



Early View

Original research article

The science behind the nose: correlating volatile organic compound characterization with canine biodetection of COVID-19

Marthe Charles, Dorota Ruszkiewicz, Eric Eckbo, Elizabeth Bryce, Teresa Zurberg, Austin Meister, Lâle Aksu, Leonardo Navas, Renelle Myers

Please cite this article as: Charles M, Ruszkiewicz D, Eckbo E, *et al.* The science behind the nose: correlating volatile organic compound characterization with canine biodetection of COVID-19. *ERJ Open Res* 2024; in press (<https://doi.org/10.1183/23120541.00007-2024>).

This manuscript has recently been accepted for publication in the *ERJ Open Research*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJOR online.

Copyright ©The authors 2024. This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact permissions@ersnet.org

Title: The science behind the nose: correlating volatile organic compound characterization with canine biodetection of COVID-19.

Authors: Marthe Charles MD, FRCPC^{1,3}, Dorota Ruszkiewicz, PhD^{2,3}, Eric Eckbo MD, FRCPC^{1,3}, Elizabeth Bryce MD, FRCPC^{1,4}, Teresa Zurberg, RCT⁴, Austin Meister, BA³, Lâle Aksu, BA⁴, Leonardo Navas, BA⁴, Renelle Myers MD, FRCPC^{2,3}

1. Division of Medical Microbiology, Vancouver Coastal Health, Vancouver, British Columbia, Canada

2. British Columbia Cancer Research Institute, Vancouver, British Columbia, Canada

3. University of British Columbia, Faculty of Medicine, Vancouver, British Columbia, Canada

4. Quality and Patient Safety, Vancouver Coastal Health, Vancouver, British Columbia, Canada

Keywords: Volatile Organic Compounds, Canine Biodetection, COVID-19, GC-MS,

Take home message: As the pandemic numbers are receding, the requirement for fast, mobile mass screening for respiratory virus remain of interest. This study is the first to establish that training aids by Getxent® retains VOC content that can be measured and characterized allowing to distinguish positive from negative cases of COVID-19. It also confirms a correlation between the trained dogs' alert and the identified VOCs. Further studies are required to understand how the dogs can accurately identify positive cases of COVID-19. This study could be the basis for quality assurance in aid production for canines' scent detection training; not only for COVID-19 but other medical target of interest.

Dr. Marthe Charles
Division Head, Medical Microbiology and Infection Prevention
Vancouver General Hospital
JPN 1110 - 899 West 12th Avenue
Vancouver, British Columbia V5Z 1M9

Phone: 604-875-4547
Fax: 604-875-4359

ABSTRACT

Background: The SARS-CoV-2 pandemic stimulated the advancement and research in the field of canine scent detection of COVID-19 and volatile organic compound (VOC) breath sampling. It remains unclear which VOCs are associated with positive canine alerts. This study aimed to confirm that the training aids used for COVID-19 canine scent detection were indeed releasing discriminant COVID-19 VOCs detectable and identifiable by gas chromatography (GC-MS).

Methods: Inexperienced dogs (two Labradors and one English Springer Spaniel) were trained over 19 weeks to discriminate between COVID-19 infected and uninfected individuals and then independently validated. Getxent® tubes, impregnated with the odours from clinical gargle samples, used during the canines' maintenance training process were also analysed using GC-MS.

Results: Three dogs were successfully trained to detect COVID-19. A PCA-X model was created and confirmed the ability to discriminate between VOCs from positive and negative COVID-19 Getxent® tubes with a sensitivity of 78% and a specificity of 77%. Two VOCs were found to be very predictive of positive COVID-19 cases. When comparing the dogs with GC-MS, F1 and MCC correlation scores of 0.69 and 0.37, were observed respectively, demonstrating good concordance between the two methods.

Interpretation: This study provides analytical confirmation that canine training aids can be safely and reliably produced with good discrimination between positive samples and negative controls. It is also a further step towards better understanding of canine odour discrimination of COVID-19 as the scent of interest and defining what VOC elements the canines interpret as "essential".

Introduction

Throughout the COVID-19 pandemic, the world relied on microbiology laboratories to provide accurate and timely test results as part of the global public health management strategy [1]. Laboratory diagnosis was initially based on expensive, static and laborious molecular testing requiring the establishment of collection centres and the recruitment of a highly specialised laboratory workforce. Subsequently, various iterations of rapid COVID-19 antigen tests (RAT) were released on the market [2]. These point-of-care tests were useful but were hampered by a lower sensitivity of detection and subject to the various waves of variants [3].

During this tumultuous period, investigators assessed the potential of other means of rapid and sensitive point of care testing for COVID -19 detection. Biodetection by dogs is mobile and has the advantage to quickly screen large crowds in settings such as sporting events, concerts, airports and sites of congregate living [4]. Internationally, canine scent detection teams independently trained dogs for COVID-19 biodetection and demonstrated that canines could successfully detect the virus from several sources including sweat (e.g. worn garments), breath and aerosols (e.g. used masks) and body fluids (e.g. urine, saliva, sputum) samples [5-7]. The Vancouver Coastal Health (VCH) Canines for Care team trained three inexperienced dogs to detect COVID-19 from facemask, sweat (Getxent® tubes) and gargle (Getxent® tubes) [8]. To this date, there are no local or international standards for the training of canine scent detection for human diseases.[9] Concurrently, breath testing for the detection of COVID-19 was also evaluated by several groups using different modalities for volatile organic compound comparison [10-14], with gas chromatography-mass spectrometry (GC-MS), remaining the gold standard for breath analysis. Several breath studies were able to demonstrate about two dozen key molecular features via GC-MS that separated healthy controls, unwell individuals, and COVID-19 positive patients with reasonable sensitivity and specificity [11-15]. Furthermore, the US Food and Drug Administration recently approved an emergency use authorization for a test to diagnose COVID-19 based on exhaled volatile organic compounds.[16]

The extensive work performed at VCH using GC-MS [15, 17] for COVID-19 detection presented a unique opportunity to verify if the Getxent® tubes used for training and validation of the biodetection dogs were in fact releasing discriminant COVID-19 VOCs detectable and identifiable with GC-MS, thus objectively confirming the VOC differences between negative and positive training aids. To date, a limited number of studies have compared dog's olfaction to GC-MS [18, 19]. Moreover, no other group has attempted to compare and correlate directly the ability to detect COVID-19 cases by both GC-MS and biodetection dogs from the same substrate.

This article a) describes the ability of GC-MS to separate positive and negative COVID-19 odours obtained from clinical samples and captured on GetXent® tubes; b) compares sensitivity and specificity of the dogs with GC-MS and; c) establishes the level of correlation between the two methods using key VOCs from positive samples identified by GC-MS.

Methods

Participants: Outpatient participants were recruited at community COVID-19 testing facilities, and in-patient participants were recruited from three hospitals in the Greater Vancouver area. COVID-19 status was determined by nucleic acid amplification (PCR), and positive participants were deemed eligible if their first positive COVID-19 test occurred within the ten day period prior to collection. From the three clinical samples available, a gargle sample on a Getxent® tube was selected as the closest to breath (which has an extensive body of literature available). Volunteer healthcare workers and their household members provided negative controls; saline gargle samples from these individuals were negative for all targets when tested by the BioFire® RP2.1 (bioMerieux, Lyon, France) panel that is capable of detecting COVID-19 as well as 21 other respiratory pathogens.

Ethics: The canine scent detection evaluation was deemed a quality improvement project by our bureau of ethics and the work on GC-MS breath analysis was registered under the University of British Columbia Clinical Research Ethics Office (H20-01234).

Odour Preparation: Positive saline gargle specimens were identified from clinical samples submitted to the laboratory that had a cycle threshold (Ct) value less than 27 (a strong positive test) by real-time PCR. Using sterile technique, an aliquot of the sample was transferred to a sterile plastic container and placed in a glass jar over which a Getxent® (Strasbourg, France) odour collection tube was suspended. The jar was sealed for 24 hours at room temperature, after which the tube was removed and stored in a sterile manner. Care was taken to include gargle samples from patients of various ages and sexes, infected with either circulating B.1.1.7 (alpha) variants or strains with N501Y and E484K mutations.

Canine Training & Validation: Three dogs (an English Springer Spaniel and two Labrador Retrievers) were dedicated to COVID-19 scent detection and trained over a 19-week period using scent stands combined with a remote activated treat release as previously described

[8]. The dogs chosen for participation in this study had undergone strict selection criteria to determine their suitability for detection work. Selection criteria included work ethic (the dog stuck with solving the problem and did not give up), hunt drive (the dog's innate desire to search regardless of distractions), and biddability (the dog was willing to learn and work with their human partners). Hunting breeds were utilized due to their genetic propensity to search. Formal evaluation/validation was then performed by a third-party professional canine scent detection handler using a double blind, randomised format with a fresh set of odours that had not been frozen nor used for training. Results were recorded and analyzed using Excel 2016 (Microsoft) and the categorical agreement, analytic sensitivity and specificity was calculated based on comparison to the real-time PCR results and using a COVID-19 population prevalence of 10%. Confidence intervals were calculated using "exact" Clopper-Pearson's confidence intervals. Inter-rater reliability (Kappa agreement) was calculated using GraphPad Prism ® (Boston, Massachusetts, USA)

Canine Maintenance Training: Between February 19th, 2021 and October 18th, 2021, the canine detection specialist (CDS) team continued to run five to ten scent stands weekly to preserve the accuracy and interest of the dogs. The CDS team used 20 COVID-19 positive tube and 20 negative for COVID-19 tubes. These odours were not previously used during the pre-validation training nor during the validation. Maintenance training sessions were video recorded and logged into an Excel 2016 sheet.

Sample preparation for VOC Analysis via GC-TOF-MS: A duplicate of the aforementioned Getxent® tubes used for maintenance training, 20 Getxent® tubes exposed to COVID-19, 20 negative COVID-19 gargle specimens, as well as six blank tubes were submitted for GC-TOF-MS VOC analysis. The Getxent® tubes were received in 15 cm³ sterile glass containers. After calibration and method development, 18 positive tubes and 17 negative Getxent® tubes were available for analysis. Subsequently, the tubes were placed in 20cm³ headspace vials and spiked with one µl of internal standard, consisting of toluene-D8 (CAS no: 2037-26-5) and

chloroform (CAS no:865-49-6) at a contraction level of 20µg/ml and 15µg/ml, respectively diluted in methanol (MEOH). Samples were then heated to 70°C and agitated for 20 minutes using a HS-Centri (Markes International) autosampler before one cm³ of headspace was injected onto a focusing cold-trap. This was repeated four times for each sample. To avoid possible carry-over, all samples were run in sandwich mode with two high-temperature 10-minute-long system blanks in-between. A VOC check standard consisting of 52 compounds (Sigma Aldrich, 4357-U) at 10 µg mL⁻¹ in MeOH: H₂O (95:5) was run in between every five samples. A total of 20 out of 52 VOCs were used to monitor system performance.

Instrument Analysis: Samples were then analysed using the HS-Centri system coupled to an Agilent 7890A (Agilent, UK) gas chromatogram with a Bench ToF mass spectrometer (Markes International) with electron ionisation. A DB-5MS capillary column 60 m x 0.25 mm x 0.25 µm was used for the chromatographic separation. Full method details can be found in Supplementary Tables S1 and S2 and S3.

Data Processing and Analysis: Samples were first processed using a Dynamic Baseline Correction (DBC) algorithm (TOF-DS Markes International LTD), followed by deconvolution and integration used to produce a peak table with 500 and 800 VOC features per sample recovered. From 35 tubes available, MATLAB was used to produce a sample matrix with 647 tentatively assigned VOCs with matching identification from the National Institute of Standards and Technology (NIST) library, with a match factor >700. Further features were discriminated using the following conditions: 1) the VOC was present in at least 30% of gargle samples and 2) the component was outside an average ± 2* standard deviation of the blank tubes (95% confidence). The VOCs meeting both conditions were subjected to further multivariate analysis using SIMCA® (Umetrics, UK) software. The data was log₁₀ transformed and Pareto scaled before orthogonal partial least squares-discriminant analysis (OPLS-DA) was performed to remove background and select peaks with the highest degree of discrimination using VIP and p-corr values. The remaining candidate compound's peaks were taken for principal

components analysis (PCA-X) that has an integrated seven-fold cross validation. Independent sample t-test was also performed on each individual VOC used in the PCA-X model.

The final PCA-X model was further tested using chemometric scoring. For each participant sample, a cumulative score of peak areas was produced according to the equation below:

$$\text{Score} = \sum (\text{Log(Peak Area) of features with increase in Covid - 19 pos}) - \sum (\text{Log(Peak Area) of features with decrease in Covid - 19 pos}) \quad \text{Equation 1.}$$

Where, “Log(Peak Area)” refers to peak area of each selected VOC feature (which corresponds to the levels of the VOC in the sample) which either showed increase or decrease in the Covid-19 positive samples in comparison to Covid-19 negative. The results are presented in the form of Box Whisker plots, with an Anova p-value of 0.001 (Figure 3A) when comparing scores of COVID-19 negative vs positive participants. The aforementioned scores were also used to confirm the accuracy of the PCA model separation.

GC/MS comparison to Dog’s Alert: Out of the 35 Getxent® tubes analyzed, 24 were used for training of at least two of the three dogs. These 24 tubes were compared to their corresponding GC-MS results. In addition, individual Kappa agreements were calculated from the first response of each dog with the GC-MS results. In order to further compare the dogs’ overall training responses with the GC-MS results, a “discrepancy matrix” was created using the dogs’ responses as the reference. A Pearson correlation was run for both a) the individual VOC levels (integrated Log (10)) peak areas and the dogs’ percentage of error, and b) the summarised chemometric score of the GC-MS results and dogs’ percentage error for both the COVID-19 positive and negative group.

Results

Pre-validation training

Prior to the CDS team validation for COVID-19, 51 unique gargle odours (samples) had been used for training (34 positive COVID-19 and 17 negative COVID-19 tubes). Overall, the team completed more than 100 training sessions each.

CDS validation and maintenance sessions

Validation sessions: Twelve runs over two days were conducted and 120 new samples were used for validation. Twenty-four were positive odours (8 gargle, 8 sweat and 8 breath); 21 were negative (5 gargle, 8 sweat, and 8 breath) and the remaining 75 were associated odours. Associated odours are scents related to patients' samples production, for example saline, gloves, detergent, unused Getxent® tube, etc. The sensitivity and specificity of dog A was 100% (95% CI [85.75, 100%]) and 92.7% (95%CI [85.55, 97.02%]). Using a population prevalence of 10%, the PPV would be 60.4% and the NPV 100%. The sensitivity and specificity of dog B was 100% (95% CI [85.75, 100%]) and 93.8% (95%CI [86.89, 97.67%]) respectively with a PPV of 64% and a NPV of 100%. The sensitivity and specificity of dog C was 100% (95% CI [80.49, 100%]) and 95.24% (95%CI [86.71, 99.01%]) respectively with a PPV of 70% and NPV of 100%. [8] (Table 1)

Table 1: Comparison of Dog overall performance and gargles performance for COVID-19 detection

	Dog A	Dog B	Dog C
Number of training sessions prior to validation	165	134	127
Overall sensitivity	100 (95%CI [85.75,100%])	100 (95% CI [85.75, 100%])	100 (95% CI [80.49, 100%])
Gargle sensitivity	100 (95%CI [63.06-100])	100 (95%CI [63.06-100])	100 (95%CI [54.07-100])
Overall specificity	92.7 (95%CI [85.55, 97.02%])	93.8 (95%CI [86.89, 97.67%])	95.24 (95%CI [86.71, 99.01%])
Gargle specificity	60 (95%CI [14.66-94.73])	80 (95%CI [28.36-99.49])	100 (95%CI [47.82-100])

Maintenance training sessions: For the maintenance sessions, each dog completed more than 100 sessions specifically using 40 gargle samples using (20 positive, 20 negative). Dog A failed to alert 4 times on a positive (twice on the same tube) and alerted 6 times on an expected negative case. Dog B failed to alert once on a positive case and alerted once on an expected negative case. Dog C failed to alert once and alerted 4 times on an expected negative case. (see Figure 1)

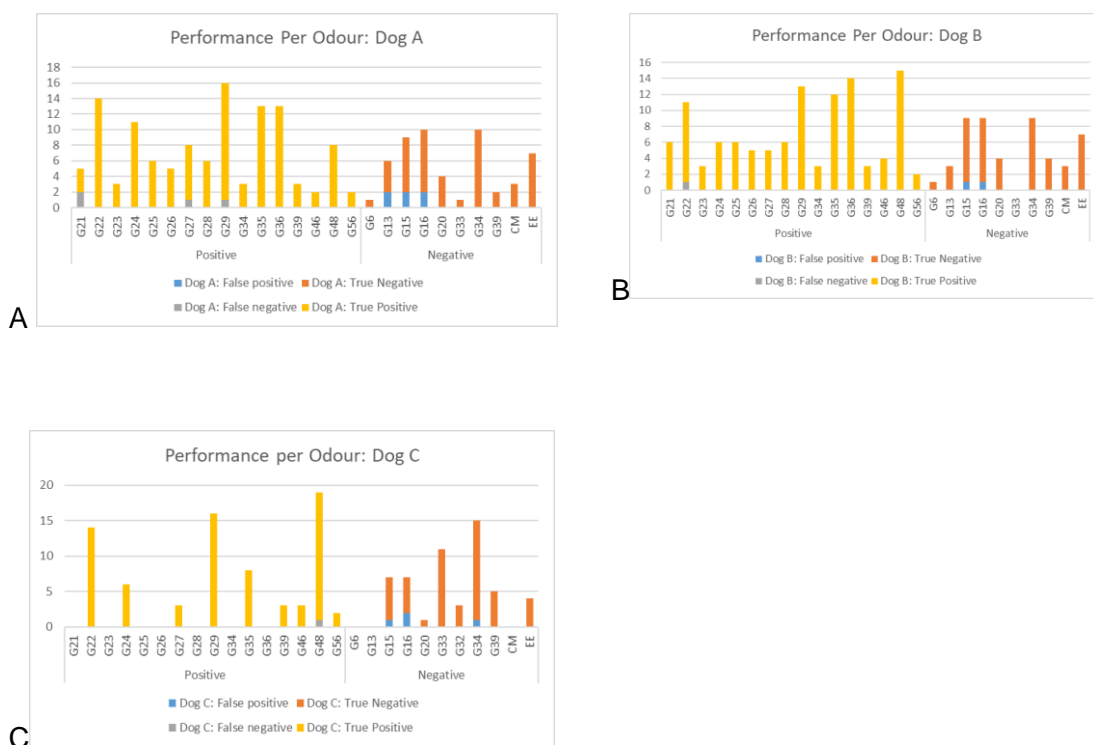


Figure 1: Comparison of performance of GetXent® tubes per number of sessions (Y-axis) per gargles with and without COVID-19 (X-axis) per Dog: A) dog A B) dog B and C) dog C.

GC-MS Analysis and Modelling

Multivariate analysis and modeling: A total of 36 Getxent® tubes were available for analysis. One gargle sample was removed from further investigation due to a low internal standard and

35 tubes (18 positive and 17 negative for COVID-19) were assessed. A two component PCA-X model was created using 9/13 features selected in OPLS (See Figure 2). There was no improvement in the model accuracy or performance at cross validation by adding further components. Hierarchical analysis based on Ward size value distinguished three main branches: positive, negative and a mixed cluster that further separated into two branches: one fully negative and one fully positive (See Figure 3). The ROC curve created on this basis provided accuracy of the classification to be 77.1% with a fitted AUROC of 0.761 (95%CI: 0.550 – 0.901) and sensitivity and specificity of 77.8% and 76.5%, respectively (see Figure 4B). When comparing chemometric scores of COVID-19 negative vs positive gargle samples, the accuracy of the PCA model separation (Figure 4A) was confirmed, as similar results were obtained with an AUROC of 0.781 (95% CI: 0.626 – 0.936).

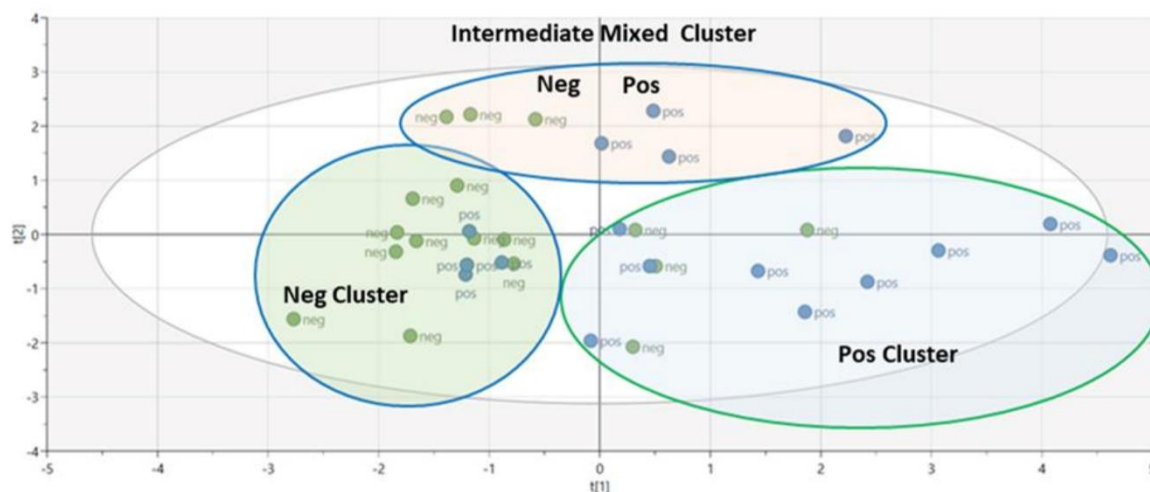


Figure 2. 2-Component PCA-X model. PCA-X model of 35 Getxent® tubes for COVID-19 positives (pos) vs negative (neg) gargle samples, based on nine features selected by OPLS-DA; Green cluster – samples assigned as COVID-19 negative; blue cluster samples assigned as COVID-19 positive, orange cluster – intermediate mixed cluster.

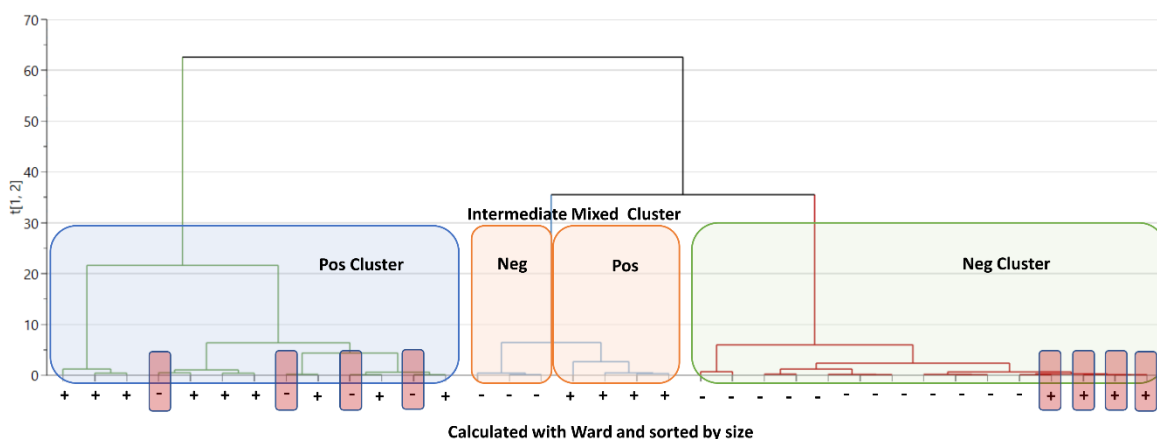


Figure 3. Dendrogram from hierarchal analysis of 35 COVID-19 positive (+) and negative (-) gargle samples, showing three distinctive classes: positive, negative and intermediate mixed cluster with two branches with positive and negative separation.

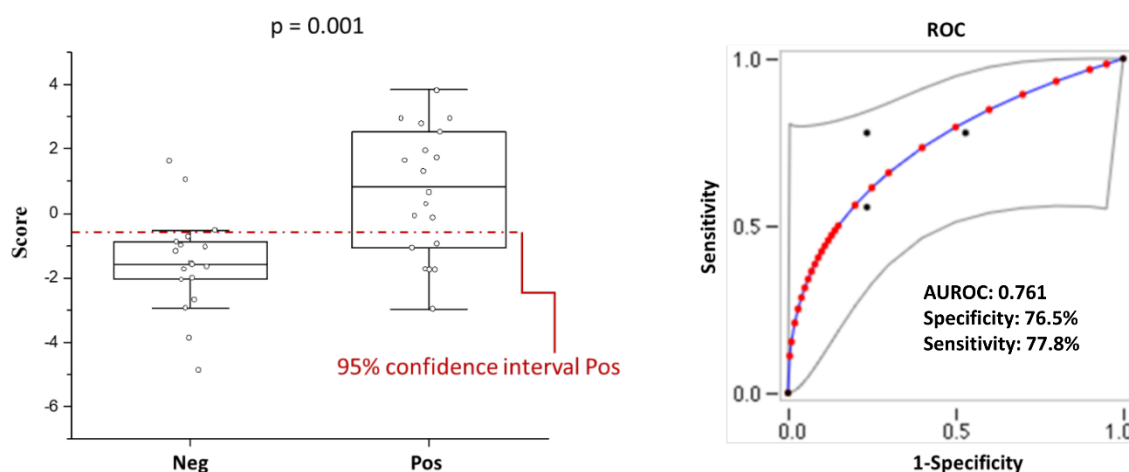


Figure 4. Chemometric score and predictive model: A) Box Whisker plot of chemometric score based on peak areas of nine features selected from multivariate analysis, according to Equation 1. B) ROC curve based on PCA dendrography model classification; with accuracy of separation between positive and negative gargle samples collected on Getxent® tubes of 77.1 % (Sensitivity of 78% and specificity of 77%).

Independent sample t-testing was also performed on each individual compound's peak used in the PCA-X model. Overall, 2/9 VOCs: benzaldehyde and one unidentified compound; were statistically significant with p-value of <0.05 (See Table 2). Differences in the peak areas are also demonstrated in the form of Box-Whisker plots presented in supplemental Figure S1.

Table 2. Features selected by OPLS-DA analysis and in PCA-X model with their CAS numbers.

Feature number / CAS Number	OPLS-DA	PCA-X	tr (min) / top 3 m/z
1 (↑)	Feature 1	Feature 1	4.21 / 41, 55, 56
2 / 590-36-3 (↑)	2_Pentanol_2_methyl	2_Pentanol_2_methyl	9.14 / 59, 45, 87
3 / 100-52-7 (↑)	#Benzaldehyde	#Benzaldehyde*	16.9 / 105, 77, 106
4 (↓)	Feature 4	Feature 4	21.5 / 119, 91, 148
5 / 565-75-3 (↑)	Pentane_2_3_4_trimethyl	Pentane_2_3_4_trimethyl	10.1 / 43, 70, 71
6 / 629-94-7 (↓)	Heneicosane	Heneicosane	46.98 / 57, 71, 43
7 (↑)	Feature 7	Feature 7	13.19 / 69, 111, 55
8 (↑)	Feature 8	Feature 8	31.5 / 83, 82, 55
9 (↓)	Feature 9	Feature 9*	46.2 / 236, 193, 167
10	Feature 10		20.56 / 56, 55, 42
11 / 64-19-7	#Acetic Acid		6.09 / 45, 43, 60
12 / 64-17-5	#Ethanol		4.45 / 45, 46, 43
13 / 592-27-8	#Heptane-2-methyl		12.6 / 43, 57, 42

*Note: Features with given names were identified using NIST library and achieved high match factor of >800 and are highlighted in bold; VOCs with # symbol, were identified to level 1 of metabolic standard recommendations. Arrow highlights direction of the concentration change in COVID-19 positive participants. * highlights the features with statistically significant ($p < 0.05$) differences in levels between SARS-CoV-2 positive and negative samples, as determined in independent sample t-test. tr – retention time; m/z – mass to charge ratio of main ions*

Performance of the model vs Dog's alert

Of the 35 Getxent® tubes sent for GCMS analysis, 24 (10 negative and 14 positive for COVID-19) had been used for at least two of the three dogs during the maintenance training. These two canines were exposed to the samples over a total number of 426 trials with an overall 95.56% (CI95 92.37-97.68) sensitivity and 96.15% (CI95 91.82-98.58) specificity. (Table 3)

Table 3. Summary of dog training on 24 Getxent® tube used in GC-MS analysis

	Dog 1	Dog 2	Dog 3	Total (%)	Mean (Q3-Q1)
Total runs on specified positive gargle samples / N = 14	99	94	71	264 (100)	20 (23)
Total runs with false positives	4	1	1	6 (2.3)	NA
Total runs on specified negative gargle samples / N = 10	55	60	47	162 (100)	16 (16)
Total runs with false negatives	6	2	4	12 (7.4)	NA

Using the discrepancy matrix, samples with a failure rate of dog alert above 15% were considered negative for the purpose of this analysis. Altogether, the model correctly classified 68% of the samples with F1 and MCC correlation scores of 0.69 and 0.37, respectively, considered to be a good concordance between the two methods (Table 4) Individually, the

Kappa agreement between the dogs and GC-MS was moderate: 0.503 for dog A, 0.510 dog B and 0.412 for dog C.

Table 4 Summary of Dogs alert versus GC-MS responses.

	Dogs alerted (pos) TP	Dogs not alerted (neg) TN
Positive by GC-MS	9 (TP)	2 (FP)
Negative by GC-MS	6 (FN)	7 (TN)
Sensitivity	0.60	$TPR = TP / (TP + FN)$
Specificity	0.78	$SPC = TN / (FP + TN)$
Precision	0.82	$PPV = TP / (TP + FP)$
Negative Predictive Value	0.54	$NPV = TN / (TN + FN)$
False Positive Rate	0.22	$FPR = FP / (FP + TN)$
False Discovery Rate	0.18	$FDR = FP / (FP + TP)$
False Negative Rate	0.40	$FNR = FN / (FN + TP)$
Accuracy	0.68	$ACC = (TP + TN) / (P + N)$
F1 Score	0.69	$F1 = 2TP / (2TP + FP + FN)$
Matthews Correlation Coefficient / MCC	0.37	$\frac{TP*TN - FP*FN}{\sqrt{((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN))}}$

TP: true positive, TN: true negative, FP: false positive and FN false negative

Correlation between Dogs and GC-MS VOCs

All three dogs were accurate (95.77% 93.40-97.48%) in detecting COVID-19 positive samples with a tendency to overcall positive samples (7.4% of total trials). Importantly, the dogs rarely missed a positive COVID-19 sample with a failure to detect at 2.3% of total trials. A Pearson correlation showed no significant correlation between individual VOCs levels and incorrect alerts by the dogs. However, there was a significant effect between the chemometric score ($p=0.04$; Pearson score -0.655) and a dog's response in the COVID-19 negative group, which better characterized the VOCs profile. The lower the chemometric score, the more "errors" the dogs made. Of note, there was no significant correlation between the dog's incorrect alert and the chemometric score within the positive COVID-19 group of Getxent® tubes.

Discussion

The ability to use odours to detect disease dates back to Hippocrates [20], who described *fetor oris* as a sign of liver disease as well as the fruity smell of urine in patients with diabetic ketoacidosis [21, 22]. Modern-day VOC tests such as nitric oxide monitoring in asthma [23] and the C-urea breath test for the diagnosis of *Helicobacter pylori* [24] illustrate the incorporation of VOC based testing into clinical practice, and demonstrate the potential of VOCs for disease identification and monitoring. Analysis of VOCs is classically done by highly sensitive analytic equipment such as GC-ToF-MS, however over the past two decades, canine biodetection is being explored for clinical practice and has been incorporated into clinical pathways such as screening facilities for environmental contamination by *Clostridiodes difficile* [25]. Canines' olfactory acuity to detect disease states and even individual microorganisms is thought to be due to an ability to detect unique VOCs or VOC profiles that are continuously emitted from body fluids such as breath, body fluids (15). The medical field has been slow in the advancement of VOC based clinical tests, because of challenges with VOC collection standardisation and analytic platform sensitivity and standardization. [26, 27]. Although the sensory capability of canines to detect discrete odours is not disputed, training dogs to specifically detect a VOC signature unique for a disease process in a standardized manner remains a challenge. Quality guidelines regarding safe handling and storage of training aids and prevention of cross contamination are still required [18, 28]. Comparing detection rates on samples between GC-MS and the dogs not only validates the biodetection training method used but identifies key molecular features that further the understanding of the components required for accurate detection of COVID-19. Ultimately, this will aid in developing universally standardized quality control processes for canine training aids for this disease, and potentially others.

Of the 35 gargle samples prepared on Getxent® tubes and analysed via GC-ToF-MS, selected VOCs were able to discriminate between COVID-19 positive and negative samples, allowing for detection of disease with 77% accuracy using the PCA-X model. The ROC curve produced

on the basis of the hierarchical analysis provided sensitivity and selectivity of 78% and 77%, respectively. The model was based on nine major VOC peaks, of which two were individually statistically significant: benzaldehyde and one unidentified molecular feature.

Many of the VOCs detected here by OPLS; acetic acid, ethanol, and benzaldehyde, have been previously identified as key markers of COVID-19 [10, 12, 29, 30], in this described case benzaldehyde made it to the final PCA. [15] Interestingly, the analyzed training aids were also associated with this VOC pattern and triggered very clear canine alerts strengthening the theory that dogs recognize specific VOCs and the patterns associated with disease.

The Pearson correlation between the dogs and the GC-MS demonstrated no significant correlation between individual VOCs and canine incorrect alerts. However, there was a significant relationship between the chemometric score ($p=0.04$; Pearson score -0.655) and the dogs' responses in the COVID-19 negative group, which suggests that there is a chemical overlap between what the dogs alert on and the panel of features detected by GC-MS.

The lower the chemometric score, the more "errors" the dogs were making. A chemometric score is related to VOC concentration levels – a lower score results from generally lower levels of VOCs in the panel and /or the expected ratio between VOCs is altered. Dogs are able to detect odours at very low levels, parts per trillion [31] vs the current GC-MS technology that can detect levels at parts per billion. Therefore, it is unlikely that the levels of VOCs were below the dogs' detection limits. The change in a combination of VOCs from COVID-19 positive samples and their presence at specific ratios may be the signal dogs perceive as most important. There was no significant correlation between a dog's "error" alert and the chemometric score within the COVID -19 positive samples, which was generally higher than in negative group. This further suggests, that the individual concentration levels of the nine key VOCs were not the most important feature for the response of the dogs towards COVID-19 samples. This may also reflect the limits of the current analytic equipment to detect a complete VOC profile, also demonstrated by fact that the dogs performed better at detecting COVID-19 in general.

Understanding that specific ratios of VOCs are more important than individual VOCs concentrations for the dog's response is an important finding. Dogs have the ability to learn and the significance or importance of a VOC profile can be increased by positive reinforcement, ie treats, toys, cheers, etc. Consequently, the observed Kappa agreement between the dogs first response to the PCA-X classification could increase overtime.

This study had some evident limitations, the number of Getxent® tubes was limited and an independent cross validation would be required to strengthen the PCA-X model findings. The discrepancy matrix was set at a failure rate of 15%, a common threshold for laboratory equipment, meanwhile from our experience the dogs failed to alert on positive sample only 2.3% of the time. This decision might have biased towards the negative samples. Specimen/sample collection was dependent upon patient presentation and thus there were different storage times for the samples. In addition, the three dogs were not all trained on the same set of odours at the same time. Not all samples could be immediately stored in air-tight containers. The latter two factors may have influenced the chemistry of the samples; the number of compounds recovered may have diminished with less stable VOCs lost prior to analysis.

CONCLUSION

This study is a step towards further standardization of the methods used to train canines in biodetection. It confirms that ultra clean tubes like Getxent® can reliably be used to safely and accurately train dogs in COVID19 detection. Furthermore, key VOC signature features were identified, seemingly involved in canines COVID-19 detection as demonstrated by a strong correlation between training aids with canine positive alerts and the VOC model generated by GC-MS. The potential to use both canine scent detection and GC-MS in a synergistic manner to further refine disease detection could have enormous impact in several areas of clinical medicine, including infectious disease and oncology practices.

Declaration of Conflict of Interest

The Authors acknowledge that part of the work presented hereby has been funded by Health Canada as part of The Safe Restart Agreement program. The Authors declare that there is no conflict of interest pertinent to this work.

References:

- [1] J. Blondeau, "Clinical microbiology laboratories and COVID-19: an interview with Joseph Blondeau," *Future Microbiol*, vol. 16, pp. 615-618, Jun 2021, doi: 10.2217/fmb-2021-0113.
- [2] J. Tsao, A. L. Kussman, C. Costales, B. A. Pinsky, G. D. Abrams, and C. E. Hwang, "Accuracy of Rapid Antigen vs Reverse Transcriptase-Polymerase Chain Reaction Testing for SARS-CoV-2 Infection in College Athletes During Prevalence of the Omicron Variant," *JAMA Netw Open*, vol. 5, no. 6, p. e2217234, Jun 1 2022, doi: 10.1001/jamanetworkopen.2022.17234.
- [3] M. Bekliz *et al.*, "Analytical Sensitivity of Eight Different SARS-CoV-2 Antigen-Detecting Rapid Tests for Omicron-BA.1 Variant," *Microbiol Spectr*, vol. 10, no. 4, p. e0085322, Aug 31 2022, doi: 10.1128/spectrum.00853-22.
- [4] N. A. Ten Hagen *et al.*, "Canine real-time detection of SARS-CoV-2 infections in the context of a mass screening event," *BMJ Global Health*, vol. 7, no. 11, p. e010276, 2022-11-01 2022, doi: 10.1136/bmjgh-2022-010276.
- [5] C. Guest *et al.*, "Using trained dogs and organic semi-conducting sensors to identify asymptomatic and mild SARS-CoV-2 infections: an observational study," *Journal of Travel Medicine*, vol. 29, no. 3, 2022-05-31 2022, doi: 10.1093/jtm/taac043.
- [6] O. Vesga *et al.*, "Highly sensitive scent-detection of COVID-19 patients in vivo by trained dogs," *PLoS one*, vol. 16, no. 9, p. e0257474, 2021.
- [7] D. Grandjean *et al.*, "Identifying SARS-COV-2 infected patients through canine olfactive detection on axillary sweat samples; study of observed sensitivities and specificities within a group of trained dogs," *PLoS One*, vol. 17, no. 2, p. e0262631, 2022, doi: 10.1371/journal.pone.0262631.
- [8] M. Charles *et al.*, "In search of COVID-19: The ability of biodetection canines to detect COVID-19 odours from clinical samples," *Official Journal of the Association of Medical Microbiology and Infectious Disease Canada*, p. e20220017, 2022, doi: 10.3138/jammi-2022-0017.
- [9] T. Dickey and H. Junqueira, "Toward the use of medical scent detection dogs for COVID-19 screening," *J Osteopath Med*, vol. 121, no. 2, pp. 141-148, Feb 1 2021, doi: 10.1515/jom-2020-0222.
- [10] W. Ibrahim *et al.*, "Diagnosis of COVID-19 by exhaled breath analysis using gas chromatography-mass spectrometry," *ERJ Open Res*, vol. 7, no. 3, Jul 2021, doi: 10.1183/23120541.00139-2021.
- [11] D. M. Ruskiewicz *et al.*, "Diagnosis of COVID-19 by analysis of breath with gas chromatography-ion mobility spectrometry - a feasibility study," *EClinicalMedicine*, vol. 29, p. 100609, Dec 2020, doi: 10.1016/j.eclinm.2020.100609.
- [12] H. Chen *et al.*, "COVID-19 screening using breath-borne volatile organic compounds," *J Breath Res*, vol. 15, no. 4, Oct 22 2021, doi: 10.1088/1752-7163/ac2e57.
- [13] A. Z. Berna *et al.*, "Reproducible Breath Metabolite Changes in Children with SARS-CoV-2 Infection," *ACS Infect Dis*, vol. 7, no. 9, pp. 2596-2603, Sep 10 2021, doi: 10.1021/acsinfecdis.1c00248.
- [14] S. Grassin-Delyle *et al.*, "Metabolomics of exhaled breath in critically ill COVID-19 patients: A pilot study," *EBioMedicine*, vol. 63, p. 103154, Jan 2021, doi: 10.1016/j.ebiom.2020.103154.
- [15] R. Myers *et al.*, "Breath testing for SARS-CoV-2 infection," *EBioMedicine*, vol. 92, p. 104584, Apr 28 2023, doi: 10.1016/j.ebiom.2023.104584.
- [16] FDA, "Coronavirus (COVID-19) Update: FDA Authorizes First COVID-19 Diagnostic Test Using Breath Samples," in *Test provides results in less than three minutes*, FDA, Ed., ed. Online, 2022.
- [17] R. Myers *et al.*, "Breath collection protocol for SARS-CoV-2 testing in an ambulatory setting," *J Breath Res*, vol. 16, no. 2, Mar 4 2022, doi: 10.1088/1752-7163/ac4e2c.
- [18] J. Mendel *et al.*, "Preliminary accuracy of COVID-19 odor detection by canines and HS-SPME-GC-MS using exhaled breath samples," (in eng), *Forensic Sci Int Synerg*, vol. 3, p. 100155, 2021, doi: 10.1016/j.fsisyn.2021.100155.

- [19] J. D. Pleil, M. A. G. Wallace, J. McCord, M. C. Madden, J. Sobus, and G. Ferguson, "How do cancer-sniffing dogs sort biological samples? Exploring case-control samples with non-targeted LC-Orbitrap, GC-MS, and immunochemistry methods," *J Breath Res*, vol. 14, no. 1, p. 016006, Nov 19 2019, doi: 10.1088/1752-7163/ab433a.
- [20] A. Amann *et al.*, "The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva," *J Breath Res*, vol. 8, no. 3, p. 034001, Sep 2014, doi: 10.1088/1752-7155/8/3/034001.
- [21] L. Laffel, "Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes," *Diabetes Metab Res Rev*, vol. 15, no. 6, pp. 412-26, Nov-Dec 1999, doi: 10.1002/(sici)1520-7560(199911/12)15:6<412::aid-dmrr72>3.0.co;2-8.
- [22] L. Pauling, A. B. Robinson, R. Teranishi, and P. Cary, "Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography," *Proc Natl Acad Sci U S A*, vol. 68, no. 10, pp. 2374-6, Oct 1971, doi: 10.1073/pnas.68.10.2374.
- [23] M. Hogman, "Extended NO analysis in health and disease," *J Breath Res*, vol. 6, no. 4, p. 047103, Dec 2012, doi: 10.1088/1752-7155/6/4/047103.
- [24] D. Y. Graham *et al.*, "Campylobacter pylori detected noninvasively by the ¹³C-urea breath test," *Lancet*, vol. 1, no. 8543, pp. 1174-7, May 23 1987, doi: 10.1016/s0140-6736(87)92145-3.
- [25] E. Bryce, T. Zurberg, M. Zurberg, S. Shajari, and D. Roscoe, "Identifying environmental reservoirs of Clostridium difficile with a scent detection dog: preliminary evaluation," *J Hosp Infect*, vol. 97, no. 2, pp. 140-145, Oct 2017, doi: 10.1016/j.jhin.2017.05.023.
- [26] T. Risby, "Critical issues for breath analysis," *Journal of Breath Research*, vol. 2, no. Editorial, 2008, doi: 10.1088/1752-7155/2/3/030302.
- [27] J. Herbig and J. Beauchamp, "Towards standardization in the analysis of breath gas volatiles," *J Breath Res*, vol. 8, no. 3, p. 037101, Sep 2014, doi: 10.1088/1752-7155/8/3/037101.
- [28] S. Meller *et al.*, "Expert considerations and consensus for using dogs to detect human SARS-CoV-2-infections," *Front Med (Lausanne)*, vol. 9, p. 1015620, 2022, doi: 10.3389/fmed.2022.1015620.
- [29] R. Remy *et al.*, "Profiling of exhaled volatile organics in the screening scenario of a COVID-19 test center," *iScience*, vol. 25, no. 10, p. 105195, Oct 21 2022, doi: 10.1016/j.isci.2022.105195.
- [30] S. Laird *et al.*, "Breath Analysis of COVID-19 Patients in a Tertiary UK Hospital by Optical Spectrometry: The E-Nose CoVal Study," *Biosensors (Basel)*, vol. 13, no. 2, Jan 20 2023, doi: 10.3390/bios13020165.
- [31] C. Guest and C. M. Otto, "Editorial: Canine Olfactory Detection," *Front Vet Sci*, vol. 7, p. 100, 2020, doi: 10.3389/fvets.2020.00100.

Supplementary

Table S1. Summary of Markes International Headspace Sampling and preconcentration/desorption Centri Conditions.

Parameter	Setting
Mode	Headspace
Agitator Temp	70 °C
Agitation Time	20.0 min
Agitation Speed	100 rpm
Agitation on time	10 s
Agitation off time	2 s
Mode	Headspace Trap
Injector temperature	250 °C
Injection time	2.0 min
Trap Flow	50 mL/min
Injection depth	50 mm
Pre-injection delay	1 s
Post-Injection delay	1 s
Post-Injection Syringe Flush	10 s
Injector penetration speed	25 mm s ⁻¹
Injection speed	50 cm ³ s ⁻¹
Syringe Temperature	70
Syringe Sample Flush volume	1.0 mL
Syringe Sample Flush counts	1
Syringe Sample Volume	1.0 mL
Syringe Vial penetration speed	25 mm s ⁻¹
Syringe Filling Speed	50 mm s ⁻¹
Total number of injections of cold trap	4
Enrichment delay	0.5 min

Cold Trap Purge Time	1 min
Cold Trap Purge Flow	50 cm ³ min ⁻¹
Cold Trap low temperature	-10 °C
Cold trap heating rate	Max °C min ⁻¹
Cold trap high temperature	300 °C
Cold trap Desorb Time	5.0 min
Cold Trap split flow	0 cm ³ min ⁻¹
Flow Path temperature	150°C

Table S2: Summary of GC conditions used for Getxent™ tubes analysis.

Parameter	Setting
He carrier gas flow (constant flow)	1.3 cm ³ min ⁻¹
Front inlet temperature	150 °C
Initial Oven Temperature	30 °C
Initial hold time	2 min
Oven temperature program	5 °C min ⁻¹ to 300 °C, hold 8 minutes
Total run time	64 min
Post Run Temperature	30 °C
Post Run Time	0 min

Table S3: Summary of Mass Spectrometric parameters for Bench ToF Markes International used to perform sample analysis.

Parameter	Setting
Scan Type	Continuous
Mass Range	40 to 500 m/z
Ionization Type	EI -70V
Scan Time	4 Scans/second
Transfer line temperature	310 °C
Ion Source temperature	200 °C
Solvent delay time	0.0 min

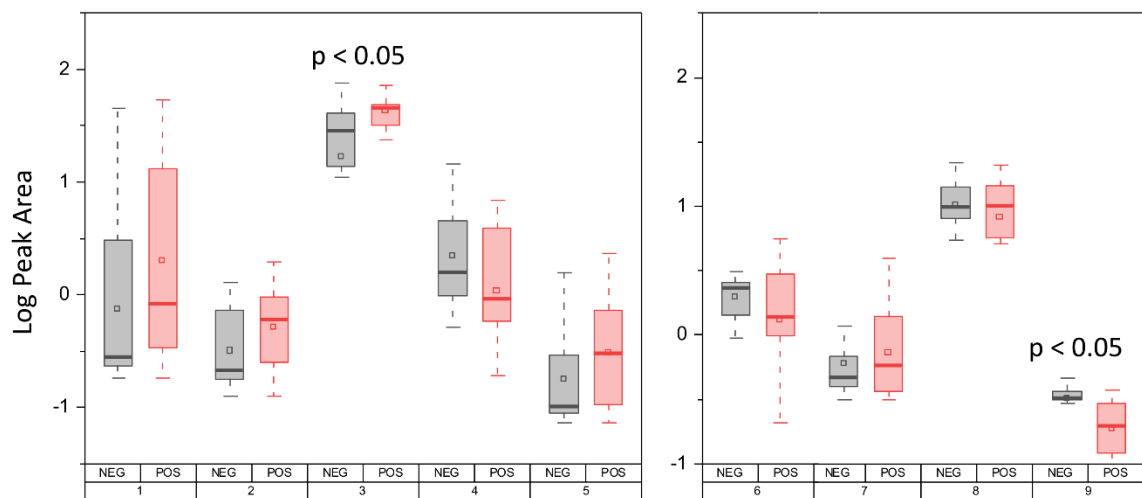


Figure S1. Box Whisker plot of peak areas distribution for 9 VOCs used to produce PCA-X model for COVID-19 negatives (black) vs positives (red) group. The box plots show median and 25% and 75% IQR with $1.5 \times$ IQR, indicated by the whiskers.