Supplementary Information (SI) for Sensors & Diagnostics.

SNAPSHOT OF LONG COVID IN YOUNG ADULTS: ELECTRONIC NOSE AS AN ALTERNATIVE FOR FAST SCREENING

Table 1 Supplementary Material. Clinical characteristics of COVID-19 disease.

Parameters	All	Long COVID	Healthy	p
N (%)	78	23 (29.5%)	55 (70.5%)	
Headache, n (%)	62 (79.5%)	21 (91.3%)	41 (74.5%)	0.128
Myalgias, n (%)	54 (69.2%)	18 (78.3%)	36 (65.5%)	0.264
Fever, n (%)	49 (62.8%)	17 (73.9%)	32 (58.2%)	0.190
Nasal discharge, n (%)	49 (62.8%)	17 (73.9%)	32 (58.2%)	0.190
Throat pain, n (%)	47 (60.3%)	17 (73.9%)	30 (54.5%)	0.111
Fatigue, n (%)	47 (60.3%)	18 (78.3%)	29 (52.7%)	**0.036**
Sneezing, n (%)	43 (55.1%)	13 (56.5%)	30 (54.5%)	0.783
Disgeusia, n (%)	39 (50.0%)	15 (65.2%)	24 (43.6%)	0.082
Anosmia, n (%)	38 (48.7%)	14 (60.9%)	24 (43.6%)	0.165
Back pain, n (%)	26 (46.2%)	11 (47.8%)	25 (45.5%)	0.848
Fever >38ºC, n (%)	34 (43.6%)	15 (65.2%)	19 (34.5%)	0.013
Dyspnea, n (%)	31 (39.7%)	16 (69.6%)	15 (27.3%)	**<0.001**
Arthalgia, n (%)	31 (39.7%)	11 (47.8%)	20 (36.4%)	0.346
Dry cough, n (%)	26 (33.3%)	9 (39.1%)	17 (30.9%)	0.482
Cough with phlegm, n (%)	25 (32.1%)	9 (39.1%)	16 (29.1%)	0.386
Conjunctivitis, n (%)	25 (32.1%)	9 (39.1%)	16 (29.1%)	0.386
Earache, n (%)	16 (20.5%)	6 (26.1%)	10 (18.2%)	0.540
Diarrhea, n (%)	16 (20.5%)	4 (17.4%)	12 (21.8%)	0.766
Nausea, n (%)	13 (16.7%)	7 (30.4%)	6 (10.9%)	**0.035**
Vomit, n (%)	5 (6.4%)	4 (17.4%)	1 (1.8%)	**0.025**
Any sequelae, n (%)	46 (59.0%)	19 (82.6%)	27 (49.1%)	**0.006**
*The chi-square and Fisher's exact test were performed.				

Table 2 Supplementary Material. Long COVID symptoms

Parameters	All	Long COVID	Healthy	p
N (%)	78	23 (29.5%)	55 (70.5%)	N.C.
Fatigue, n (%)	27 (34.6%)	8 (34.8%)	19 (34.5%)	0.984
Hair loss, n (%)	17 (21.8%)	6 (26.1%)	11 (20.0%)	0.553
Headache, n (%)	16 (20.5%)	5 (21.7%)	11 (20.0%)	1.000
Dyspnea, n (%)	16 (20.5%)	6 (26.1%)	10 (18.2%)	0.540
Anxiety, n (%)	16 (20.5%)	8 (34.8%)	8 (14.5%)	0.064
Reduced lung function, n (%)	16 (20.5%)	8 (34.8%)	8 (14.5%)	0.064
Eating disorder, n (%)	15 (19.2%)	6 (26.1%)	9 (16.4%)	0.354
Anosmia, n (%)	15 (19.2%)	8 (34.8%)	7 (12.7%)	**0.032**
Sleep disorders, n (%)	14 (17.9%)	8 (34.8%)	6 (10.9%)	**0.021**
Altered mental health, n (%)	12 (15.6%)	6 (26.1%)	6 (10.9%)	0.165
Cough, n (%)	12 (15.6%)	1 (4.5%)	11 (20.0%)	0.162
Percordial pain, n (%)	12 (15.4%)	5 (21.7%)	7 (12.7%)	0.322
Gastrointestinal symptoms, n (%)	10 (12.8%)	6 (26.1%)	4 (7.3%)	0.056
Fatigue after exercise, n (%)	10 (12.8%)	5 (21.7%)	5 (9.1%)	0.149
Depression, n (%)	9 (11.5%)	3 (13.0%)	6 (10.9%)	1.000
Skin symptoms, n (%)	4 (5.1%)	3 (13.0%)	1 (1.9%)	0.074
Sputum, n (%)	4 (5.1%)	3 (13.0%)	1 (1.9%)	0.074
Bradycardia, n (%)	3 (3.8%)	3 (13.0%)	0 (0.0%)	**0.023**
*The chi-square, and Fisher's exact test were performed.				

Potential Sources of Bias

This study was observational and cross-sectional, and thus, some sources of bias cannot be completely ruled out. Selection bias may have occurred since participants were recruited from a single medical school, which may not fully represent the general young adult population. Information bias and recall bias are possible because some data—particularly acute-phase and post-acute symptoms—relied on self-reported questionnaires completed retrospectively. To mitigate this, standardized symptom checklists and clear definitions based on the WHO Delphi consensus were used. Measurement bias was minimized through the use of calibrated spirometry equipment and a standardized e-Nose sampling protocol performed under controlled environmental conditions with duplicate measurements. Analytical bias was reduced by using cross-validation, permutation tests, and independent algorithms (PCA, PLS-DA, CAP, and Random Forest) to confirm reproducibility of results. Despite these measures, the cross-sectional nature of the study precludes establishing causal relationships, and future longitudinal studies are needed to confirm temporal associations and long-term outcomes.

General Information about Metal-Oxide Sensors in the e-Nose

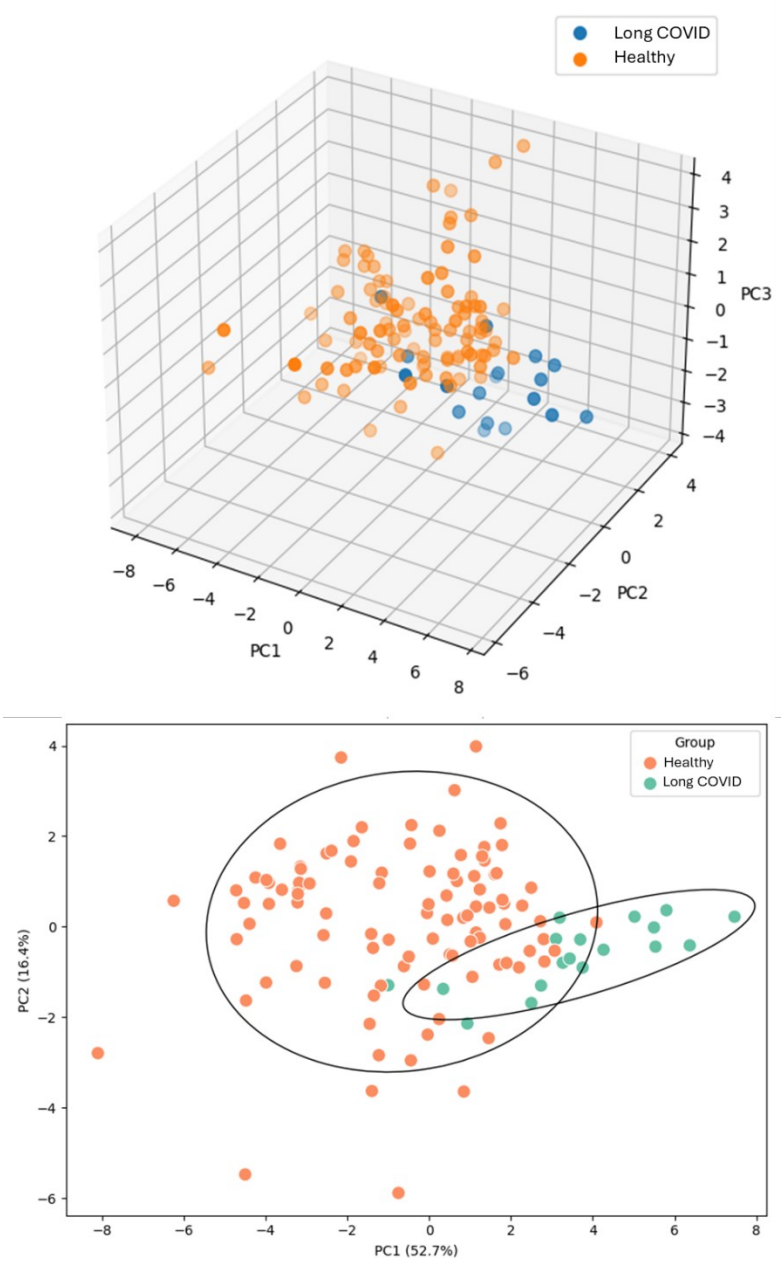
The electronic nose employed in this study utilizes an array of commercially available metal-oxide semiconductor (MOx) gas sensors based on established sensing materials commonly used in breath analysis applications. Although specific sensor-to-material assignments cannot be disclosed due to intellectual property constraints, the sensing elements within the array are representative of widely used MOx compositions reported in the literature.

In general, MOx sensors are fabricated using semiconducting metal oxides such as tin oxide (SnO_2), zinc oxide (ZnO), tungsten oxide (WO_3), and mixed-metal oxide formulations. These materials may be modified with catalytic additives or dopants (e.g., noble metals or transition metal oxides) to enhance sensitivity, response kinetics, and selectivity toward different classes of volatile organic compounds, including alcohols, aldehydes, ketones, hydrocarbons, and nitrogen-containing compounds.

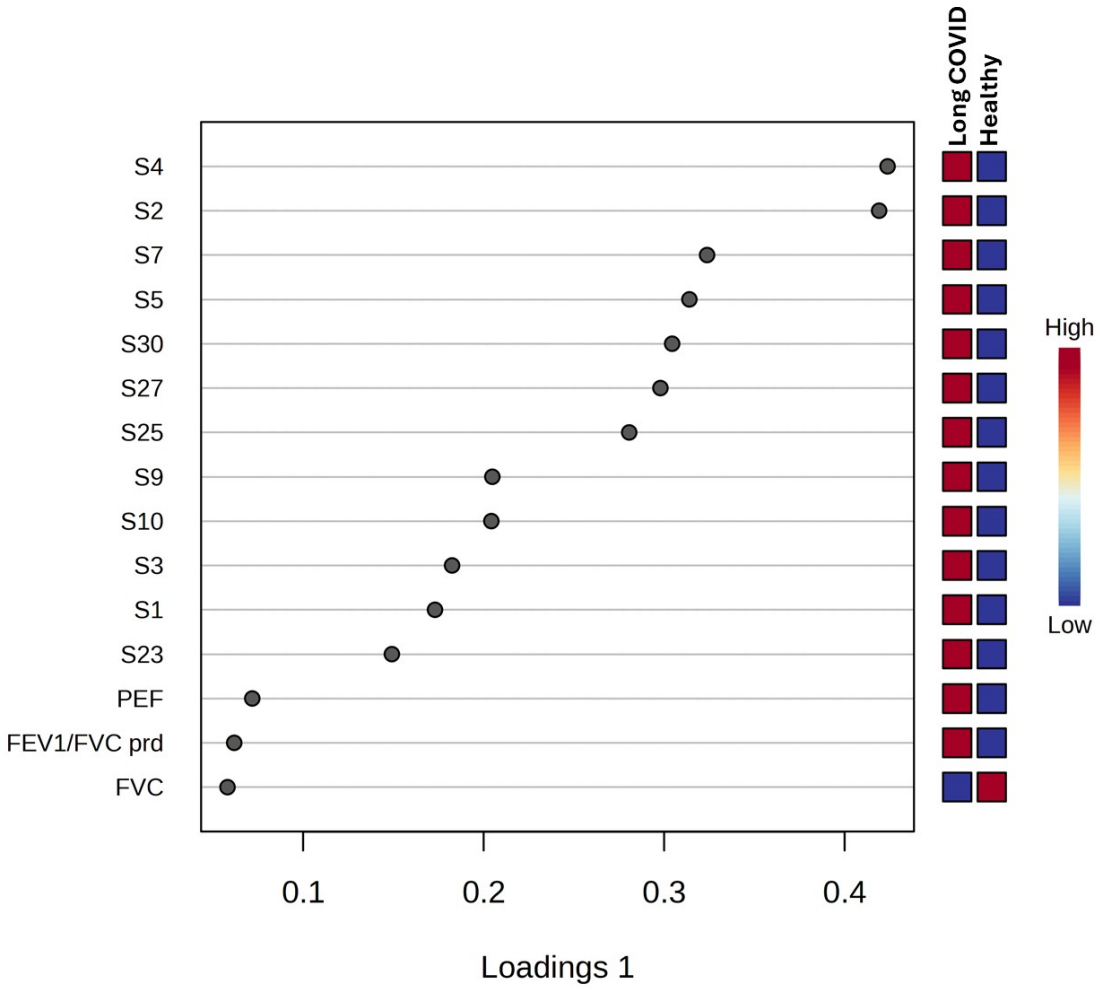
Gas detection is based on surface redox reactions between adsorbed oxygen species and target analytes, which induce changes in the electrical resistance of the sensing layer. To promote these surface reactions, MOx sensors operate at elevated temperatures, typically in the range of approximately 200–400 °C, generated by integrated micro-heaters within each sensor package. This operating temperature is independent of the external temperature of the sensing chamber.

Electrical operation generally involves low-voltage excitation and monitoring of resistance or conductance changes as the analytical signal. Some sensors in the array provide raw analog resistive outputs, while others integrate on-chip signal conditioning, analog-to-digital conversion, and environmental compensation. These are referred to as “analog” and “digital” sensors, respectively; however, it should be emphasized that the underlying gas-sensing mechanism is inherently analog for all MOx sensors.

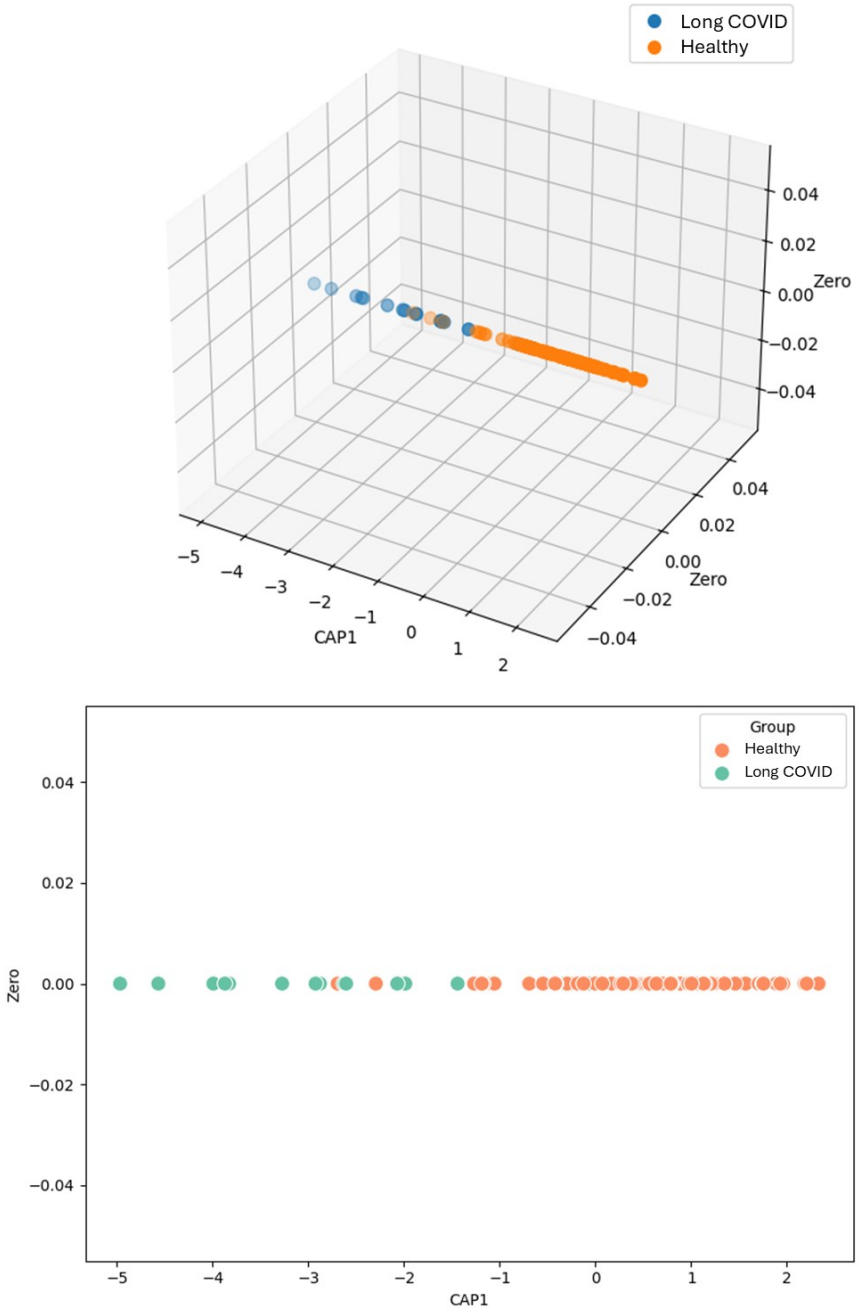
The combination of multiple MOx materials with partially overlapping sensitivities enables the generation of composite response patterns rather than compound-specific detection, which is the basis of electronic nose operation.



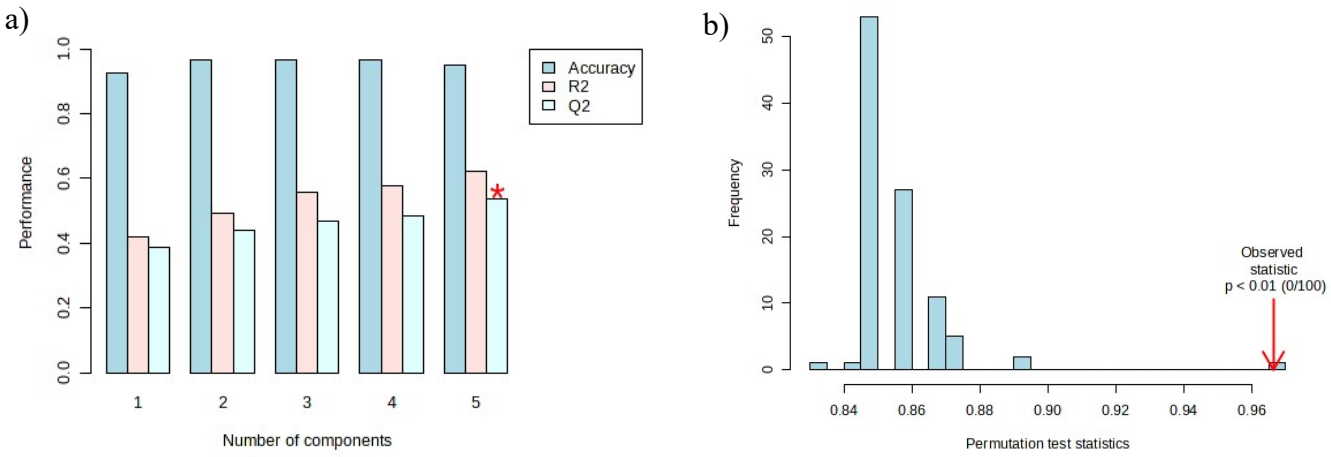
Supplementary Figure 1. Principal Component Analysis (PCA) score plot of exhaled breath sensor responses for the study groups. Each point represents one participant. Long COVID participants are shown in blue and green tones and healthy participants in orange. PCA is an unsupervised method used here for exploratory analysis and visualization of variance structure; group labels were not used to build the model.



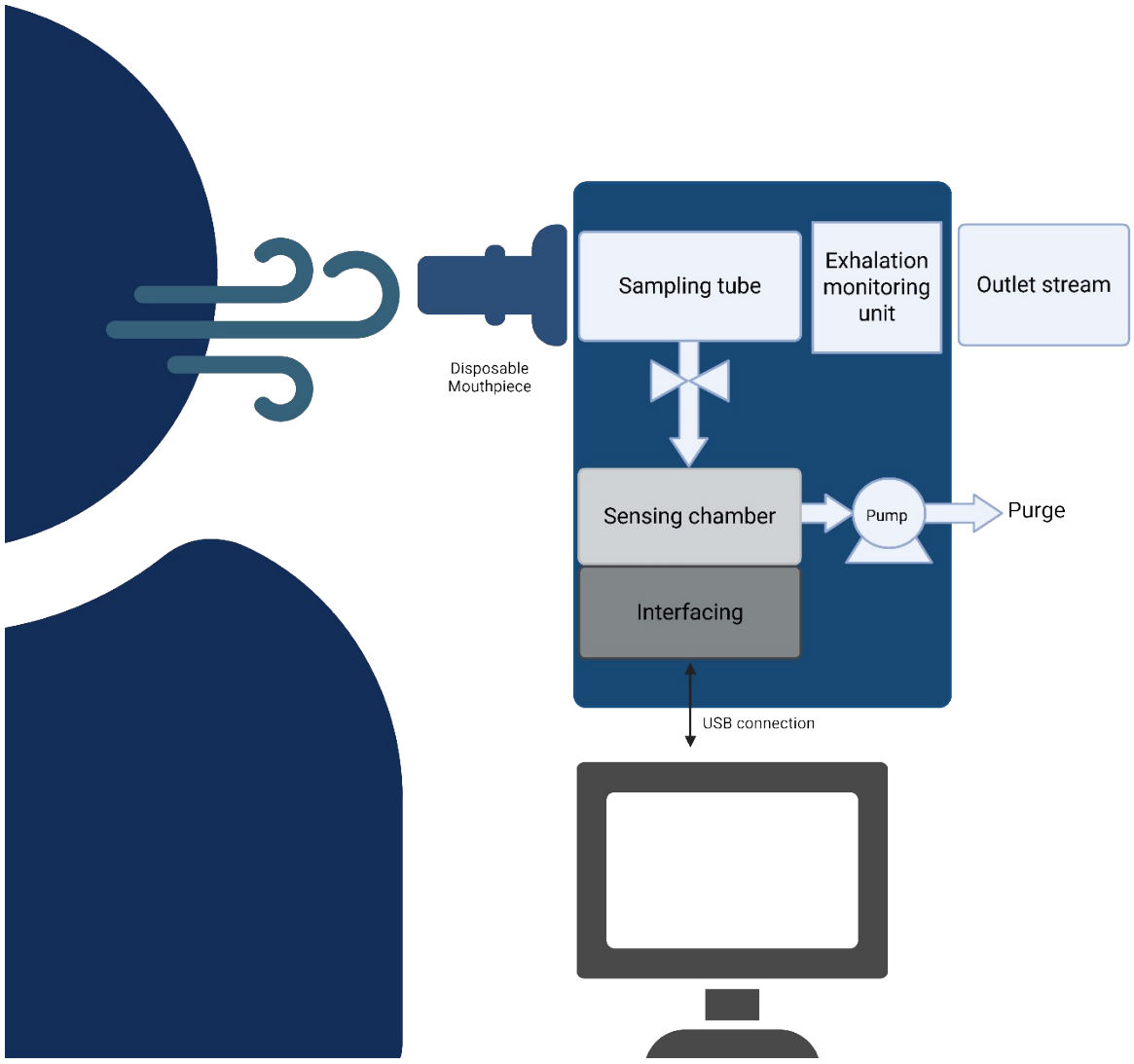
Supplementary Figure 2. Variable Importance in Projection (VIP) scores derived from supervised classification models (PLS-DA and Random Forest), showing the 15 most informative sensors and parameters contributing to separation between Long COVID and healthy participants. The x-axis represents the relative importance of each variable for class discrimination. Color coding indicates the direction of association within the discriminant model rather than exclusive importance for a single class. Because the classification task is binary, variables contributing to separation may appear associated with one class as a consequence of model orientation, while classification inherently implies non-membership in the other class. Variable importance reflects multivariate contributions and should not be interpreted as univariate effect size or mean group difference.



Supplementary Figure 3. Canonical Analysis of Principal Coordinates (CAP) score plot based on exhaled breath sensor responses. CAP is a constrained ordination method that maximizes separation between predefined groups (Long COVID vs healthy). Points represent individual participants, with LC shown in blue and green tones and healthy participants in orange. Separation along the first canonical axis reflects group-structured variation in the multivariate dataset.



Supplementary Figure 4. Partial Least Squares–Discriminant Analysis (PLS-DA) model validation. (a) Cross-validation results showing model performance across increasing numbers of latent components. R^2 represents the proportion of variance explained by the model, and Q^2 represents predictive ability estimated by cross-validation. (b) Permutation test assessing the statistical significance of the observed class separation relative to randomly permuted class labels.



Supplementary Figure 5. Illustrative scheme of the e-Nose designed for this study, as well as the representation of the sampling procedure. Scheme based on the publication of Jaeschke et al., 2021 [1]

1. C. Jaeschke, M. Padilla, J. Glöckler, I. Polaka, M. Leja, V. Veliks, J. Mitrovics, M. Leja and B. Mizaikoff, *Molecules*, 2021, **26**, 3776.