

[1] J.B. Soriano, S. Murthy, J.C. Marshall, P. Relan, J.V. Diaz, A clinical case definition of post-COVID-19 condition by a Delphi consensus, the Lancet, Infectious diseases 22 (4) (2022) e102–e107.
[2] N. Zeng, Y.-M. Zhao, W. Yan, C. Li, Q.-D. Lu, L. Liu, S.-Y. Ni, H. Mei, K. Yuan, L. Shi, P. Li, T.-T. Fan, J.-L. Yuan, M.V. Vitiello, T. Kosten, A.L. Kondratiuk, H.-Q. Sun, X.-D. Tang, M.-Y. Liu, A. Lalvani, J. Shi, Y.-P. Bao, L. Lu, A systematic review and meta-analysis of long term physical and mental sequelae of COVID-19 pandemic: call for research priority and action, Mol. Psychiatr. 28 (1) (2023) 423–433.
[3] A.K. Luger, T. Sonnweber, L. Gruber, C. Schwabl, K. Cima, P. Tymoszuk, A. K. Gerstner, A. Pizzini, S. Sahanic, A. Boehm, M. Coen, C.J. Strolz, E. Wöll, G. Weiss, R. Kirchmair, G.M. Feuchtner, H. Prosch, I. Tancevski, J. Löffler-Ragg, G. Widmann, Chest CT of lung injury 1 Year after COVID-19 pneumonia: the CovILD study, Radiology 304 (2) (2022) 462–470.
[4] C. Chen, S.R. Haupert, L. Zimmermann, X. Shi, L.G. Fritsche, B. Mukherjee, Global prevalence of post-coronavirus disease 2019 (COVID-19) condition or long COVID: a meta-analysis and systematic review, J. Infect. Dis. 226 (9) (2022) 1593–1607.
[5] B. Oronsky, C. Larson, T.C. Hammond, A. Oronsky, S. Kesari, M. Lybeck, T.R. Reid, A review of persistent post-COVID syndrome (PPCS), Clin. Rev. Allergy Immunol. (2021) 1–9.
[6] A. Nalbandian, K. Sehgal, A. Gupta, M.V. Madhavan, C. McGroder, J.S. Stevens, J. R. Cook, A.S. Nordvig, D. Shalev, T.S. Sehrawat, N. Ahluwalia, B. Bikdeli, D. Dietz, C. Der-Nigoghossian, N. Liyanage-Don, G.F. Rosner, E.J. Bernstein, S. Mohan, A. A. Beckley, D.S. Seres, T.K. Choueiri, N. Uriel, J.C. Ausiello, D. Accili, D. E. Freedberg, M. Baldwin, A. Schwartz, D. Brodie, C.K. Garcia, M.S.V. Elkind, J. M. Connors, J.P. Bilezikian, D.W. Landry, E.Y. Wan, Post-acute COVID-19 syndrome, Nat. Med. 27 (4) (2021) 601–615.
[7] F. Wang, R.M. Kream, G.B. Stefano, Long-Term respiratory and neurological sequelae of COVID-19, Med. Sci. Mon. Int. Med. J. Exp. Clin. Res. 26 (2020), e928996.
[8] A.A. Rabaan, S.H. Al-Ahmed, J. Muhammad, A. Khan, A.A. Sule, R. Tirupathi, A. A. Mutair, S. Alhumaid, A. Al-Omari, M. Dhawan, R. Tiwari, K. Sharun, R. K. Mohapatra, S. Mitra, M. Bilal, S.A. Alyami, T.B. Emran, M.A. Moni, K. Dhama, Role of inflammatory cytokines in COVID-19 patients: a review on molecular mechanisms, immune functions, immunopathology and immunomodulatory drugs to counter cytokine storm, Vaccines 9 (5) (2021).
[9] D.K. Rai, S. Kumar, N. Sahay, Post-COVID-19 pulmonary fibrosis: a case series and review of literature, J. Fam. Med. Prim. Care 10 (5) (2021) 2028–2031.
[10] M. Xiang, H. Jing, C. Wang, V.A. Novakovic, J. Shi, Persistent lung injury and prothrombotic state in long COVID, Front. Immunol. 13 (2022), 862522.
[11] G. Ramírez-Martínez, L.A. Jiménez-Álvarez, A. Cruz-Lagunas, S. Ignacio-Cortés, I. A. Gómez-García, T.S. Rodríguez-Reyna, J.A. Choreño-Parra, J. Zúñiga, Possible role of matrix metalloproteinases and TGF-β in COVID-19 severity and sequelae, J. Interferon Cytokine Res. : the official journal of the International Society for Interferon and Cytokine Research 42 (8) (2022) 352–368.
[12] B.J. Hama Amin, F.H. Kakamad, G.S. Ahmed, S.F. Ahmed, B.A. Abdulla, S. H. Mohammed, T.M. Mikael, R.Q. Salih, R.K. Ali, A.M. Salh, D.A. Hussein, Post COVID-19 pulmonary fibrosis; a meta-analysis study, Annals of medicine and surgery 77 (2022) (2012), 103590.
[13] X. Zhao, J.M. Nicholls, Y.G. Chen, Severe acute respiratory syndrome-associated coronavirus nucleocapsid protein interacts with Smad3 and modulates transforming growth factor-beta signaling, J. Biol. Chem. 283 (6) (2008) 3272–3280.
[14] A. Fortini, A. Rosso, P. Cecchini, A. Torrigiani, A. Lo Forte, P. Carrai, C. Alessi, F. Fabbrizzi, E. Lovicu, S. Sbaragli, A. Faraone, One-year evolution of DLCO

changes and respiratory symptoms in patients with post COVID-19 respiratory syndrome, Infection 50 (2) (2022) 513–517.
[15] D.C. Bretas, A.S. Leite, E.V. Mancuzo, T.A. Prata, B.H. Andrade, J. Oliveira, A. P. Batista, G.L.L. Machado-Coelho, V.M. Augusto, C.C. Marinho, Lung function six months after severe COVID-19: does time, in fact, heal all wounds? Braz. J. Infect. Dis. : an official publication of the Brazilian Society of Infectious Diseases 26 (3) (2022), 102352.
[16] L. Díaz de León-Martínez, M. Rodríguez-Aguilar, P. Gorocica-Rosete, C. A. Domínguez-Reyes, V. Martínez-Bustos, J.A. Tenorio-Torres, O. Ornelas-Rebolledo, J.A. Cruz-Ramos, B. Balderas-Segura, R. Flores-Ramírez, Identification of profiles of volatile organic compounds in exhaled breath by means of an electronic nose as a proposal for a screening method for breast cancer: a case-control study, J. Breath Res. 14 (4) (2020), 046009.
[17] K.B. Méndez-Rodríguez, N. Figueroa-Vega, C.A. Ilizaliturri-Hernandez, M. Cardona-Alvarado, J.A. Borjas-García, C. Kornhauser, J.M. Malacara, R. Flores-Ramírez, F.J. Pérez-Vázquez, Identification of metabolic markers in patients with type 2 Diabetes by Ultrafast gas chromatography coupled to electronic nose, A pilot study, Biomedical chromatography : BMC (Biomed. Chromatogr.) 34 (12) (2020), e4956.
[18] M. Rodríguez-Aguilar, L. Díaz de León-Martínez, P. Gorocica-Rosete, R.P. Padilla, I. Thirión-Romero, O. Ornelas-Rebolledo, R. Flores-Ramírez, Identification of breath-prints for the COPD detection associated with smoking and household air pollution by electronic nose, Respir. Med. 163 (2020), 105901.
[19] M. Rodríguez-Aguilar, L. Díaz de León-Martínez, P. Gorocica-Rosete, R. Pérez-Padilla, C.A. Domínguez-Reyes, J.A. Tenorio-Torres, O. Ornelas-Rebolledo, G. Mehta, B.N. Zamora-Mendoza, R. Flores-Ramírez, Application of chemoresistive gas sensors and chemometric analysis to differentiate the fingerprints of global volatile organic compounds from diseases. Preliminary results of COPD, lung cancer and breast cancer, Clinica chimica acta, international journal of clinical chemistry 518 (2021) 83–92.
[20] M. Rodríguez-Aguilar, S. Ramírez-García, C. Ilizaliturri-Hernández, A. Gómez-Gómez, E. Van-Brussel, F. Díaz-Barriga, S. Medellín-Garibay, R. Flores-Ramírez, Ultrafast gas chromatography coupled to electronic nose to identify volatile biomarkers in exhaled breath from chronic obstructive pulmonary disease patients: a pilot study, Biomed. Chromatogr. : BMC (Biomed. Chromatogr.) 33 (12) (2019), e4684.
[21] K.B.M. Rodríguez, L.M.R. Gómez, L.C. Yáñez, R.F. Ramírez, O. Ornelas-Rebolledo, J.A. Borjas-García, F. Pérez-Vázquez, M.R. Aguilar, Application of the electronic nose in predicting preeclampsia in high-risk pregnancies, Pilot Study, Archives of medical research 52 (5) (2021) 561–568.
[22] M. Phillips, J. Herrera, S. Krishnan, M. Zain, J. Greenberg, R.N. Cataneo, Variation in volatile organic compounds in the breath of normal humans, J. Chromatogr. B Biomed. Sci. Appl. 729 (1–2) (1999) 75–88.
[23] M. Rodríguez-Aguilar, L. Díaz de León-Martínez, B.N. Zamora-Mendoza, A. Comas-García, S.E. Guerra Palomares, C.A. García-Sepúlveda, L.E. Alcántara-Quintana, F. Díaz-Barriga, R. Flores-Ramírez, Comparative analysis of chemical breath-prints through olfactory technology for the discrimination between SARS-CoV-2 infected patients and controls, Clinica chimica acta, international journal of clinical chemistry 519 (2021) 126–132.
[24] B.N. Zamora-Mendoza, L. Díaz de León-Martínez, M. Rodríguez-Aguilar, B. Mizaikoff, R. Flores-Ramírez, Chemometric analysis of the global pattern of volatile organic compounds in the exhaled breath of patients with COVID-19, post-COVID and healthy subjects. Proof of concept for post-COVID assessment, Talanta 236 (2022), 122832.
[25] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, Routledge, New York, 2013.
[26] CDC, NHANES III, Reference Values, Spirometry Training Program, 2011. https://www.cdc.gov/niosh/topics/spirometry/nhanes.html. (Accessed 15 July 2021). Accessed.
[27] Z. Ye, G. Auner, Principal component analysis approach for biomedical sample identification, in: IEEE International Conference on Systems, Man and Cybernetics, IEEE Cat. No. 04CH37583), 2004, pp. 1348–1353. IEEE, 2004.
[28] J. Xia, N. Psychogios, N. Young, D.S. Wishart, MetaboAnalyst: a web server for metabolomic data analysis and interpretation, Nucleic Acids Res. 37 (suppl_2) (2009) W652–W660.
[29] K. Snitz, M. Andelman-Gur, L. Pinchover, R. Weissgross, A. Weissbrod, E. Mishor, R. Zoller, V. Linetsky, A. Medhanie, S. Shushan, E. Jaffe, N. Sobel, Proof of concept for real-time detection of SARS CoV-2 infection with an electronic nose, PLoS One 16 (6) (2021), e0252121.
[30] S.N. Hidayat, T. Julian, A.B. Dharmawan, M. Puspita, L. Chandra, A. Rohman, M. Julia, A. Rianjanu, D.K. Nurputra, K. Triyana, H.S. Wasisto, Hybrid learning method based on feature clustering and scoring for enhanced COVID-19 breath analysis by an electronic nose, Artif. Intell. Med. 129 (2022), 102323.
[31] N. Phukkaphan, T. Eamsa-Ard, W. Aunsa-Ard, C. Khunarak, T. Nitivanichsakul, B. Roongpuvapaht, T. Kerdcharoen, Detection of COVID-19 Infection Based on Electronic Nose Technique: Preliminary Study, 2022 International Electrical Engineering Congress, iEECON), 2022, pp. 1–4.
[32] A.D. Subali, L. Wiyono, M. Yusuf, M.F.A. Zaky, The potential of volatile organic compounds-based breath analysis for COVID-19 screening: a systematic review & meta-analysis, Diagn. Microbiol. Infect. Dis. 102 (2) (2022), 115589.
[33] C.N.S. Allen, M. Santerre, S.P. Arjona, L.J. Ghaleb, M. Herzi, M.D. Llewellyn, N. Shcherbik, B.E. Sawaya, SARS-CoV-2 causes lung inflammation through metabolic reprogramming and RAGE, Viruses 14 (5) (2022).
[34] V. Visco, C. Vitale, A. Rispoli, C. Izzo, N. Virtuoso, G.J. Ferruzzi, M. Santopietro, A. Melfi, M.R. Rusciano, A. Maglio, P. Di Pietro, A. Carrizzo, G. Galasso, A. Vatrella, C. Vecchione, M. Ciccarelli, Post-COVID-19 syndrome: involvement and interactions between respiratory, cardiovascular and nervous systems, J. Clin. Med. 11 (3) (2022) 524.
[35] A. Vianello, G. Guarnieri, F. Braccioni, S. Lococo, B. Molena, A. Cecchetto, C. Giraudo, L. Bertagna De Marchi, M. Caminati, G. Senna, The pathogenesis, epidemiology and biomarkers of susceptibility of pulmonary fibrosis in COVID-19 survivors, Clin. Chem. Lab. Med. 60 (3) (2022) 307–316.
[36] G.-Y. Cui, B.-C. Rao, Z.-H. Zeng, X.-M. Wang, T. Ren, H.-Y. Wang, H. Luo, H.-Y. Ren, C. Liu, S.-Y. Ding, J.-J. Tan, Z.-G. Liu, Y.-W. Zou, Z.-G. Ren, Z.-J. Yu, Characterization of oral and gut microbiome and plasma metabolomics in COVID-19 patients after 1-year follow-up, Military Medical Research 9 (1) (2022) 32.
[37] C. Donati, F. Cencetti, C. Bernacchioni, V. Vannuzzi, P. Bruni, Role of sphingosine 1-phosphate signalling in tissue fibrosis, Cell. Signal. 78 (2021), 109861.
[38] S.E. Tanni, A.T. Fabro, A. de Albuquerque, E.V.M. Ferreira, C.G.Y. Verrastro, M.V. Y. Sawamura, S.M. Ribeiro, B.G. Baldi, Pulmonary fibrosis secondary to COVID-19: a narrative review, Expet Rev. Respir. Med. 15 (6) (2021) 791–803.
[39] G. Currie, E. Rohren, Social asymmetry, artificial intelligence and the medical imaging landscape, Semin. Nucl. Med. 52 (4) (2022) 498–503.
[40] E.A. Akl, I. Blažić, S. Yaacoub, G. Frija, R. Chou, J.A. Appiah, M. Fatehi, N. Flor, E. Hitti, H. Jafri, Z.-Y. Jin, H.U. Kauczor, M. Kawooya, E.A. Kazerooni, J.P. Ko, R. Mahfouz, V. Muglia, R. Nyabanda, M. Sanchez, P.B. Shete, M. Ulla, C. Zheng, E. van Deventer, M.d.R. Perez, Use of chest imaging in the diagnosis and management of COVID-19: a WHO rapid advice guide, Radiology 298 (2) (2020) E63–E69.
[41] R.N. V, A.K. Mohapatra, K.U. V, J. Lukose, V.B. Kartha, S. Chidangil, Post-COVID syndrome screening through breath analysis using electronic nose technology, Anal. Bioanal. Chem. 414 (12) (2022) 3617–3624.
[42] R. Torres-Castro, L. Vasconcello-Castillo, X. Alsina-Restoy, L. Solis-Navarro, F. Burgos, H. Puppo, J. Vilaró, Respiratory function in patients post-infection by COVID-19: a systematic review and meta-analysis, Pulmonology 27 (4) (2021) 328–337.