A feasibility study of Covid-19 detection using breath analysis by high-pressure photon ionization time-of-flight mass spectrometry

Peize Zhang1,2, Tantant Ren1,2, Haibin Chen1,2,∗, Qingyun Li2, Mengqi He1, Yong Feng1, Lei Wang1, Ting Huang1, Jing Yuan2, Guofang Deng1,2 and Hongzhou Lu1,2

1 Department of Pulmonary Disease and Tuberculosis, The Third People’s Hospital of Shenzhen, No. 29, Bulan Road, Longgang District, Shenzhen 518112, Guangdong, People’s Republic of China
2 Breax Laboratory, PCAB Research Center of Breath and Metabolism, Beijing, People’s Republic of China
3 Department of Disease Control, The Third People’s Hospital of Shenzhen, No. 29, Bulan Road, Longgang District, Shenzhen 518112, Guangdong, People’s Republic of China
4 Department of Infectious Disease, The Third People’s Hospital of Shenzhen, No. 29, Bulan Road, Longgang District, Shenzhen 518112, Guangdong, People’s Republic of China
5 Contributed equally to this study.

∗ Authors to whom any correspondence should be addressed.
E-mail: haibinc@hotmail.com, jxxk1035@yeah.net and luhongzhou@fudan.edu.cn

Keywords: COVID-19, volatile organic compounds, machine learning, breath test

Supplementary material for this article is available online

Abstract
Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has caused a tremendous threat to global health, polymerase chain reaction (PCR) and antigen testing have played a prominent role in the detection of SARS-CoV-2-infected individuals and disease control. An efficient, reliable detection tool is still urgently needed to halt the global COVID-19 pandemic. Recently, the food and drug administration (FDA) emergency approved volatile organic component (VOC) as an alternative test for COVID-19 detection. In this case-control study, we prospectively and consecutively recruited 95 confirmed COVID-19 patients and 106 healthy controls in the designated hospital for treatment of COVID-19 patients in Shenzhen, China. Exhaled breath samples were collected and stored in customized bags and then detected by high-pressure photon ionization time-of-flight mass spectrometry for VOCs. Machine learning algorithms were employed for COVID-19 detection model construction. Participants were randomly assigned in a 5:2:3 ratio to the training, validation, and blinded test sets. The sensitivity (SEN), specificity (SPE), and other general metrics were employed for the VOCs based COVID-19 detection model performance evaluation. The VOCs based COVID-19 detection model achieved good performance, with a SEN of 92.2% (95% CI: 83.8%, 95.6%), a SPE of 86.1% (95% CI: 74.8%, 97.4%) on blinded test set. Five potential VOC ions related to COVID-19 infection were discovered, which are significantly different between COVID-19 infected patients and controls. This study evaluated a simple, fast, non-invasive VOCs-based COVID-19 detection method and demonstrated that it has good sensitivity and specificity in distinguishing COVID-19 infected patients from controls. It has great potential for fast and accurate COVID-19 detection.

1. Introduction
Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has been mutating continuously with stronger transmissibility and virulence and has caused a tremendous threat to global health [1]. Quantitative reverse transcription polymerase chain reaction (RT-qPCR), which can detect SARS-CoV-2 ribonucleic acid on nasopharyngeal or oropharyngeal swabs, has been used frequently to identify individuals infected with the virus and disease control [2]. However, the cost of nucleic acid testing is relatively high even though some self-administered test kits have been developed [3]. The whole procedure from sample collection to the test consumes many staffing and material resources, generates a large amount of
medical waste, and puts medical providers who perform nucleic acid testing at great risk of COVID-19 infection at the same time. Thus, an efficient, reliable, and low-cost detection tool is urgently required to halt the global COVID-19 pandemic, especially in low-resource countries.

In recent years, exhaled breath testing technologies have drawn great attention due to their non-invasiveness, convenience, and rapidity [4]. Volatile organic compounds (VOCs) in the air exhaled by humans are the end products of various metabolic pathways, which contain important information, such as the body’s metabolic state, pathological status, and pollution exposure. VOCs are useful for diagnosing a variety of infectious diseases, including respiratory [5, 6] and gastrointestinal infections [7, 8]. Some studies have proven the sensitivity and specificity of using VOCs from exhaled breath for COVID-19 detection to some extent [9–11]. Recently, FDA emergency approved the first COVID-19 diagnostic test using breath samples [12, 13]. For VOC detection, online mass spectrometry could provide more rapid breath detection and analysis results. Recently, different online mass spectrometry technologies have been developed to analyze exhaled breath, such as proton transfer reaction MS (PTR-MS) [14], secondary electrospray ionization MS (SESI-MS) [15, 16], and high-pressure photon ionization time-of-flight mass spectrometry (HPPI-TOFMS) [17]. The HPPI-TOFMS platform has been used for lung cancer and esophageal cancer detection [18–20] and achieved good performances with sensitivity (SEN) and specificity (SPE) > 90%.

In this study, we used self-designed and developed HPPI-TOFMS [17] to analyze the disease-related small-molecule volatile characteristic metabolites in exhaled breath and performed RT–qPCR testing in parallel using nasopharyngeal swabs as the reference standard to confirm SARS-CoV-2 infection. We aimed to investigate the performance of HPPI-TOFMS based VOC analysis in detecting COVID-19 via the exhaled breath.

2. Materials and method

The study was approved by the Ethics Committee of Shenzhen Third People’s Hospital (number: 2020-00251), and written informed consent was obtained from all of the participants. All clinical data were extracted from the medical system, and the personal information of all patients was masked. The hospital undertook that in using these data, no personal information of any patient was revealed and this study complied with the Declaration of Helsinki regarding confidentiality and ethical standards. We enrolled 95 patients with Delta variants of SARS-COV-2 infection and 106 health care providers as healthy controls (HCs) in the isolation ward from 1 December 2021, to 7 February 2022, in The Third People’s Hospital of Shenzhen, which was the designated hospital for COVID-19 patients’ care and management. All patients were hospitalized with positive PCR test results. Then, the breath samples were synchronously collected on the day the PCR samples were collected at an interval of 3–7 d for all enrolled patients. For every health care provider, breath samples and PCR samples were synchronously collected on the day of enrollment. All PCR samples were collected via nasopharyngeal swabs.

3. The procedure of sampling collection and nosocomial infection control

All breath samples were collected according to the standard operation and study protocol. The detailed procedure was as follows: First, all participants were asked not to have any spicy and pungent foods, alcohol, or coffee the night before the test and they rinsed their mouths with clean water before sampling. Then, a sampling bag made of polyether-ether-ketone (PEEK) was labeled by the health provider. The seal cap was removed, and the air valve was adjusted to an open state. A single-use mouthpiece was connected to the inlet of the sampling bag for breath collection. While the participants exhaled the air into the sampling bag by mouth, the health provider kept a distance of one meter and instructed the participant to hold the bag and breath into the bag by mouth one or more times until the bag was about 2/3 inflated. After that, the valve was adjusted to a closed state, and a seal cap was placed over the inlet of the sampling bag. The used mouthpiece would be removed and discarded in the medical waste bin for sterilization and destruction. The sample bag would be placed in a plastic bag and sterilized with 75% alcohol-containing disinfectant spray. The sealed samples were then transported in a transfer container to a negative pressure testing room for VOC detection on HPPI-TOFMS.

To meet the needs of nosocomial infection prevention in hospitals, VOC detection was finished in the isolation ward for COVID-19 patients within 24 h. All post-test sample bags with residual air was discarded immediately in a high-temperature disinfection box. Routine disinfection was carried out at any time before and after the test.

4. VOC detection

VOC detection was conducted on our developed HPPI-TOFMS platform, which is consisted of a vacuum ultra violet (VUV) lamp-based HPPI ion source and an orthogonal acceleration TOF mass analyzer. The HPPI ion source was designed with two ionization models: soft HPPI ionization and collision-induced dissociation ionization. The soft-HPPI model is adopted in this study, it will
predominantly produce radical cations \((M^+\)) which are formed as \(M + h\gamma \rightarrow M^+ + e\). A commercial VUV-Kr lamp with a photon energy of 10.6 eV and a photon flux of 1011 photons s\(^{-1}\) was adopted for the gas-phase sample ionization. Breath samples were directly introduced through a 250 \(\mu\)m i.d. 0.60 m long stainless-steel capillary. Most VOCs with an ionization potential lower than 10.6 eV were ionized in the ionization region directly [21]. Then, the ion transmission system effectively transferred these ions from the ion source into the orthogonal acceleration, reflection TOFMS mass analyzer, where mass spectrum peaks with \(m/z< 500\) were detected by HPPI-TOFMS and a spectrogram with 31 666 data pairs was extracted from each exhaled breath sample. The TOFMS signals were recorded by a 400 ps time-to-digital conversion ratio at 25 kHz, and all the mass spectra were accumulated for 60 s. Thus, it took 1 min for one sample to go through a detection. The TOFMS signals were positively correlated with the concentration of the VOC ions. The detection limit is down to 0.015 ppbv (parts per billion by volume) for aliphatic and aromatic hydrocarbons [17]. The test result was shown on the computer screen connected to the platform. As shown in supplementary figure 1, the results of our experiments demonstrate that the PEEK breath bags and storage for 24 h had a limited impact on the collected samples.

5. COVID-19 detection modeling

Figure 1 illustrates the main steps of the proposed breath analysis-based COVID-19 detection study. The mass spectrometry obtained via breath detection is a representation of the current metabolic status of the body. Before building the COVID-19 detection model, noise reduction, baseline correction, and VOC ions detection were completed. Besides the background subtraction taken in the built-in data processing of HPPI-TOFMS, anti-symmetric wavelet transformation based noise-reducing and baseline drift correction was applied to avoid randomized noise by the Python package pywavelets [22]. To transfer the discrete signal of mass spectra data to standard VOC ions data, we calculated the area of the strongest peak in the range of \((x-0.1, x + 0.1)\) as the feature of VOC ion with \(m/z\) close to \(x\). Considering no signal detected for \(m/z< 20\) and \(m/z> 320\), 1500 VOC ions were detected from the \(m/z\) range of \([20, 320]\) with an interval of 0.2, which were regarded as breath-omics features for machine learning (ML) model construction.

Considering that the cohort is small and VOC ions are many in this study, a statistical and correlation-based feature selection method was executed to avoid model over-fitting. VOC ions without significant differences (\(p\)-value > 0.05) and VOC ions with a smaller standard deviation between two VOC ions with a high correlation coefficient (CC > 0.9) were excluded. As the ML model was constructed, the most important ions could be confirmed based on the feature importance or coefficient in ML model training. Ions difference analysis was also implemented on the relative density of VOCs among different patient groups.

In this work, several popular ML models, including random forest (RF) [23], support vector machine (SVM) [24], logistic regression (LR) [25], eXtreme gradient boosting (XGB) [26], K nearest neighbors (KNN) [27], and decision tree (DT) [28], were employed as classifiers for COVID-19 detection. The brief description and the main parameter settings of all employed ML models were illustrated in supplement table 1.

As shown in figure 1(c), the datasets of all of the enrolled participants were randomly split into three groups by patients: 50% (COVID-19: 47, HC: 53) for model construction, 20% (COVID-19: 19, HC: 21) for model validation and confirming the cut-off value with optimal Youden-index [29], and the remaining 30% (COVID-19: 28, HC: 32) for testing with clinical information and label blinded. There were between 1 and 13 matched PCR and breath tests performed for each patient. Thus, the number of samples involved in training, validation, and test was not as same as the number of patients. A threshold cycle value (C\(_t\) value \(\leq 40\)) was set to confirm the positive PCR results for a single PCR test. Figure 2 illustrated the PCR test records and results of patients in the train, validation, and test sets. No. 0 PCR test was conducted before enrolling in the hospital, which was positive for all enrolled patients. However, there are some ‘false-negative’ results between two positive PCR test results on 18/95 (18.9%) patients. Thus, we have to confirm the infection status based on follow-up PCR tests. In this study, all the PCR tests were identified as four types: (a) single positive; (b) follow-up positive (single negative but followed by positive PCR test result); (c) single negative; (d) follow-up negative (with continuing negative PCR test results for at least 2 weeks), which were illustrated with different colors in figure 2. Since single negative samples were uncertain, these samples were excluded for model training in this study. Samples with single positive and follow-up positive PCR results were regarded as true positive samples (termed as POSITIVE). Samples with follow-up negative PCR results and samples from HCs were regarded as true negative samples (termed as NEGATIVE).

In this study, we constructed COVID-19 detection models via six popular ML methods, RF, SVM, LR, XGB, KNN, and DT, on POSITIVE and NEGATIVE samples and evaluated their performance in the validation and test sets.
Figure 1. The main steps of the proposed VOCs based COVID-19 detection study. (a) Breath sample collection, (b) VOC ions detection on the HPPI-TOFMS platform and a spectrogram example, and (c) training, validation, and test of the COVID-19 detection model.

Figure 2. The PCR test records and results of train, validation, and test sets. The horizontal axis represents the serial number of patients in the train, validation and test sets in figures (a)–(c), where patients were ordered by the number of PCR tests taken. The vertical axis represents the serial number of PCR tests taken for each patient. No. 0 PCR test was conducted before enrolling in the hospital. Wine Red: single positive. Red: follow-up positive (single negative but followed by positive PCR test result). Orange: single negative. Pink: follow-up negative (with continuing negative PCR test results for at least 2 weeks). Light pink: No sample.

6. Performance evaluation and statistical analysis

To evaluate the performance of the constructed ML and VOCs based COVID-19 detection models, the model detection results were compared with the PCR confirmed diagnosis results. SEN, SPE, positive predictive value (PPV), negative predictive value (NPV), and accuracy (ACC) were calculated to evaluate the COVID-19 detection performance. A receiver operating characteristic curve (ROC) was plotted, and the area under the ROC curve (AUC) was calculated to assess the overall performance of the VOCs based COVID-19 detection model. 95% confidence interval (CI) was also calculated for all performance evaluation metrics.

All statistical analyses were performed on statistical analysis system (SAS) version 9.4 (SAS Institute Inc., Cary, NC, USA) and Origin software (version 2018). Descriptive statistics are reported as frequencies (percentages) for categorical variables or medians (interquartile range, IQR) for continuous variables. We compared the demographic characteristics among different patient groups using the Mann–Whitney U test for continuous variables and the chi-square test for categorical variables. The ANOVA test was employed for the performance comparison between validation and test sets. A $p$-value < 0.05 was
considered statistically significant for all of the analyses. All of the tests were two-tailed.

7. Results

As illustrated in table 1, the mean age of the collected COVID-19 patients was 39 years, and 67.4% of them were male. The mean age of the HCs was 33 years, and 41.5% of them were male. Significant differences in age, sex, and body mass index (BMI) between the two groups were identified, while no significant difference in smoking status, alcohol use, or comorbidities was noticed.

The classification results of six ML models were illustrated in table 2. The SVM based model failed in discrimination between POSITIVE and NEGATIVE samples. KNN and DT based models have ordinary performances with AUC < 0.90 in validation and test sets. RF, LR, and XGB based models achieved superior performances than other ML models with AUC > 0.90 in validation and test sets. Overall, the XGB based model has the best performance (cutoff = 0.691) in the validation set with a SEN of 91.3% (95% CI: 82.1%, 100%), and an AUC of 0.967 (95% CI: 0.916, 1.00). Moreover, it also performs best in test set with a SEN, a SPE and an AUC of 92.2% (95% CI: 83.8%, 100%), 86.1% (95% CI: 74.8%, 97.4%) and 0.944 (95% CI: 0.896, 0.993) respectively. Thus, the XGB based model was further analyzed in this study. As illustrated in supplementary table 1, SVM is aiming to find the classifier with a maximum geometric interval, and difficulty in finding a suitable kernel function. Thus, SVM is prone to over-fitting. The data set enrolled in this study is discrete, balanced, and with a median feature scale, DT, KNN, and LR are suitable basic classifiers for model training. RF and XGB are the meta and boosting classifiers of DT, respectively. Thus, the performances of these ML models are as expected.

The performances of the XGB-based model were compared between validation and test sets, and analyzed in different sub-groups: POSITIVE vs. HCs, POSITIVE vs. follow-up negative. As shown in figure 2, the following properties were observed: (a) the XGB-based model performs superior in the validation set than test set; (b) the XGB-based model achieved better performances in POSITIVE vs. HC
Figure 3. Performances of the XGB based COVID-19 detection model in validation and test sets. (a) Performances in the validation set; (b) performances in the test set. The Blue full line represents the ROC in POSITIVE vs. NEGATIVE discrimination. Green and pink dotted lines represent the ROC in POSITIVE vs. HC, and POSITIVE vs. follow-up negative discriminations, respectively.

Table 3. Performance comparison of the XGB based COVID-19 detection model between validation and test sets in different subgroups.

| Experiment    | Data sets                | SEN (%)     | SPE (%)     | PPV (%)     | NPV (%)     | ACC (%)     | AUC (p-value) |
|---------------|--------------------------|-------------|-------------|-------------|-------------|-------------|---------------|
| POSITIVE vs.  | Validation (23 vs. 26)   | 91.3(78.6, 100) | 92.3(82.1, 100) | 91.3(81.2, 100) | 92.3(80.3, 100) | 91.8(84.2, 100) | 0.967(0.916, 1.00) | 0.681 |
| POSITIVE vs.  | Test (51 vs. 36)         | 92.2(83.8, 100) | 86.1(74.8, 97.4) | 90.4(81.7, 99.1) | 86.6(79.1, 98.1) | 89.7(83.3, 96.1) | 0.944(0.896, 0.993) | 0.577 |
| POSITIVE vs.  | Validation (23 vs. 20)   | 91.3(78.6, 100) | 100(100, 100) | 100(100, 100) | 100(100, 100) | 95.3(89.1, 100) | 0.985(0.948, 1.00) | 0.892 |
| POSITIVE vs.  | Validation (23 vs. 6)    | 91.3(78.6, 100) | 66.7(28.9, 100) | 91.3(68.8, 100) | 66.7(47.4, 85.9) | 86.2(73.2, 100) | 0.906(0.799, 1.00) | 0.577 |
| POSITIVE vs.  | Test (51 vs. 32)         | 92.2(83.8, 100) | 93.8(85.4, 100) | 95.9(89.3, 100) | 88.2(79.2, 97.3) | 92.8(87.2, 98.3) | 0.966(0.927, 1.00) | 0.577 |
| POSITIVE vs.  | Validation (23 vs. 6)    | 91.3(78.6, 100) | 66.7(28.9, 100) | 91.3(68.8, 100) | 66.7(47.4, 85.9) | 86.2(73.2, 100) | 0.906(0.799, 1.00) | 0.577 |
| POSITIVE vs.  | Test (51 vs. 4)          | 92.2(83.8, 100) | 25.0(0.0, 67.4) | 94.0(73.2, 100) | 20.0(8.9, 31.1) | 87.3(78.5, 96.1) | 0.770(0.658, 0.881) | 0.577 |

than POSITIVE vs. follow-up negative; (c) the performance in distinguishing POSITIVE vs. follow-up negative declined more than in other scenarios.

As the quantitative results illustrated in figure 3 and table 3, there is no significant difference between the performances of XGB based COVID-19 detection model in validation and test sets. In the test set, the XGB based COVID-19 detection model could distinguish the POSITIVE samples with HCs well with a SEN of 92.2% (95% CI: 83.8%, 100%), a SPE of 93.8% (95% CI: 85.4%, 100%), an ACC of 92.8% (95% CI: 87.2%, 98.3%), and an AUC of 0.966 (95% CI: 0.927, 1.00). However, the XGB based COVID-19 detection model performs poorly with low specificity in distinguishing follow-up negative samples from POSITIVE samples, although the overall ACC and AUC metrics are good. The main reason lies in the small amount of follow-up negative samples.

Although some breath samples were uncertain of whether COVID-19 was infected or cured, the XGB based COVID-19 detection model was executed on all breath samples in validation and test sets. The results were exhibited in figure 4. For single and follow-up positive samples, most of them were classified as positive via a breath test. Almost all samples of HCs were classified as negative via breath test with only two false-positive cases in test sets. These results demonstrated the XGB-based COVID-19 detection model was reliable in distinguishing between COVID-19 patients and HCs. 12/37(32.4%) and 8/56(14.3%) of single negative samples were classified as negative via breath test in validation and test sets, respectively. 50% of the follow-up negative samples in validation and test sets were classified as negative via a breath test, which is not very reliable because of the small sample size.

To explore the breath-omics evidence in the COVID-19 detection model, the model used VOC ions were further analyzed. After statistical and correlation based feature selection, 62 VOC ions were
Table 4. Mass spectrum signal comparison of the five most important VOC ions between COVID-19 POSITIVE and NEGATIVE groups. ‘↑’ and ‘↓’ represent the variation tendency in the COVID-19 POSITIVE group relative to the NEGATIVE group. Bold p-value(<0.05) represent there is significant differences between Positive and Negative groups. Chemical Abstracts Service (CAS).

| m/z | Chemicals (CAS number) | Positive (Mean ± STD) | Negative (Mean ± STD) | p-value |
|-----|------------------------|-----------------------|-----------------------|---------|
| 150↑| Phenyl propionic acid/(501-52-0/492-37-5/7782-24-3/2328-24-7) | 421.1 ± 486.3 | 185.4 ± 162.5 | <0.001 |
| 92↓ | Toluene(108-88-3) | 2003.2 ± 1060.2 | 3143.1 ± 2459.2 | <0.001 |
| 78↓ | Benzene(71-43-2) | 395.1 ± 234.5 | 513.9 ± 318.1 | <0.001 |
| 87↑ | Pentanone(107-87-9) | 232.0 ± 163.9 | 132.2 ± 101.3 | <0.001 |
| 88↓ | Isobutyric acid/Butyric Acid/(79-31-2/107-92-6) | 1138.7 ± 1942.9 | 1587.7 ± 1725.5 | <0.001 |

included in model training. To find the important and related VOC ions for COVID-19 detection, we ordered the 62 VOC ions from large to small based on the feature coefficients in modeling, which represent the contribution value of the ions in the model. Then, we fit the models as the ions increased one by one until the ACC tends to be stable. As shown in figure 5, the ACC of the model tends to be stable with the fluctuation range of [−0.05, 0.05] when the five most important VOC ions were involved in the model fitting. Thus, we analyzed the top five most important VOC ions between POSITIVE and NEGATIVE breath samples.

As shown in table 4, the five most important VOC ions with m/z values of 150, 92, 78, 87, and 88 were compared between the mass spectrometer data of POSITIVE and NEGATIVE breath samples. 2/5(40%) VOC ions increased and 3/5(60%) VOC ions decreased in the POSITIVE breath samples compared with the NEGATIVE samples. We infer the possible chemicals of these ions related to COVID-19 based on their m/z, the potential biomarkers published, and the human breath-omics database [30]. The VOC ion with m/z of 88 would be isobutyric acid/butyric acid (CAS: 79-31-2/107-92-6). It is identified as the potential biomarker of COVID-19-induced metabolic dysregulation in he is study [31]. However, they demonstrated that isobutyric acid significantly trended up in non-severe COVID-19 patients in both infection and recovery phases, but with no significant tendency in severe COVID-19 patients when compared with ten HCs. There is no significant difference in butyric acid among HC, non-severe COVID-19 patients, and severe COVID-19 patients. The different variation tendencies may be the results of the severity of infection, different controls, and different sample sizes, which need to be further studied. The VOC ion with m/z of 150 would be phenyl propionic acid, which is a potential indicator of health status [32]. Although the melting point and boiling point of phenyl propionic acid are high. It can be evaporated with water vapor and then be detected from the breath sample. Thus, its relative concentration is low. The VOC ions with m/z of 92 and 78 would be toluene (CAS: 108-88-3) and benzene (CAS: 71-43-2), which are common environmental pollutants from mobile sources in urban settings. Toluene and benzene were proved related to M. tuberculosis strains or tuberculosis [33–35]. The VOC ion with m/z of 87 would be the protonated cation of pentanone (CAS:107-87-9), which relates to colon diseases [36]. Thus, these three ions may be related to the infection reaction of the body.

8. Discussion and conclusion

After the outbreak of COVID-19, scientists from all over the world explored speedy screening tools for COVID-19. SARS-COV-2 could be detected by analyzing VOCs in exhaled breath of patients by gas chromatography ion mobility spectrometry. Early exploratory studies from Edinburgh and Dortmund showed that COVID-19 patients could be distinguished from non-COVID-19 patients using gas chromatography with 80% and 81.5% accuracy in 25 COVID-19 patients in Edinburgh and 65 COVID-19 patients in Dortmund, respectively [37]. However, their analysis process was complex and time-consuming, which makes it difficult to apply in the clinic. Another study using electronic noses for SARS-CoV-2 screening in
preoperative patients also demonstrated that electronic noses could distinguish COVID-19 positive from COVID-19 negative participants even though they could not quantify the detailed contents of the VOCs in the mixture [38]. A study from Zamora-Mendoza BN et al also showed that the use of electronic noses and chemometric analysis could identify the overall profile of VOCs in patients with COVID-19 patients and controls [39]. A recent study from Singapore also demonstrated that VOC detection could identify COVID-19 infected individuals, which achieved >95% SEN and SPE across 501 participants regardless of their displayed symptoms [40]. A meta-analysis demonstrated that the cumulative SEN of the analysis based on breath VOCs was 98.2% (97.5% CI: 93.1%–99.6%), and the SPE was 74.3% (97.5% CI: 66.4%–80.9%) [41]. The first test that uses breath samples to diagnose COVID-19 received emergency use authorization from the FDA on 14 April 2022 [12, 13].

HPPI-TOFMS is a spectrometer operating at VUV-Kr lamp-based ion source. It does not require preprocessing of exhaled air but with a low detection limit of 0.015 ppbv. It has been employed for lung cancer and esophageal cancer detection and achieved good performances with SEN and SPE >90 [18–20]. As the first rapid VOCs based diagnostic test for COVID-19 in China, we found that COVID-19 patients could be easily distinguished from controls using VOCs with HPPI-TOFMS. The SEN and SPE of this technology in detecting COVID-19 was 92.2% (95% CI: 83.8%, 100%), 86.1% (95% CI: 74.8%, 97.4%), respectively. Our study further demonstrated the great potential of exhaled breath detection for COVID-19 screening. Most importantly, it only took a few minutes for sample detection and result interpretation. As a highly contagious and pandemic disease, it is very important to identify the possible source of infection as soon as possible. With high accuracy and efficiency for COVID-19, exhaled breath detection on HPPI-TOFMS is a promising approach for COVID-19 screening.

It has been reported that SARS-CoV-2 enters the body through the respiratory tract and then spreads to different parts of the body. It can be detected in the nasopharynx after the onset of the disease. As the disease progresses, the virus multiplies in the lungs and other organs but disappears from the nasopharynx. Sampling timing and disease development time both affect the accuracy of the PCR test. This explains why the positive PCR tests dropped from 94.39% to 67.15% at 0–10 d after SARS-COV-2 infection and became negative during the late stage of the disease [42]. In addition, insufficient organisms in the sample arising from inappropriate collection, transportation, or handling contribute to ‘false-negative’ PCR results [43]. In our study, we also noticed that some patients experienced positive PCR tests after 1–3 consecutive negative results. The patients had no clinical symptoms but the PCR test became positive. We supposed that the possible causes of the negative PCR results were a lower viral load in the body or insufficient sample collection. Figure 4 illustrated the breath test results on single negative breath samples in validation and test sets. Only 12/37 (32.4%) and 8/56 (14.3%) of single negative samples were classified as negative via breath test in validation and test sets, respectively. We supposed that part PCR results of these single negative samples were ‘false-negative’. These results may imply VOCs based COVID-19 detection method is more sensitive to COVID-19 infection with low virus load, which also occurs in the early stage of infection. A recent autopsy found that viral ribo-nucleic acid (RNA) in body tissues decreased in the middle and late stages of infection, but a low level of viral RNA persisted in organ tissue [44]. It has been demonstrated that some patients and their close contacts are repeatedly negative by RT-qPCR during the first stage of disease onset and finally have a confirmed diagnosis when SARS-CoV-2 RT-qPCR became positive [45], indicating that a more sensitive method to detect SARS-CoV-2 virus at an early stage is required to reduce the possible spread of disease. VOCs tested on HPPI-TOFMS could detect metabolic molecules of COVID-19 infection even at a trace viral level. Since it is a highly contagious disease, this finding might provide a clue for the early detection of COVID-19 simply and easily, which is meaningful for disease control, especially in densely populated and low-resource settings.

There are some limitations of our study. First, the enrolled participants were limited, (a) the sample size was relatively small, (b) there is no patient with other pulmonary disease enrolled, and (c) the method was not validated with external data from independent clinical centers. Secondly, the age, sex, and BMI differences between COVID-19 patients and HCs may make a difference in VOCs in the breath. Thirdly, patients who received different types of food combined with some other diseases might have an impact on the metabolites in VOCs, which might also affect the precision of the results. Fourthly, the choice of ML methods is crude. We just selected the best ML model based on the performance in our data set. It also implies the model on a small data set is not robust. Finally, the chemical characteristics of COVID-19 related VOCs were not completely confirmed. It could be a barrier to subsequent research by other researchers. We expect to resolve or mitigate the above limitations in our future works.

In summary, our study developed and evaluated a simple, fast, non-invasive VOCs based COVID-19 detection method. The results demonstrated that VOCs based COVID-19 detection model has good sensitivity and specificity of 91.3% and 92.3% in distinguishing COVID-19 infected patients from
controls. This technology is potentially to be a rapid, sensitive, and non-invasive tool for the early screening of COVID-19. It may promptly identify suspected COVID-19 patients and alarm the medical system to respond quickly in communities and resource-limited settings.

Data availability statement
The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments
We thank all investigators and caregivers in The Third People’s Hospital in Shenzhen whose names were not listed here, for their contribution to sample collection and patient care. This work was supported by the National Natural Science Foundation of China (82070016), and Projects funded by Shenzhen Third People’s Hospital (G2022003, G2022051).

ORCID iD
Haibin Chen https://orcid.org/0000-0002-5542-683X

References
[1] Sanyaola A, Okorie C, Marinkovic A, Haider N, Abbasi A F, Jaferi U, Prakash S and Balendra V 2021 The emerging SARS-CoV-2 variants of concern Ther. Adv. Infect. Dis. 8 204993612111024372
[2] Bögler B, Fachi M M, Vilhena R O, Cobre A F, Tonin F S and Pontarollo R 2021 Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19 Am. J. Infect. Control 49 21–29
[3] Tang Y W, Schmitz J E, Persing D H and Stratton C W 2020 Laboratory diagnosis of COVID-19: current Issues and challenges J. Clin. Microbiol. 58 e00512–20
[4] Bruderer T, Gaisl T, Gaugg M T, Nowak N, Streckenbach B, Muller S, Moeller A, Kohler M and Zenobi R 2019 On-line analysis of exhaled breath focus review Chem. Rev. 119 10803–28
[5] Li Z-T et al 2021 Exhaled volatile organic compounds for identifying patients with chronic pulmonary aspergillosis Front. Med. 8 720119
[6] Kolik A H, van Berkel J J, Claassens M M, Walters E, Kuijper S, Dallinga J W and van Schooten F J 2012 Breath analysis as a potential diagnostic tool for tuberculosis Int. J. Tuberc. Lung Dis. 16 777–82
[7] Ulansowska A, Kowalowski T, Hryniewicz K, Jackowski M and Buszewski B 2011 Determination of volatile organic compounds in human breath for Helicobacter pylori detection by SPME-GC/MS Biomed. Chromatogr. 25 391–7
[8] Garner C E, Smith S, Bardhan P K, Ratcliffe N M and Probert C S 2009 A pilot study of faecal volatile organic compounds in faeces from cholera patients in Bangladesh to determine their utility in disease diagnosis Trans. R. Soc. Trop. Med. Hyg. 103 1171–3
[9] Wintjens A, Hintsen K F H, Engelen S M E, Lubbers T, Savalkoul P H M, Wesseling G, van der Palen J A M and Bouvy N D 2021 Applying the electronic nose for pre-operative SARS-CoV-2 screening Surg. Endosc. 35 6671–8
[10] Grassin-Delyle S et al 2021 Metabolomics of exhaled breath in critically ill COVID-19 patients: a pilot study EBioMedicine 63 103154
[11] Berna A Z, Akaho E H, Harris R M, Congdon M, Korn E, Neher S, M’Farrej M, Burns J and John A R O 2021 Reproducible breath metabolite changes in children with SARS-CoV-2 infection ACS Infect. Dis. 7 2596–403
[12] FDA 2022 Coronavirus (COVID-19) update: FDA authorizes first COVID-19 diagnostic test using breath samples FDA News Release
[13] Rubin R 2022 First breathalyzer test to diagnose COVID-19 JAMA 327 1860
[14] Trefz P, Schmidt M, Oertel P, Obermeier J, Brock B, Kamysyk S, Dunkl J, Zimmermann R, Schubert J K and Mieckish W 2013 Continuous real time breath gas monitoring in the clinical environment by proton-transfer-reaction-time-of-flight-mass spectrometry Anal. Chem. 85 10321–9
[15] Gaugg M T, Bruderer T, Nowak N, Eifert L, Martinez-Lozano Sinues P, Kohler M and Zenobi R 2017 Mass-spectrometric detection of omega-oxidation products of aliphatic fatty acids in exhaled breath Anal. Chem. 89 10529–34
[16] Singh K D, Del Miguel G V, Gaugg M T, Ibanez A J, Zenobi R, Kohler M, Frey U and Sinues P M L 2018 Translating secondary electrospray ionization–high-resolution mass spectrometry to the clinical environment J. Breath Res. 12 027113
[17] Wang Y, Jiang J, Hua L, Hou K, Xie Y, Chen P, Liu W, Li Q, Wang S and Li H 2016 High-pressure photon ionization source for TOFMS and its application for online breath analysis Anal. Chem. 88 9047–55
[18] Huang Q et al 2021 Assessment of breathomics testing using high-pressure photon ionization time-of-flight mass spectrometry to detect esophageal cancer JAMA Netw. Open 4 e2127042–e
[19] Meng S et al 2021 Assessment of an exhaled breath test using high-pressure photon ionization time-of-flight mass spectrometry to detect lung cancer JAMA Netw. Open 4 e213486–e
[20] Wang P, Huang Q, Meng S, Mu T, Liu Z, He M, Li Q, Zhao S, Wang S and Qiu M 2022 Identification of lung cancer biomarkers based on perioperative breathomics testing: a prospective observational study eClinicalMedicine 47 101384
[21] Zhao X, Liu X, Liu J, Chen J, Fu S and Zhong F 2019 The effect of ionization energy and hydrogen weight fraction on the non-thermal plasma volatile organic compounds removal efficiency J. Phys. D: Appl. Phys. 52 145201
[22] Lee G, Gommers R, Waselewski F, Wohlhaft K and Aaron O 2019 PyWavellets a Python package for wavelet analysis J. Open Source Softw. 4 1237
[23] Breiman L 2001 Random forests Mach. Learn. 45 5–32
[24] Pizer D and Schnyer D M 2020 Chapter 6—Support vector machine Machine Learning A Mechelli and S Vieira (New York: Academic) pp 101–21
[25] Bewick V, Cheek L and Ball J 2005 Statistics review 14: logistic regression J. Crit. Care 9 112
[26] Azarra N A and Juneja A 2019 Extreme gradient boosting with squared logistic loss function Machine Intelligence and Signal Analysis ed M Tanveer and R B Pachori (Singapore: Springer) pp 313–22
[27] Guo G et al 2003 KNN model-based approach in classification On the Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE ed R Meersman (Berlin: Springer) pp 986–96
[28] Jordan M 1994 A statistical approach to decision tree modeling (https://doi.org/10.1145/180339.175372)
[29] Ruopp M D, Perkins N I, Whitecomb B W and Schisterman E F 2008 Youden index and optimal cut-point estimated from observations affected by a lower limit of detection Biom. J. 50 419–30
[30] Kuo T-C et al 2020 Human breathomics database Database 2020
[31] He X et al 2021 COVID-19 induces new-onset insulin resistance and lipid metabolic dysregulation via regulation of secreted metabolic factors Signal Transduct. Targeted Ther. 6 427

[32] Gutiérrez-Díaz I, Fernández-Navarro T, Salazar N, Bartolomé B, Moreno-Arribas M V, López P, Suárez A, de Los Reyes-gavilán C G, Gueimonde M and González S 2018 Could fecal phenylacetic and phenylpropionic acids be used as indicators of health status? J. Agric. Food Chem. 66 10438–46

[33] Tay S T, Hemond H F, Polz M F, Cavanaugh C M, Dejesus I and Krumholz L R 1998 Two new Mycobacterium strains and their role in toluene degradation in a contaminated stream Appl. Environ. Microbiol. 64 1715–20

[34] Vishinkin R et al 2021 Profiles of volatile biomarkers detect tuberculosis from skin Adv. Sci. 8 2100235

[35] Phillips M, Catanese R N, Condos R, Ring Erickson G A, Greenberg J, La B V, Munawar M I and Tietje O 2007 Volatile biomarkers of pulmonary tuberculosis in the breath Tuberculosis 87 44–52

[36] Ahmed I, Greenwood R, Costello B, Ratcliffe N and Probert C S 2016 Investigation of faecal volatile organic metabolites as novel diagnostic biomarkers in inflammatory bowel disease Aliment. Pharmacol. Ther. 43 596–611

[37] Ruszkiewicz D M et al 2020 Diagnosis of COVID-19 by analysis of breath with gas chromatography-ion mobility spectrometry—a feasibility study EClinical Medicine 29 100009

[38] Rodríguez-Aguilar M, Díaz de León-martínez L, Zamora-Mendoza B N, Comas-García A, Guerra Palomares S E, García-Sepúlveda C A, Alcántara-Quintana I E, Díaz-Barriga F and Flores-Ramírez R 2021 Comparative analysis of chemical breath-prints through olfactory technology for the discrimination between SARS-CoV-2 infected patients and controls Clin. Chim. Acta 519 126–32

[39] Zamora-Mendoza B N, Díaz de León-martínez L, Rodríguez-Aguilar M, Mizaikoff B and Flores-Ramírez R 2022 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 122832

[40] Leong S X et al 2022 Noninvasive and point-of-care surface-enhanced raman scattering (SERS)-based breathalyzer for mass screening of coronavirus disease 2019 (COVID-19) under 5 min ACS Nano 16 2629–39

[41] Subali A D, Wiyono L, Yusuf M and Zaky M F A 2022 The potential of volatile organic compounds-based breath analysis for COVID-19 screening: a systematic review & meta-analysis Diag. Microbiol. Infect. Dis. 102 115589

[42] Wikramaratna P S, Paton R S, Ghafari M and Lourenço J 2020 Estimating the false-negative test probability of SARS-CoV-2 by RT-PCR Eurosurveillance 25 2000568

[43] Tahamtan A and Ardabili A 2020 Real-time RT-PCR in COVID-19 detection: issues affecting the results Expert Rev. Mol. Diagn. 20 453–4

[44] Daniel C et al 2022 SARS-CoV-2 infection and persistence throughout the human body and brain Nat. Portfolio (https://doi.org/10.21203/rs.3.rs-1139035/v1)

[45] Long C, Xu H, Shen Q, Zhang X, Fan B, Wang C, Zeng B, Li Z, Li X and Li H 2020 Diagnosis of the Coronavirus disease (COVID-19): rRT-PCR or CT? Eur. J. Radiol. 126 108961