



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

Marthe Charles ^{1,2}, Dorota Ruszkiewicz^{2,3}, Eric Eckbo^{1,2}, Elizabeth Bryce^{1,4}, Teresa Zurberg ⁴, Austin Meister², Lâle Aksu⁴, Leonardo Navas⁴ and Renelle Myers^{2,3}

¹Division of Medical Microbiology, Vancouver Coastal Health, Vancouver, BC, Canada. ²University of British Columbia, Faculty of Medicine, Vancouver, BC, Canada. ³British Columbia Cancer Research Institute, Vancouver, BC, Canada. ⁴Quality and Patient Safety, Vancouver Coastal Health, Vancouver, BC, Canada.

Corresponding author: Marthe Charles (marthe.charles@vch.ca)



Shareable abstract (@ERSpublications)

Trained canines can accurately identify positive cases of COVID-19. GC-MS can be used to confirm the presence of volatile organic compounds (VOCs) on training aids, which can be compared to previously reported key VOCs in COVID-19 breath research. <https://bit.ly/4bTmPYU>

Cite this article as: Charles M, Ruszkiewicz D, Eckbo E, *et al.* The science behind the nose: correlating volatile organic compound characterisation with canine biodetection of COVID-19. *ERJ Open Res* 2024; 10: 00007-2024 [DOI: 10.1183/23120541.00007-2024].

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

Received: 5 Jan 2024
Accepted: 20 Feb 2024

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 principal components analysis 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 Matthew's correlation coefficient, 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 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 of



quickly screening 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 face mask, sweat (Getxent tubes) and gargle (Getxent tubes) [8]. To 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 (VOC) 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 authorisation for a test to diagnose COVID-19 based on exhaled VOCs [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: 1) describes the ability of GC-MS to separate positive and negative COVID-19 odours obtained from clinical samples and captured on Getxent tubes; 2) compares sensitivity and specificity of the dogs with GC-MS; and 3) 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 inpatient 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 10-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 (bioMérieux, Lyon, France) panel, which 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 value <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 h 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 (α) variants or strains with N501Y and E484 K mutations.

Canine training and 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 been subject to 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

partner). Hunting breeds were utilised 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 analysed using Excel 2016 (Microsoft) and the categorical agreement, analytical 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 confidence intervals. Inter-rater reliability (kappa agreement) was calculated using GraphPad Prism (GraphPad Software, Boston, MA, USA).

Canine maintenance training

Between 19 February and 18 October 2021, the canine detection specialist (CDS) team continued to run five to 10 scent stands weekly to preserve the accuracy and interest of the dogs. The CDS team used 20 COVID-19 positive tubes and 20 COVID-19 negative 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 and 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 20-cm³ headspace vials and spiked with 1 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 20 µg·mL⁻¹ and 15 µg·mL⁻¹, respectively, diluted in methanol (MeOH). Samples were then heated to 70°C and agitated for 20 min using a HS-Centri (Markes International) autosampler before 1 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-min-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 × 0.25 mm × 0.25 µm was used for the chromatographic separation. Full details of the method can be found in supplementary tables S1–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 ±2×SD of the mean of the blank tubes (95% confidence interval). 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 (Variable Influence on Projection) and p-corr values. The remaining candidate compound's peaks were taken for principal components analysis (PCA-X), which 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:

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

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 and whisker plots, with an ANOVA p-value of 0.001 (figure 4a) when comparing scores of COVID-19 negative versus 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 analysed, 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 1) the individual VOC levels (integrated Log (10)) peak areas and the dogs’ percentage of error, and 2) 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 >100 training sessions each.

CDS validation and maintenance sessions

Validation sessions

12 runs over 2 days were conducted and 120 new samples were used for validation. 24 were positive odours (eight gargle, eight sweat and eight breath); 21 were negative (five gargle, eight sweat and eight 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 positive predictive value (PPV) would be 60.4% and the negative predictive value (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).

Maintenance training sessions

For the maintenance sessions, each dog completed more than 100 sessions specifically using 40 gargle samples (20 positive, 20 negative). Dog A failed to alert four times on a positive case (twice on the same tube) and six 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 four times on an expected negative case (figure 1).

GC-MS analysis and modelling

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 nine of 13 features selected in OPLS (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 (figure 3). The receiver operating characteristic (ROC) curve created on this basis provided

TABLE 1 Comparison of dogs’ overall performance and gargle performance for COVID-19 detection

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

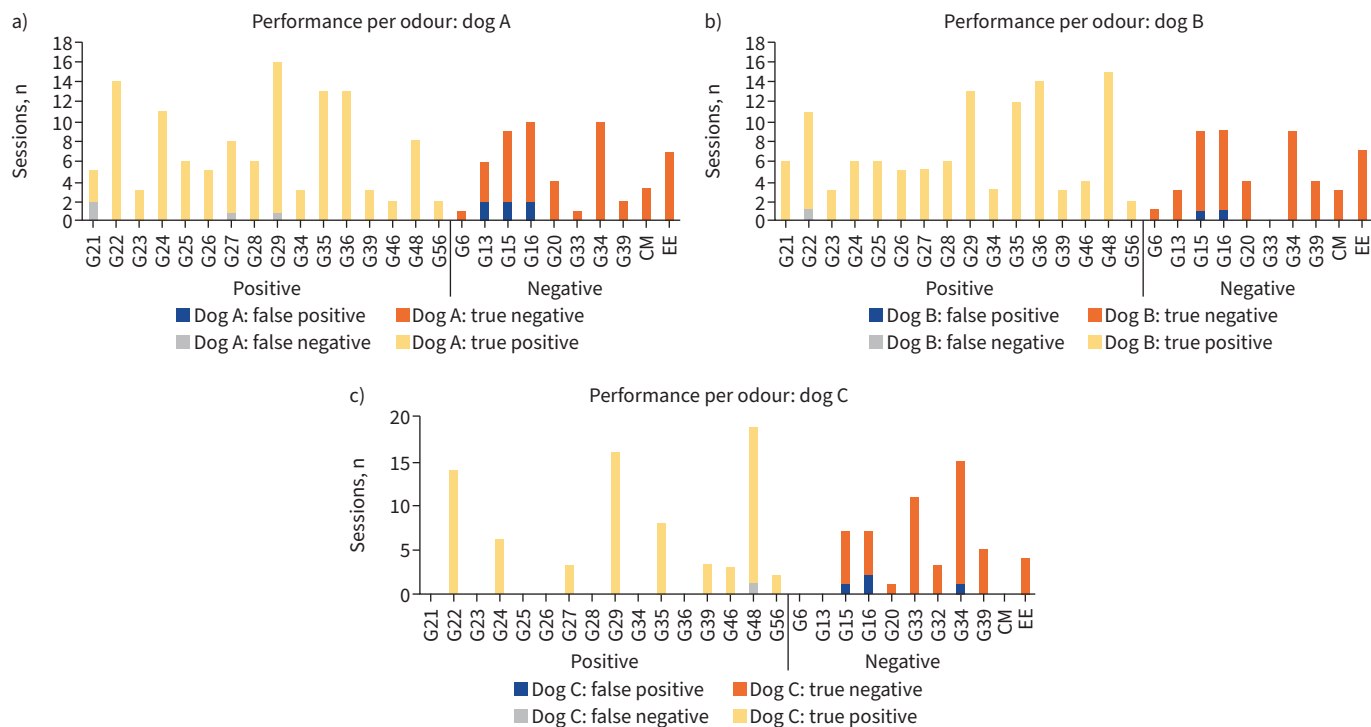


FIGURE 1 Comparison of performance of Getxent tubes per number of sessions (y-axis) per gargle with and without COVID-19 (x-axis) per dog.

accuracy of the classification to be 77.1% with a fitted area under the ROC (AUROC) of 0.761 (95% CI 0.550–0.901) and sensitivity and specificity of 77.8% and 76.5%, respectively (figure 4b). When comparing chemometric scores of COVID-19 negative *versus* 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).

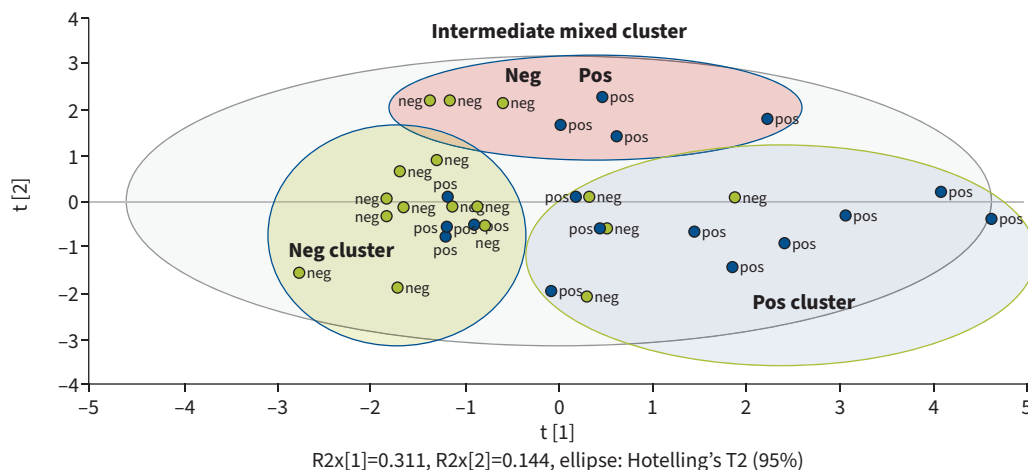


FIGURE 2 Two-component principal components analysis (PCA-X) model. PCA-X model of 35 Getxent tubes for COVID-19 positive (pos) *versus* negative (neg) gargle samples, based on nine features selected by orthogonal partial least squares-discriminant analysis. Green cluster: samples assigned as COVID-19 negative; blue cluster: samples assigned as COVID-19 positive; orange cluster: intermediate mixed cluster.

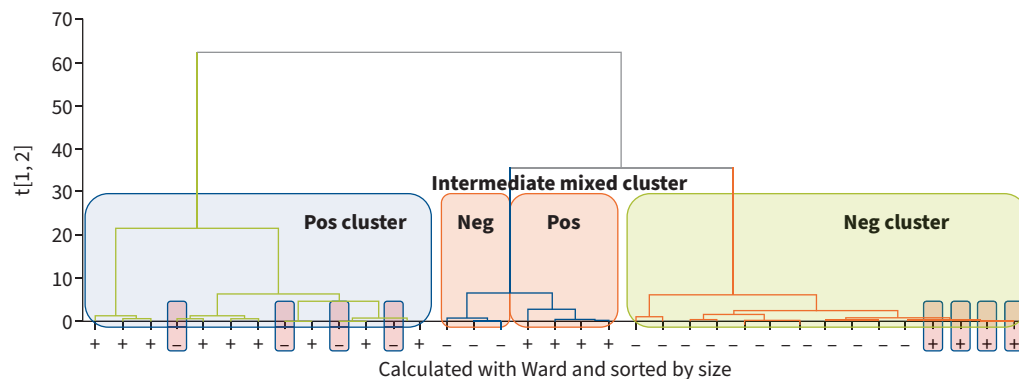


FIGURE 3 Dendrogram from hierarchical analysis of 35 COVID-19 positive (pos; +) and negative (neg; -) gargle samples, showing three distinctive classes: positive, negative and intermediate mixed cluster with two branches with positive and negative separation.

Independent sample t-testing was also performed on each individual compound's peak used in the PCA-X model. Overall, two of nine VOCs, benzaldehyde and one unidentified compound, were statistically significant with a p-value of <0.05 (table 2). Differences in the peak areas are also demonstrated in the form of box and whisker plots presented in supplementary figure S1.

Performance of the model versus dog's alert

Of the 35 Getxent tubes sent for GC-MS 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. Overall, these three canines were exposed to the samples over a total number of 426 trials with an overall 95.56% (95% CI 92.37–97.68%) sensitivity and 96.15% (95% CI 91.82–98.58%) specificity (table 3).

Correlation between dogs and GC-MS VOCs

All three dogs were accurate (95.77%, 95% CI 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 VOC levels and incorrect alerts by the dogs. However, there was

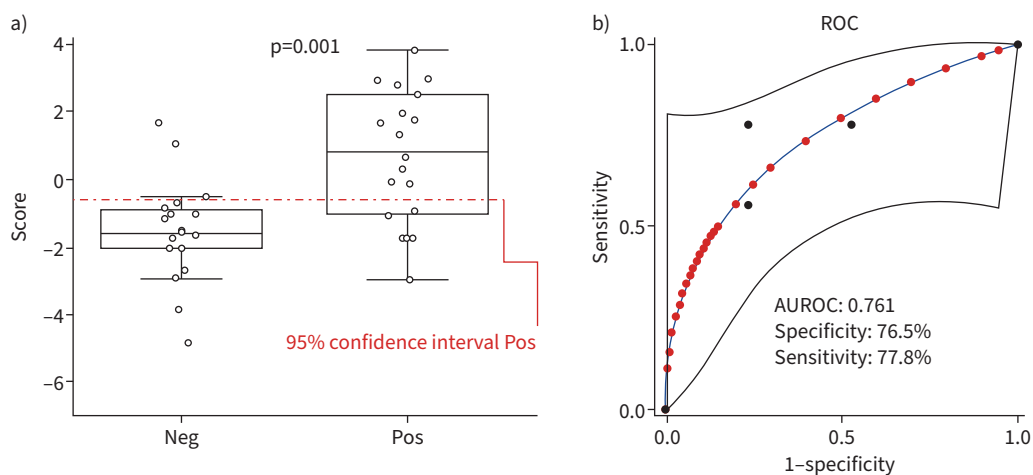


FIGURE 4 Chemometric score and predictive model. a) Box and whisker plot of chemometric score based on peak areas of nine features selected from multivariate analysis, according to Equation 1. b) Receiver operating curve (ROC) based on principal components analysis dendrography model classification; with accuracy of separation between positive (Pos) and negative (Neg) gargle samples collected on Getxent tubes of 77.1% (sensitivity of 78% and specificity of 77%).

TABLE 2 Features selected by orthogonal partial least squares-discriminant analysis (OPLS-DA) analysis and in principal components analysis (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

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

TABLE 3 Summary of dog training on 24 Getxent tubes used in gas chromatography-mass spectrometry (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 Matthew's correlation coefficient 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 for dog B and 0.412 for dog C.

TABLE 4 Summary of dogs alert versus gas chromatography-mass spectrometry (GC-MS) responses

	Dogs alerted (positive) TP	Dogs not alerted (negative) 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	$F1 = 2TP / (2TP + FP + FN)$
Matthew's correlation coefficient	0.37	$TP * TN - FP * FN / \sqrt{((TP + FP) * (TP + FN) * (TN + FN) * (TN + FN))}$

TP: true positive; TN: true negative; FP: false positive; FN: false negative; TPR: TP rate; SPC: specificity; PPV: positive predictive value; NPV: negative predictive value; FPR: FP rate; FDR: false discovery rate; FNR: FN rate; ACC: accuracy.

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 characterised the VOC 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 analytical 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 *Clostridioides 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 [15]. The medical field has been slow in the advancement of VOC-based clinical tests because of challenges with VOC collection standardisation and analytical platform sensitivity and standardisation [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 standardised 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 standardised 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], and in one recent publication benzaldehyde made it to the final PCA [15]. Interestingly, the analysed training aids were also associated with this VOC pattern and triggered very clear canine alerts strengthening the theory that dogs recognise 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, *i.e.* parts per trillion [31], *versus* 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 the 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 analytical equipment to detect a complete VOC profile, also demonstrated by the fact that the dogs performed better at detecting COVID-19 in general.

Understanding that specific ratios of VOCs are more important than individual concentrations of VOCs 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, *i.e.* treats, toys, cheers, *etc.* Consequently, the observed kappa agreement between the dogs' first response to the PCA-X classification could increase over time.

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 standardisation of the methods used to train canines in biodetection. It confirms that ultra clean tubes like Getxent can reliably be used to train dogs safely and accurately in COVID-19 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.

Provenance: Submitted article, peer reviewed.

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.

Support statement: This study was supported by Health Canada. Funding information for this article has been deposited with the Crossref Funder Registry.

References

- 1 Blondeau J. Clinical microbiology laboratories and COVID-19: an interview with Joseph Blondeau. *Future Microbiol* 2021; 16: 615–618.
- 2 Tsao J, Kussman AL, Costales C, *et al.* 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* 2022; 5: e2217234.
- 3 Bekliz M, Adea K, Puhach O, *et al.* Analytical sensitivity of eight different SARS-CoV-2 antigen-detecting rapid tests for omicron-BA.1 variant. *Microbiol Spectr* 2022; 10: e0085322.
- 4 Ten Hagen NA, Twele F, Meller S, *et al.* Canine real-time detection of SARS-CoV-2 infections in the context of a mass screening event. *BMJ Glob Health* 2022; 7: e010276.
- 5 Guest C, Dewhirst SY, Lindsay SW, *et al.* Using trained dogs and organic semi-conducting sensors to identify asymptomatic and mild SARS-CoV-2 infections: an observational study. *J Travel Med* 2022; 29: taac043.
- 6 Vesga O, Agudelo M, Valencia-Jaramillo AF, *et al.* Highly sensitive scent-detection of COVID-19 patients *in vivo* by trained dogs. *PLoS One* 2021; 16: e0257474.
- 7 Grandjean D, Gallet C, Julien C, *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* 2022; 17: e0262631.
- 8 Charles M, Eckbo E, Zurberg T, *et al.* In search of COVID-19: the ability of biodetection canines to detect COVID-19 odours from clinical samples. *J Assoc Med Microbiol Infect Dis Can* 2022; 7: 343–349.
- 9 Dickey T, Junqueira H. Toward the use of medical scent detection dogs for COVID-19 screening. *J Osteopath Med* 2021; 121: 141–148.
- 10 Ibrahim W, Cordell RL, Wilde MJ, *et al.* Diagnosis of COVID-19 by exhaled breath analysis using gas chromatography-mass spectrometry. *ERJ Open Res* 2021; 7: 00139–2021.
- 11 Ruszkiewicz DM, Sanders D, O'Brien R, *et al.* Diagnosis of COVID-19 by analysis of breath with gas chromatography-ion mobility spectrometry: a feasibility study. *EClinicalMedicine* 2020; 29: 100609.
- 12 Chen H, Qi X, Zhang L, *et al.* COVID-19 screening using breath-borne volatile organic compounds. *J Breath Res* 2021; 15: 047104.

- 13 Berna AZ, Akaho EH, Harris RM, *et al.* Reproducible breath metabolite changes in children with SARS-CoV-2 infection. *ACS Infect Dis* 2021; 7: 2596–2603.
- 14 Grassin-Delyle S, Roquencourt C, Moine P, *et al.* Metabolomics of exhaled breath in critically ill COVID-19 patients: a pilot study. *EBioMedicine* 2021; 63: 103154.
- 15 Myers R, Ruskiewicz DM, Meister A, *et al.* Breath testing for SARS-CoV-2 infection. *EBioMedicine* 2023; 92: 104584.
- 16 US Food and Drug Administration (FDA). Coronavirus (COVID-19) Update: FDA Authorizes First COVID-19 Diagnostic Test Using Breath Samples. Test Provides Results in Less Than Three Minutes. 2022. Date last accessed: 11 March 2024. <https://public4.pagefreezer.com/browse/FDA/11-03-2024T11:12/https://www.fda.gov/emergency-preparedness-and-response/about-mcmi/mcmi-news-archive>
- 17 Myers R, Ruskiewicz D, Meister A, *et al.* Breath collection protocol for SARS-CoV-2 testing in an ambulatory setting. *J Breath Res* 2022; 16.
- 18 Mendel J, Frank K, Edlin L, *et al.* Preliminary accuracy of COVID19 odor detection by canines and HSSPMEGCMS using exhaled breath samples. *Forensic Sci Int Synerg* 2021; 3: 100155.
- 19 Pleil JD, Wallace MAG, McCord J, *et al.* 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* 2019; 14: 016006.
- 20 Amann A, Costello BL, Miekisch W, *et al.* The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J Breath Res* 2014; 8: 034001.
- 21 Laffel L. Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes. *Diabetes Metab Res Rev* 1999; 15: 412–426.
- 22 Pauling L, Robinson AB, Teranishi R, *et al.* Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proc Natl Acad Sci U S A* 1971; 68: 2374–2376.
- 23 Hogman M. Extended NO analysis in health and disease. *J Breath Res* 2012; 6: 047103.
- 24 Graham DY, Klein PD, Evans DJ, *et al.* *Campylobacter pylori* detected noninvasively by the ¹³C-urea breath test. *Lancet* 1987; 1: 1174–1177.
- 25 Bryce E, Zurberg T, Zurberg M, *et al.* Identifying environmental reservoirs of *Clostridium difficile* with a scent detection dog: preliminary evaluation. *J Hosp Infect* 2017; 97: 140–145.
- 26 Risby T. Critical issues for breath analysis. *J Breath Res* 2008; 2: 030302.
- 27 Herbig J, Beauchamp J. Towards standardization in the analysis of breath gas volatiles. *J Breath Res* 2014; 8: 037101.
- 28 Meller S, Al Khatri MSA, Alhammadi HK, *et al.* Expert considerations and consensus for using dogs to detect human SARS-CoV-2-infections. *Front Med (Lausanne)* 2022; 9: 1015620.
- 29 Remy R, Kemnitz N, Trefz P, *et al.* Profiling of exhaled volatile organics in the screening scenario of a COVID-19 test center. *iScience* 2022; 25: 105195.
- 30 Laird S, Debenham L, Chandla D, *et al.* Breath analysis of COVID-19 patients in a tertiary UK hospital by optical spectrometry: the E-Nose CoVal study. *Biosensors (Basel)* 2023; 13: 165.
- 31 Guest C, Otto CM. Editorial: canine olfactory detection. *Front Vet Sci* 2020; 7: 100.