Ultrafast Preliminary Screening of COVID-19 by Machine Learning Analysis of Exhaled NO

Li Yang

Shandong University <https://orcid.org/0000-0001-6570-5704>

Jing Ma

Liyuan Hospital, Tongji Medical College, Huazhong University of Science and Technology

Wei Zhou

Shandong University <https://orcid.org/0000-0003-0006-0671>

Lei Sun (✉ slei@sdu.edu.cn)

Shandong University <https://orcid.org/0000-0001-9960-205X>

Dong Zhai

Shandong University

Guoqiang Ren

Shandong University

Tie Yu

Shandong University

Jie Chen

Liyuan Hospital, Tongji Medical College, Huazhong University of Science and Technology

Jinping Zhang

Liyuan Hospital, Tongji Medical College, Huazhong University of Science and Technology

Morong Wang

Liyuan Hospital, Tongji Medical College, Huazhong University of Science and Technology

Jie Wang

Liyuan Hospital, Tongji Medical College, Huazhong University of Science and Technology

Weiqiao Deng

Shandong University <https://orcid.org/0000-0002-3671-5951>

Research Article

Keywords: COVID-19, Machine Learning, Fractional Exhaled Nitric Oxide (FeNO), Preliminary Screening

Posted Date: July 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-56416/v2>

Abstract

Background A new coronavirus, SARS-CoV-2, has caused the coronavirus disease-2019 (COVID-19) epidemic. Current diagnostic methods mainly include nucleic acid detection, antibody detection, antigen detection, and chest computed tomography (CT) imaging. Although these methods are crucial for the diagnosis of COVID-19, there is a lack of a rapid and economical method for preliminary screening COVID-19.

Methods We measured the FeNO concentrations of 103 subjects without COVID-19 and 46 patients with COVID-19. Using machine learning (ML) method, we build a ML model based on fractional exhaled nitric oxide (FeNO) concentration and features of age, and body size for rapid preliminary screening COVID-19 suspects with low-cost.

Findings The statistical analysis t-test show that there is a significant difference between the FeNO of healthy people and patients with COVID-19. The ML model can screen out the patients with COVID-19 or other diseases, which show abnormal FeNO distributions. An area under the curve of 0.982 and a sensitivity 0.917 have been achieved for preliminary screening COVID-19 suspects. This non-invasive detection method which takes in two minutes and costs less than a dollar could provide a direction for the control of the rapid spread COVID-19.

Interpretation During the COVID-19 pandemic, large numbers and extensive testing of COVID-19 patients remains a problem. Public healthy efforts to limit SARS-CoV-2 spread need to find a more economical and faster screening method.

Research In Context

Evidence before this study

Fractional exhaled nitric oxide (FeNO) has been considered an important biomarker for the detection and diagnosis of respiratory diseases. After the severe acute respiratory syndrome coronavirus (SARS-CoV) outbreak in 2002-2003, relevant studies showed that the NO concentration in the organism changes with the degree of infection by a coronavirus

Added value of this study

This work, we tested the FeNO of 93 healthy people and 46 COVID-19 patients. To our knowledge, we provide the first analysis of FeNO in COVID-19 patients. Using statistic method we found FeNO level is different in healthy people and COVID-19 patients. However, FeNO alone could not be used to identify COVID-19 patients because of the statistic distribution between FeNO in COVID-19 patients and healthy subjects overlaps. Using ML algorithm to integrate FeNO and other human characteristics including age and body size can make a better classification of healthy people and COVID-19 patients.

Implications of all the available evidence

Our finding that FeNO combining with age and body size can be used to preliminary screening COVID-19 could help to control the wide spread COVID-19.

1. Introduction

Transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative virus of the coronavirus disease-2019 (COVID-19) epidemic, can be effectively prevented by developing a rapid and large-scale screening method for the SARS-CoV-2-infected population¹⁻³. Screening suspected patients for isolation and maintaining social distancing can effectively prevent the spread of the virus.⁴⁻⁵ At present, the detection of SARS-CoV-2 infection mainly depends on nucleic acid detection, antibody detection, antigen detection, and chest computed tomography imaging^{6,7}. In the control of the rapid spread of SARS-Cov-2, nucleic acid detection became a standard because of its accuracy and efficiency, which usually costs tens of dollars and takes at least 30 minutes for each case. However, for the screening of large populations such as more than ten million people, this method is too expensive and time-consuming. Therefore, the development of a rapid and accurate preliminary diagnosis for large population is a grand challenge.⁸

Fractional exhaled nitric oxide (FeNO) has been considered an important biomarker for the detection and diagnosis of respiratory diseases⁹. In 1991, Gustafsson and coworkers found that humans produced FeNO at a magnitude of parts per billion (ppb), and there was evidence of endogenous NO synthesis in the lower respiratory tract¹⁰. After at least 28 years of exploration, it was determined that the detection of FeNO could predict lung cancer, asthma, scleroderma, sarcoidosis, and other lung diseases^{9, 11-18}. After the severe acute respiratory syndrome coronavirus (SARS-CoV) outbreak in 2002-2003, relevant studies showed that the NO concentration in the organism changes with the degree of infection by a coronavirus¹⁹. For example, in the early stage of infection with SARS-CoV, activation of epithelial cells, such as those in alveoli, to produce cytokines results in the upregulation of inducible NO synthase, further increasing the NO concentration and thus playing an antiviral role in controlling SARS-CoV infection²⁰. At the same time, excessive NO is converted into peroxynitrite and other substances, damaging human immune cells, promoting SARS-CoV infection, and inhibiting NO production, resulting in a reduction in the NO concentration. Importantly, the NO concentration in SARS patients is significantly different from that in healthy controls. Considering that SARS-CoV-2 and SARS-CoV disease have similarities²¹, we infer that the FeNO concentration in COVID-19 patients may be different from that in healthy controls.

In this work, we demonstrated the FeNO concentration is correlated to COVID-19. However, FeNO alone could not be used to identify COVID-19 patients because of the statistic distribution between FeNO in COVID-19 patients and healthy subjects overlaps. Therefore we developed an ML model by integrating FeNO with other features including age, body mass index (BMI) and body area, which had been reported to be related to FeNO values. This model successfully identified COVID-19 suspects with AUC of 0.982 and sensitivity of 0.917.

2. Materials And Methods

Clinical data

We collected data on age, sex and anamnesis (tuberculosis, chronic obstructive pulmonary disease, lung cancer, allergic rhinitis, pharyngitis, heart disease, diabetes, and stomach disease). This analysis was conducted on a control sample of 149 subjects in the Hubei and Shandong provinces of China from March 2020 to June 2020. Among the subjects, 109 were non-COVID-19 subjects, and 46 were COVID-19 patients from LiYuan Hospital affiliated with TongJi Medical College HuaZhong University of Science and Technology, Wuhan. Among the non-COVID-19 subjects, 93 subjects were healthy without other diseases and these subjects were designated Hea-(1, 2...), 2 subjects had hypertension and were designated Hyp-(1, 2), 1 subject had gastric disorder and was designated Gastr-1, 5 subjects had pharyngitis and were designated Phar-(1, 2...), and 2 subjects had asthma and were designated Asth-(1, 2). The 46 COVID-19 patients had no marked distinctions and were designated Cov-(1, 2...). Fractional exhaled NO samples and data on bodily characteristics were obtained from each subject. All volunteers were informed of the content of the test, and informed consent was obtained from each subject. More detailed information for each subject can be found in the Supplementary Table S1-2. The demographic characteristics of all tested patients and controls in the current study are reviewed in Table 1.

Table 1. Demographics and Baseline Characteristics of COVID-19 Patients and Non-COVID-19 in the Current Study.

Variables	Non-COVID-19				COVID-19 (N=46)
	Healthy Control (N=93)	Asthma (N=2)	Pharyngitis (N=5)	Hypertension /Gastric Disorder(N=3)	
Sex (number) (%)					
Male	38 (40.86)	0 (0.00)	4 (80.00)	0 (0.00)	24 (52.17)
Female	55 (59.14)	2 (100.00)	1 (20.00)	3 (100.00)	22 (47.83)
Age (year)					
Mean±SD	37.5 ± 12.6	28.0 ± 9.9	31.6 ± 7.7	49.0 ± 2.0	61.5 ± 15.3
Range	21.0-65.0	21.0- 35.0	25.0-41.0	47.0-51.0	32.0-89.0
Height (cm)					
Mean±SD	166.66 ± 6.82	158.50 ± 4.95	172.20 ± 8.32	163.33 ± 5.77	165.70 ± 7.06
Range	150.0-183.0	155.0- 162.0	162.0- 183.0	160.0-170.0	148.0- 180.0
Weight (Kg)					
Mean±SD	64.24 ± 11.71	43.00 ± 2.83	66.60 ± 7.44	65.00 ± 13.23	64.86 ± 10.59
Range	41.0-95.0	41.0- 45.0	55.0-75.0	55.0-80.0	40.0-85.0
BMI					
Mean±SD	23.04 ± 3.44	17.18 ± 2.20	22.47 ± 2.17	24.20 ± 3.17	23.56 ± 3.15
Range	15.4-31.2	15.6- 18.7	19.4-24.2	21.5-27.7	16.0-29.4
Body Area (m²)					
Mean±SD	1.69 ± 0.18	1.36 ± 0.01	1.75 ± 0.13	1.68 ± 0.20	1.69 ± 0.16
Range	1.3-2.2	1.4-1.4	1.5-1.9	1.5-1.9	1.3-2.0
FeNO (ppb)					

Mean±SD	15.66 ± 5.07	57.50 ± 3.54	11.60 ± 1.67	35.67 ± 5.69	18.35 ± 7.08
Range	6.0-26.0	55.0-60.0	10.0-14.0	31.0-42.0	7.0-36.0

The FeNO level was measured using a Nano Coulomb Breath Analyser obtained from Wuxi Sunvou Medical Electronic Co., Ltd, China. The Sunvou Nano Coulomb Breath Analyser can measure the NO concentration in the large and small airways through different flow rates. It is mainly based on and fully conforms to the technical standards updated by the American Thoracic Society (ATS) and European Respiratory Society (ERS) in 2005 and ERS in 2017. According to this standard, the determination of NO in the exhaled breath includes the parameters eNO and nNO, where eNO = bronchus NO/flow F + alveolus NO + correction term. To ensure the accuracy of test results and meet the requirements for the use of the instruments, one hour before sampling, the subjects were not allowed to eat, smoke, exercise, or perform lung function tests. Besides, three hours before sampling, the subjects were not allowed to eat broccoli, kale, lettuce, celery, radish, or smoked or pickled food. For detailed information, see www.sunvou.com.

3. Results

Statistical analysis

As shown in Figure 1A, the FeNO value has a significant difference between healthy people and COVID-19 patients. The p value of the student's t test is 0.006. Though there is a significant statistic difference of FeNO between healthy subjects and COVID-19 patients, only analyzing the value of FeNO can't distinguish COVID-19 patients due to the overlaps of the statistical distribution, as shown in Figure 1B. FeNO levels have been reported to be affected with many factors, including age, gender, and body size parameters. Therefore, we considered introducing other features and using the ML algorithm to meet the need to identify COVID-19 patients. ML method has been widely applied in scientific research, and many scientist used it to help with controlling COVID-19 outbreaks including diagnose and anti-SARS-CoV-2 activities prediction etc.²²⁻²³

Machine learning model development

Our classification machine learning framework is schematically illustrated in Figure 2. The approach consists of three parts: feature engineering, ML model training, and the output COVID-19 status probability. Feature engineering is firstly used to evaluate the feature correlation. Pearson correlation coefficient matrices are calculated to identify the positive and negative correlations between pairs of features (e-Figure 1). Body mass index (BMI) and body area are highly related to weight and height. At the same time, age and FeNO show low linear correlations with other features. The low linear correlations and high importance of age and FeNO features indicate that it would better keep them in the ML model construction process. The redundant feature should be removed, which will improve the performance of the ML model. We randomly split the subjects into two sets: the training set and the test set (Figure

2). The training set contained 70 healthy subjects and 34 COVID-19 patients. The test set comprised 10 non-COVID-19 subjects with asthma, gastric disorders, pharyngitis and hypertension, 23 healthy subjects, and 12 COVID-19 patients. The 23 healthy subjects and 12 COVID-19 patients in the test set were selected according to a 2:1 ratio of healthy to COVID-19 subjects, and they made up 25% percent of the total number of healthy and COVID-19 subjects. Due to the small dataset, 10-fold cross-validation (CV) was performed to improve our prediction of the “out-of-sample” error. CV method is usually used in deal with small dataset in ML.²⁴⁻²⁵ We found that the best ML model came from the feature combination of 'Age', 'BMI', 'Body Area', and 'FeNO' (Figure 3, e-Figures 2 and 3). Figure 3A shows the performance of the model in the training cohort, and Figure 3B shows the receiver operating characteristic (ROC) curve of the ML model in the training cohort. This model reached an AUC of 0.996 (95% CI 0.988, 1.000) in the training set. One COVID-19 patient (Cov-37) and one healthy subject (Hea-131) were incorrectly classified. We then tested the model on the unseen 45 subjects of the test cohort. As shown in Figure 3C, only 1 COVID-19 (Cov-13) case, and one healthy case (Hea-78) was incorrectly classified. If we only consider the classification of healthy subjects and COVID-19 patients, the ML model achieved an AUC of 0.982 (95% CI 0.944, 1.000), accuracy of 0.942 (95% CI 0.866, 1.000), sensitivity of 0.917 (95% CI 0.760, 1.000), specificity of 0.957 (95% CI 0.873, 1.000), precision of 0.917 (95% CI 0.760, 1.000) and F1 score of 0.917 in the test cohort.

We further used the model to test the people with other disease (asthma, gastric disorder and hypertension) which has been reported to have abnormal FeNO compared with healthy people.²⁶⁻²⁸ All two asthma, one gastric disorder and two hypertension patients were identified as having COVID-19 with a high output probability. Asthma patients show significantly higher FeNO values than healthy subjects²⁶, and the considerably higher FeNO value makes asthma patients have a high probability of being classified as having COVID-19 under this model. The nitrate vasodilators antihypertensive drugs, such as nitroglycerin and sodium nitroprusside and stomach disease, will affect the FeNO according to prior research^{27, 28}, which cause hypertension and gastric disorder patients being classified as COVID-19 suspects. Five subjects with pharyngitis were all classified as healthy non-COVID-19 subjects. To our knowledge, there is no research showing that pharyngitis is related to FeNO. It is noteworthy that although this machine learning model will screen out people with other diseases affecting the FeNO value, these people account for a very small part of the non-COVID-19 people, so it is reasonable to use this model as a primary screening for COVID-19 suspects.

Discussion

The procedure to screen COVID-19 in our method is as follows. First, connect the data on FeNO, age, sex, body size and anamnesis of each subject. The FeNO detects process is shown in Supplementary Video. Second, the ML model is used to calculate the probability of having COVID-19 for each subject. Subjects with an output probability greater than 0.5 are best admitted to the hospital for further nucleic acid testing, and those with a probability of less than 0.5 are classified as low-risk subjects. The whole test time of this method is less than 2 minutes. The consumables include a plastic filter and a

breath collection bag, and the cost is about 0.3 US dollars. For comparison, the cost of nucleic acid testing consumables is about 10 US dollars. For example, the nucleic acid detections for 10 million people cost in the US \$100 million. In contrast, the cost of our detection method for preliminary screening is only 3 million US dollars. This detection simultaneously improves the detection efficiency and reduces expenses by multiple orders of magnitude.

Limitations and outlook.

The main limitation of the model is that the sample size is small. This limitation was mainly due to the wide regional distribution of the new cases of COVID-19 subjects impeded the data collection of the clinical specimens at an early stage. It is noteworthy that, we mainly focused on early screen for COVID-19 cases that were asymptomatic or had mild symptoms. Thus, all of the patients in this study are mild or asymptomatic cases. FeNO varies with the infection level, for severe cases, the model should be different; however, assessing FeNO in severe cases was not our purpose. On the other hand, if two additional parameters could be added to the model, namely, whether the subject had contact with a COVID-19 patient and whether the subject comes from an epidemic zone, it would improve the screen accuracy further. In the future, we will seek more cooperation to collect more samples. At now, we can only treat the small-datasets machine learning process as carefully as we can to achieve the highest possible accuracy. We believe that as more samples are collected, the accuracy of the model will be significantly improved, which will be helpful to the control of the epidemic.

4. Discussion

In conclusion, we developed a highly accurate method for the rapid screening of potential COVID-19 patients based on FeNO detection. We believe that the proposed ML model, which combines FeNO detection, could be a useful preliminary screening tool that can be used to discover the diseases with abnormal FeNO such as COVID-19 in time and that does not require more complicated technologies.

Declarations

Acknowledgements

We thank Prof. Xin Xu at Fudan University for helpful comments with the theoretical methods. This work was supported by the National Key Research and Development Program of China (No. 2017YFA0204800), the National Natural Science Foundation of China (No. 21525315 and 21721004), and the Fundamental Research Funds of Shandong University (2019HW016 to L.S., 2019GN023 to D.Z., 2019GN021 to G.Q.R., 2019GN111 to T.Y.).

Author Contributions

L.S., and W.Q.D. planned and designed the project. J.M., J.C., J.P.Z., M.R.W. and J.W. designed the breath detection experiment and contributed to collecting the COVID-19 patient data and parts healthy subject

data. J.M. guided W.Z., G.Q.R. and T.Y. to collect parts data of healthy subjects. L.Y., W.Z., L.S. and W.Q.D. contributed to the analysis and interpretation of the data. Y.L. and Z.D. established the ML model. L.Y., W.Z., L.S., and W.Q.D. prepared the draft of the manuscript. All authors reviewed, amended and approved the final version of the manuscript.

Data sharing statement

The authors declare that all the other data supporting the findings of this study are available within the article and its Supplementary Information files and from the corresponding author upon request. All of the information about COVID-19 patients and healthy subjects used in this study are included in the Supplementary Information Supplementary Table 1-3. The code used for training machine learning models and models used in this study are all available at <https://github.com/yimuyang/COVID19-ML>. Implementation of our work is based on the following open source repositories: Anaconda: <https://enterprise-docs.anaconda.com/en/latest/>; Sklearn: <https://scikit-learn.org/stable/>; Jupyter: <https://jupyter.org/documentation>.

Ethics declarations

This study was approved by the Medical Ethics Committee of Liyuan Hospital, Tongji Medical College, Huazhong University of Science and Technology ([2020]IEC(A009)). Informed consent was obtained from all COVID-19 patients and healthy subjects.

Competing interests

Shandong University filed the Chinese Patent Applications No. 202010365814.0 and 202010753816.7 based on the described technologies.

Supplementary materials

The Supplementary Tables and Video can be found in the Supplemental Materials and Multimedia sections of the online article.

References

1. Zhang JJ, Litvinova M, Liang YX, Wang Y, Wang W, Zhao S, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* 2020;368:1481-1486.
2. Tian HY, Liu YH, Li YD, Wu CH, Chen B, Kraemer MUG, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 2020;368:638-642.
3. Hao XJ, Cheng SS, Wu DG, Wu T, Lin X, Wang C. Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. *Nature* 2020;584:420-424.

4. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dorner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* 2020;368:619.
5. Leung NHL, Chu DKW, Shiu EYC, Chan KH, Mcdevitt JJ, Hau BJP, et al. Respiratory virus shedding in exhaled breath and efficacy of face masks. *Nat Med* 2020;26:676-680.
6. Zhang K, Liu XH, Shen J, Li ZH, Sang Y, Wu XW, et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* 2020;181:1423-1433.
7. Mei XY, Lee HC, Diao KY, Huang MQ, Lin B, Liu CY, et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat Med* 2020;26:1224-1228.
8. Jia JS, Lu X, Yuan Y, Xu G, Jia JM, Christakis NA. Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* 2020;582:389-394.
9. Zhou JM, Huang ZA, Kumar U, Chen DY. Review of recent developments in determining volatile organic compounds in exhaled breath as biomarkers for lung cancer diagnosis. Review of recent developments in determining volatile organic compounds in exhaled breath as biomarkers for lung cancer diagnosis. *Anal Chim Acta* 2017;996:1-9.
10. Gustafsson LE, Leone AM, Persson MG, Wiklund NP, Moncada S. Endogenous nitric oxide is present in the exhaled air of rabbits, guinea pigs and humans. *Biochem Bioph Res Co* 1991;181:852-857.
11. Tiev KP, Cabane J, Aubourg F, Kettaneh A, Ziani M, Mouthon L, et al. Severity of scleroderma lung disease is related to alveolar concentration of nitric oxide. *Eur Respir J* 2007;30:26-30.
12. Moodley YP, Chetty R, Lalloo UG. Nitric oxide levels in exhaled air and inducible nitric oxide synthase immunolocalisation in pulmonary sarcoidosis. *Eur Respir J* 1999;14:822-827.
13. Tiev KP, Hua-Huy T, Kettaneh A, Allanore Y, Le-Dong N-N, Duong-Quy S, et al. Alveolar concentration of nitric oxide predicts pulmonary function deterioration in scleroderma. *Thorax* 2012;67:157-163.
14. Hua-Huy T, Le-Dong N-N, Duong-Quy S, Bei YH, Rivière S, Tiev K-P, et al. Increased exhaled nitric oxide precedes lung fibrosis in two murine models of systemic sclerosis. *J Breath Res* 2015;9:036007.
15. Nilsson KF, Gustafsson LE, Adding LC, Linnarsson D, Agvald P. Increase in exhaled nitric oxide and protective role of the nitric oxide system in experimental pulmonary embolism. *Brit J Pharmacol* 2007;150:494-501.
16. Wu XZ, Huang XZ, Ke PF, Zheng SB, Ma Y, Zhuang JH, et al. Detection of serum MDA, NO and iNOS in patients with SARS and its clinical significance. *Chin J Pathophysiol* 2004;20:1916-1918.

17. Fernando H, Cardet JC, Chung KF, Diver S, Ferreira DS, Fitzpatrick A, et al. Management of severe asthma: a European Respiratory Society/American Thoracic Society guideline. *Eur Respir J* 2020;55:1900588.
18. Enrico H, Carpagnano GE, Favero E, Guida G, Maniscalco M, Motta A, et al. Fractional Exhaled Nitric Oxide (FENO) in the management of asthma: a position paper of the Italian Respiratory Society (SIP/IRS) and Italian Society of Allergy, Asthma and Clinical Immunology (SIAAIC). *Multidiscip Respir Med* 2020;15:36.
19. Mashir A, Paschke KM, Duin DV, Shrestha NK, Laskowski D, Storer MK et al. Effect of the influenza A (H₁N₁) live attenuated intranasal vaccine on nitric oxide (FE_{NO}) and other volatiles in exhaled breath. *J Breath Res* 2011;5:037107.
20. Zhuang JH, Huang XZ, Xu N, Ou AH, Lin L. Dynamic observation of serum NO, iNOS and SOD of patients with severe acute respiratory syndrome. *Chin J Microcircul* 2004;14:4-5.
21. Jung K, Gurnani A, Renukaradhya GJ, Saif LJ. Nitric oxide is elicited and inhibits viral replication in pigs infected with porcine respiratory coronavirus but not porcine reproductive and respiratory syndrome virus. *Vet Immunol Immunop* 2010;136:335-339.
22. Kc GB, Bocci G, Verma S, Hassan MM, Holmes J, Yang JJ, et al. A machine learning platform to estimate anti-SARS-CoV-2 activities. *Nat Mach Intell* 2021;3:527-535.
23. Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell* 2021;3:199-217.
24. Zhang Y, Ling C. A strategy to apply machine learning to small datasets in materials science. *npj Comput Mater* 2018;4.
25. Wexler RB, Martirez JMP, Rappe AM. Chemical Pressure-Driven Enhancement of the Hydrogen Evolving Activity of Ni₂P from Nonmetal Surface Doping Interpreted via Machine Learning. *J Am Chem Soc* 2018;140:4678-4683.
26. Dweik RA, Boggs PB, Erzurum SC, Lrvin CG, Leigh MW, Lundberg JO, et al. An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FE_{NO}) for clinical applications. *Am J Respir Crit Care Med* 2011;184:602-615.
27. Marczin N, Riedel B, Royston D, Yacoub M. Intravenous nitrate vasodilators and exhaled nitric oxide. *Lancet* 1997;349:1742.
28. Som S, Banik GD, Maity A, Chauduri S, Pradhan M. Exhaled nitric oxide as a potential marker for detecting non-ulcer dyspepsia and peptic ulcer disease. *J Breath Res* 2018;12:026005.

Figures

Figure 1

FeNO of COVID-19 patients and healthy subject. (A) There is a significant statistic difference of FeNO between COVID-19 patients and healthy subjects. The Student's t-test p-value is <0.05 . (B) There is significant overlaps between FeNO in healthy subjects and COVID-19 subjects indicating that FeNO alone couldn't be used to identify COVID-19 patients.

Figure 2

Illustration of the ML model framework. First, we measured the FeNO levels of each subject and obtained basic information, including age, height, weight, and anamnesis. Then, we trained the ML model on the training cohorts using a 10-fold CV method. The ML classifier was then validated in an independent test cohort. Finally, we use the trained ML model to screen COVID-19 suspects. When the predicted possibility was greater than 0.5, the subject was classified as a COVID-19 patient or as a subject with another

disease who would be better categorized into the high-risk group, and a further nucleic acid test would be recommend. If the possibility was less than 0.5, then these subjects were treated as low-risk subjects.

Figure 3

Receiver operating characteristic (ROC) curve and performance of the ML model. (A) The performance of the model in the training cohort. (B) ROC curve in the training cohort. AUC is 0.996. (C) The performance of the model in the test cohort. The test cohort, including COVID-19 patients, healthy subjects, and subjects with asthma, gastric disorder, hypertension, and pharyngitis. (D) ROC curve of test cohort. Only one healthy subject was misclassified. For the classification of healthy subjects and COVID-19 patients, the ML model achieved an AUC of 0.982, accuracy of 0.942, sensitivity of 0.917, specificity of 0.957, precision of 0.917, and F1 score of 0.917.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFeNOEBioMedicine.docx](#)
- [SupplementaryTable2.pdf](#)
- [SupplementaryTable3.pdf](#)
- [VideoOperation.mp4](#)