Correspondence

DOI: 10.5582/bst.2022.01476

Detecting latent tuberculosis infection with a breath test using mass spectrometer: A pilot cross-sectional study

Liang Fu[1,§], Yong Feng[2,§], Tantan Ren[1,§], Min Yang[1], Qianting Yang[3], Yi Lin[1], Hui Zeng[4], Jiaohong Zhang[5], Lei Liu[1], Qingyun Li[2], Mengqi He[2], Peize Zhang[1,*], Haibin Chen[2,*], Guofang Deng[1,*]

[1] Division Two of Pulmonary Diseases Department, The Third People's Hospital of Shenzhen, National clinical research center for infectious disease, Southern University of Science and Technology, Shenzhen, Guangdong, China;
[2] Breax Laboratory, PCAB Research Center of Breath and Metabolism, Beijing, China;
[3] Guangdong Key Lab for Diagnosis & Treatment of Emerging Infectious Disease, The Third People's Hospital of Shenzhen, National clinical research center for infectious disease, Southern University of Science and Technology, Shenzhen, Guangdong, China;
[4] Medical Examination Department, The Third People's Hospital of Shenzhen, National clinical research center for infectious disease, Southern University of Science and Technology, Shenzhen, Guangdong, China;
[5] Pulmonary Diseases Out-patient Department, The Third People's Hospital of Shenzhen, National clinical research center for infectious disease, Southern University of Science and Technology, Shenzhen, Guangdong, China.

SUMMARY *Mycobacterium tuberculosis* (M.tb) infects a quarter of the world's population and may progress to active tuberculosis (ATB). There is no gold standard for diagnosing latent tuberculosis infection (LTBI). Some immunodiagnostic tests are recommended to detect LTBI but can not distinguish ATB from LTBI. The breath test is useful for diagnosing ATB compared to healthy subjects but was never studied for LTBI. This proof-of-concept study (Chinese Clinical Trials Registry number: ChiCTR2200058346) was the first to explore a novel, rapid, and simple LTBI detection method *via* breath test on high-pressure photon ionization time-of-flight mass spectrometry (HPPI-TOFMS). The case group of LTBI subjects (n = 185) and the control group (n = 250), which included ATB subgroup (n = 121) and healthy control (HC) subgroup (n = 129), were enrolled. The LTBI detection model indicated that a breath test *via* HPPI-TOFMS could distinguish LTBI from the control with a sensitivity of 80.0% (95% CI: 67.6%, 92.4%) and a specificity of 80.8% (95% CI: 71.8%, 89.9%). Nevertheless, further intensive studies with a larger sample size are required for clinical application.

Keywords latent tuberculosis infection, tuberculosis, diagnosis, volatile organic compounds, breath

It is estimated that *Mycobacterium tuberculosis* (M.tb) infected a quarter of the world's population (*1*). Latent tuberculosis infection (LTBI) constitutes a broad spectrum of infection states that differ by the degree of pathogen replication, host immune response, and inflammation (*2*). Approximately 5-10% of those with LTBI will progress to active tuberculosis (ATB) (*3*). WHO recommends immunodiagnostic tests for LTBI detection, either a tuberculin skin test (TST) or interferon-gamma (IFN-γ) release assays (IGRAs) (*4*). However, these tests are not precise enough. In certain situations, TB exposure can be used as a surrogate for LTBI (*5*). Furthermore, TST and IGRAs can not differentiate LTBI from ATB (*6*). Thus, a more precise tool is urgently needed for the consecutive management of uninfected status, LTBI, and ATB.

Recent studies indicate that breathomics may be a useful rule-in or rule-out tool for diagnosing ATB (*7*), which uncovers the host-pathogen interaction *via* comprehensive exhaled breath analysis. Breathomics may hold promise to distinguish healthy subjects, LTBI and ATB (*8*) if a breath test can find the trace and tell the difference of M.tb in consecutive states in the host (*9*). High-pressure photon ionization time-of-flight mass spectrometry (HPPI-TOFMS) is designed and developed by our team, which can directly detect volatile organic compounds (VOCs) in exhaled breath (*10*). In our previous studies, this breath detection platform has been verified in lung cancer (*11,12*), esophagus cancer (*13*), and Corona Virus Disease 2019 (COVID-19) (*14*). In this study, we explored the use of this novel, rapid, simple, and inexpensive breath test to detect LTBI.

We conducted a cross-sectional study (Chinese Clinical Trials Registry number: ChiCTR2200058346) in which a breath sample was collected from 435 participants with informed consent signed at the Third

People's Hospital of Shenzhen in Shenzhen, China, between March 2020 and November 2022. This study (No.2022017) was approved by the Ethics Committee of the Third People's Hospital of Shenzhen. The study population consisted of three main groups. The LTBI group included the participants who were contacts of ATB patients and had a positive IGRA result, with normal chest imaging and no evidence of ATB (n = 185). The control group consisted of two main subcategories: 1) ATB group (n = 121): ATB subjects in whom M.tb culture or GeneXpert TB-DNA was positive, and chest imaging was suggestive of ATB; 2) healthy control (HC) group (n = 129): healthy subjects who came for physical examination and had no known contacts with ATB patients, with a negative IGRA result and a normal chest imaging. All participants were enrolled in the queue. Because of the selection offset, the enrolled HC is younger than other groups. The age is significantly ($p < 0.001$) different between the case and control groups, whose median ages are 41 and 28, respectively. There is no significant difference ($p = 0.397$) in gender between the case and control groups.

Breath samples were collected using a predefined protocol and tested in our developed HPPI-TOFMS within twenty-four hours (*10*). The sampling apparatus comprised a disposable gas nipple and a sampling bag made of polyether-ether-ketone (PEEK). In this study, we set standard sampling demands and protocols to minimize the influence of the daily diet. Firstly, we conducted sampling at a second visit if he/she was an inpatient and informed the participants to prepare for sampling in advance: no smoking, alcohol, or diets within an hour before sampling. Secondly, participants were required to rinse their mouths with purified water instantly before sampling. Thirdly, all samples must be collected in the same environment, which could minimize the effects of environmental facts. With a deep nasal inhalation, participants completely exhaled the air into the sampling bag with over 1.2 L volume.

All the enrolled participants were randomly split into three groups: 50% of them for model construction, 20% for internal validation, and the remaining 30% for model-blinded testing. Thus, 92 LTBI patients and 123 controls were randomly selected as the discovery data set for Random Forest (RF) (*15*) based LTBI detection model training, which was evaluated on an internal validation dataset (37 LTBI patients and 51 controls) and blinded test dataset (56 LTBI patients and 76 controls).

To transfer the mass spectrum data produced by HPPI-TOFMS, noise-reducing, and baseline correction were applied *via* anti-symmetric wavelet transformation after mass calibration. Then, the area of the strongest peak in the range of (x – 0.1, x + 0.1) was calculated as the feature of VOC with *m/z* close to x. In this way, a mass spectrum would be transferred into 1500 ion features in the *m/z* range of (20, 320). To avoid over-fitting in model training, the features without significant difference ($p > 0.05$) and features with high correlation coefficient but low peak area were excluded. Then, the model-based feature selection was executed based on training and validation datasets, and the top ten VOC ions were selected according to the ranked feature importance.

In this study, the receiver operating characteristic (ROC) curve analysis was implemented. The sensitivity (SEN), specificity (SPE), positive prediction value (PPV), negative prediction value (NPV), accuracy (ACC), area under the ROC curve (AUC), and their relative 95% confidence interval (CI) were calculated to evaluate the performance of LTBI detection model.

As shown in Table 1 and Figures 1a and 1b, with the cut-off value of 0.5 (over 0.5 is considered LTBI), the LTBI detection model achieved good discrimination performance with an SEN and an SPE of 78.4 (95% CI: 63.4%, 94.3%) and 84.3% (95% CI: 74.3%, 94.3%) in the internal validation dataset. In the test dataset, the model performance slightly dropped, with the AUC decreasing from 0.913 (95% CI: 0.854, 0.972) to 0.867 (95% CI: 0.809, 0.925). The SEN and SPE achieved in the test dataset were 80.4% (95% CI: 68.7%, 92.0%) and 80.3% (95% CI: 71.3%, 89.2%). Since there are two subgroups in the controls, we also evaluated the performances in discriminating LTBI with ATB and HC, respectively. The LTBI model performed better in discriminating LTBI and HC with an AUC of 0.952 (95% CI: 0.909, 0.995) than in discriminating LTBI and ATB with an AUC of 0.777 (95% CI: 0.692, 0.861).

To evaluate the selected VOC ions in LTBI detection, we trained the LTBI detection model on each single VOC ion and evaluated it in the test dataset. The ROC curve in Figure 1c demonstrates that the discrimination of a single VOC ion is also good but limited ($0.64 < AUC < 0.80$), which is much inferior to the performance (AUC = 0.867) of the combination of all ten VOCs. It implies that the panel of VOC ions is the basis for breathomics-based LTBI detection. Figure 1d illustrates the patterns of these ten VOC ions that are visually different in ATB, LTBI, and HC groups. Figure 1e illustrated there are significant differences ($p < 0.001$) among LTBI, HC, and ATB groups for almost all ten VOC ions, except for the VOC with *m/z* of 129 between LTBI and ATB ($p = 0.589$). Since the TOF mass spectrometer can only confirm the *m/z* of detected VOCs, we need to infer the possible chemicals of these LTBI related VOC ions based on their *m/z* (121, 145, 129, 135, 105, 130, 117, 93, 77, 109), peak area distribution, other published potential biomarkers, and the human breathomics database (*16*). The VOC ions with *m/z* of 145, 135, 130, and 109 should be 1,4-dimethyl-indol, benzothiazole, 2-ethyl-1-hexanol, and 4-aminophenol, respectively. The VOC ions with *m/z* of 145, 135, 121, 129, 117, and 77 would be the protonated ion of octanoic acid, 1-methyl-4-(1-methyl ethyl)-, 4-ethyltoluene, naphthalene, 2-methyl propyl acetate, and carbon-disulfide, respectively. The VOC

Table 1. Qualitative evaluation of RF-based LTBI detection model in validation and test sets

Data set (*n*)	SEN (%)	SPE (%)	PPV (%)	NPV (%)	ACC (%)	AUC
Validation (88)						
LTBI (37) *vs*. Control (51)	78.4 (63.4, 93.4)	84.3 (74.3, 94.3)	78.4 (67.1, 89.7)	84.3 (72.6, 96.0)	81.8 (73.8, 89.9)	0.913 (0.854, 0.972)
LTBI (37) *vs*.. HC (26)	78.4 (63.4, 93.4)	100 (100, 100)	100 (100, 100)	76.5 (61.0, 91.9)	87.3 (79.1, 95.5)	0.975 (0.937, 1.000)
LTBI (37) *vs*. ATB (25)	78.4 (63.4, 93.4)	68.0 (49.7, 86.3)	78.4 (62.2, 94.5)	68.0 (53.0, 83.0)	74.2 (63.3, 85.1)	0.848 (0.758, 0.937)
Test (132)						
LTBI (56) *vs*. Control (76)	80.4 (68.7, 92.0)	80.3 (71.3, 89.2)	75.0 (65.0, 85.0)	84.7 (75.6, 93.8)	80.3 (73.5, 87.1)	0.867 (0.809, 0.925)
LTBI (56) *vs*. HC (39)	80.4 (68.7, 92.0)	97.4 (92.5, 100)	97.8 (93.7, 100)	77.6 (65.5, 89.6)	87.4 (80.7, 94.0)	0.952 (0.909, 0.995)
LTBI (6) *vs*. ATB (37)	80.4 (68.7, 92.0)	62.2 (46.5, 77.8)	76.3 (62.0, 90.6)	67.6 (55.7, 79.6)	73.1 (64.1, 82.1)	0.777 (0.692, 0.861)

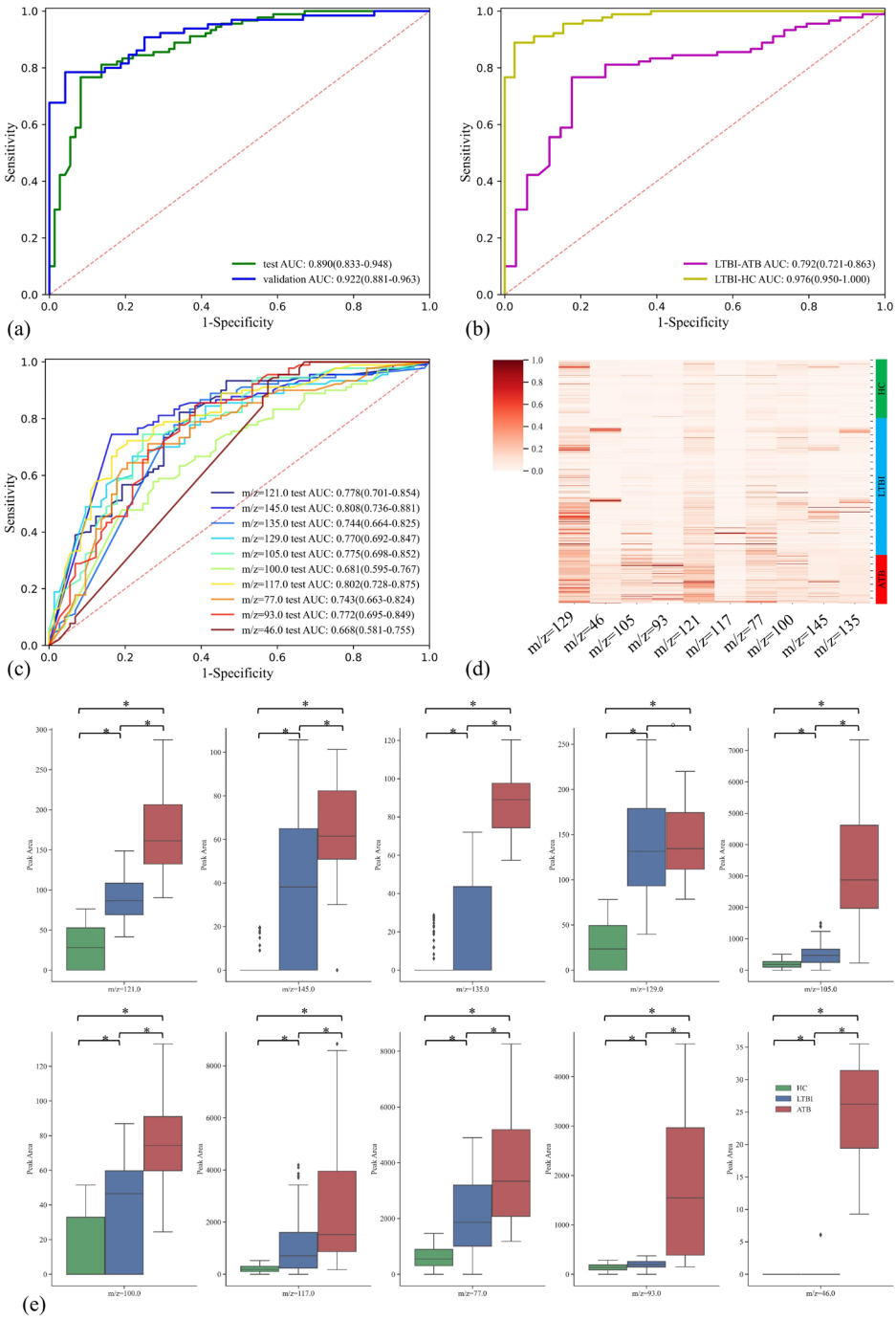

Figure 1. The ROC comparisons and the top ten VOC ions of the developed LTBI detection model. ROC comparison in validation and test datasets **(a)** and test set **(b)**. The LTBI detection power of the top ten selected VOC ions in the test dataset **(c)**. The heatmap of peak area distribution in all HC, LTBI, and ATB samples **(d)**. The box plot of peak area in all HC, LTBI, and ATB samples. * and ° represent significant or insignificant differences between the two groups **(e)**.

ions with *m/z* of 105 should be the combination ion of H_3O^+ and butanal, 3-methyl-. The VOC ion with *m/z* of 93 would be pyridine, 3-methyl- or the protonated ion of toluene. Among these related chemicals, octanoic acid (*m/z* = 144), 2-ethyl-1-hexanol (*m/z* = 130), and toluene (*m/z* = 92) were proven related to pulmonary tuberculosis (PTB) (*17*). 1,4-Dimethyl-indol (*m/z* = 145), naphthalene (*m/z* = 128), and 1-methyl-4-(1-methyl ethyl)- (*m/z* = 134) were reported as the VOC biomarkers of PTB *via* gas chromatography-mass spectrometer (GC-MS) detection in Machel Phillips's study (*18*). Their chemical would be closely related to the metabolites of M.tb or the metabolic changes accused by M.tb. However, these chemicals are not completely confirmed, and their metabolic mechanism is poorly understood.

Our study had several strengths. All breath samples of ATB subjects were taken before anti-tuberculosis treatment so that drugs did not influence our results. For ATB, we used GeneXpert as a complementary diagnostic test to verify a true-negative result in subjects found to be negative with a sputum test (*19*). Healthy subjects were sampled in the same site as LTBI and ATB subjects, which raised no concerns regarding possible geographical bias affecting the results. Since there is no gold standard for LTBI, choosing IGRA as the diagnostic test may miss LTBI people with undetectable IFN-γ response (*6*). Thus, we confirmed the LTBI and HC groups based on the IGRA results and TB exposure. For instance, we excluded close contacts with a negative IGRA from the LTBI group. On the other hand, the stronger standard for LTBI and ATB also limited the applicability and extensibility of the developed LTBI detection model. Other limitations include: 1) these VOC ions were not completed qualitatively identified, although we have extrapolated the possible chemicals based on their formula weight and published biomarkers; 2) there may be selection bias in selecting HC; 3) Participants with other respiratory diseases were not enrolled in this study, which is meaningful for the method evaluation in related application scenarios, but is unfavorable for discovering the potential biomarkers for tuberculosis infection. We expect to optimize our research and conduct more profound studies in LTBI-related VOCs in the future.

In summary, this study provided a potential noninvasive, simple, and fast method for LTBI diagnosis. The preliminary proof-of-concept results indicate that a breath test *via* HPPI-TOFMS may be a valuable tool to distinguish LTBI from HC and ATB, which achieved an accuracy of 80.3% (95% CI: 73.5%, 87.1%) and an AUC of 0.867 (95% CI: 0.809, 0.925). Before the clinical application of breath test-based LTBI diagnosis technologies, more extensive cohort studies are required.

Acknowledgements

We thank all the physicians and assistants that participated in this study and enrolled patients. We thank Jun Cheng for assistance in study design from National Center for Tuberculosis Control and Prevention of the Chinese Center for Disease Control and Prevention.

Funding: This work was supported by the National Natural Science Foundation of China (No. 82070016), the National Key Research and Development Plan (No. 2020YFA0907200, 2019YFC0840602), the Guangdong Foundation for Basic and Applied Basic Research (No. 2019B1515120041), the Guangdong Provincial Clinical Research Center for Tuberculosis Project (No. 2020B1111170014), the Shenzhen Scientific and Technological Foundation (No. KCXFZ202002011007083, JCYJ20180228162112889), Summit Plan for Foshan High-level Hospital Construction (No. FSSYKF-2020001) and Project funded by Shenzhen Third People's Hospital (No. G2022051).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Houben RM, Dodd PJ. The global burden of latent tuberculosis infection: A re-estimation using mathematical modelling. Plos Med. 2016; 13:e1002152.
2. Esmail H, Barry CE, Young DB, Wilkinson RJ. The ongoing challenge of latent tuberculosis. Philos Trans R Soc Lond B Biol Sci. 2014; 369:20130437.
3. Behr MA, Edelstein PH, Ramakrishnan L. Revisiting the timetable of tuberculosis. BMJ. 2018; 362:k2738.
4. WHO. Latent tuberculosis infection: updated and consolidated guidelines for programmatic management. Geneva: World Health Organization. 2018.
5. Stout JE, Wu Y, Ho CS, Pettit AC, Feng PJ, Katz DJ, Ghosh S, Venkatappa T, Luo R. Evaluating latent tuberculosis infection diagnostics using latent class analysis. Thorax. 2018; 73:1062-1070.
6. Heyckendorf J, Georghiou SB, Frahm N, Heinrich N, Kontsevaya I, Reimann M, Holtzman D, Imperial M, Cirillo DM, Gillespie SH, Ruhwald M. Tuberculosis treatment monitoring and outcome measures: New interest and new strategies. Clin Microbiol Rev. 2022; 35:e22721.
7. Nakhleh MK, Jeries R, Gharra A, Binder A, Broza YY, Pascoe M, Dheda K, Haick H. Detecting active pulmonary tuberculosis with a breath test using nanomaterial-based sensors. Eur Respir J. 2014; 43:1522-1525.
8. Ibrahim W, Carr L, Cordell R, Wilde MJ, Salman D, Monks PS, Thomas P, Brightling CE, Siddiqui S, Greening NJ. Breathomics for the clinician: The use of volatile organic compounds in respiratory diseases. Thorax. 2021; 76:514-521.
9. Bruderer T, Gaisl T, Gaugg MT, Nowak N, Streckenbach B, Müller S, Moeller A, Kohler M, Zenobi R. On-line analysis of exhaled breath. Chem Rev. 2019; 119:10803-10828.
10. Wang Y, Jiang J, Hua L, Hou K, Xie Y, Chen P, Liu W, Li Q, Wang S, Li H. High-pressure photon ionization

source for TOFMS and its application for online breath analysis. Anal Chem. 2016; 88:9047-9055.
11. Meng S, Li Q, Zhou. Z, Fu L. Assessment of an exhaled breath test using high-pressure photon ionization time-of-flight mass spectrometry to detect lung cancer. JAMA Netw Open. 2021; 4:e213486.
12. Wang P, Huang Q, Meng S, Mu T, Liu Z, He M, Li Q, Zhao S, Wang S, Qiu M. Identification of lung cancer breath biomarkers based on perioperative breathomics testing: A prospective observational study. EClinicalMedicine. 2022; 47:101384.
13. Huang Q, Wang S, Li Q, Wang P, Li J, Meng S, Li H, Wu H, Qi Y, Li X, Yang Y, Zhao S, Qiu M. Assessment of breathomics testing using high-pressure photon ionization time-of-flight mass spectrometry to detect esophageal cancer. JAMA Netw Open. 2021; 4:e2127042.
14. Zhang P, Ren T, Chen H, Li Q, He M, Feng Y, Wang L, Huang T, Yuan J, Deng G, Lu H. A feasibility study of Covid-19 detection using breath analysis by high-pressure photon ionization time-of-flight mass spectrometry. J Breath Res. 2022; 16. doi: 10.1088/1752-7163/ac8ea1.
15. Breiman L. Random forests. Mach Learn. 2001; 45:5-32.
16. Kuo TC, Tan CE, Wang SY, Lin OA, Su BH, Hsu MT, Lin J, Cheng YY, Chen CS, Yang YC, Chen KH, Lin SW, Ho CC, Kuo CH, Tseng YJ. Human breathomics database. Database (Oxford). 2020; 2020:baz139.
17. Vishinkin R, Busool R, Mansour E, Fish F, Esmail A, Kumar P, Gharaa A, Cancilla JC, Torrecilla JS, Skenders G, Leja M, Dheda K, Singh S, Haick H. Profiles of volatile biomarkers detect tuberculosis from skin. Adv Sci. 2021; 8:e2100235.
18. Phillips M, Cataneo RN, Condos R, Ring EG, Greenberg J, La Bombardi V, Munawar MI, Tietje O. Volatile biomarkers of pulmonary tuberculosis in the breath. Tuberculosis (Edinb). 2007; 87:44-52.
19. Theron G, Peter J, Meldau R, Khalfey H, Gina P, Matinyena B, Lenders L, Calligaro G, Allwood B, Symons G, Govender U, Setshedi M, Dheda K. Accuracy and impact of Xpert MTB/RIF for the diagnosis of smear-negative or sputum-scarce tuberculosis using bronchoalveolar lavage fluid. Thorax. 2013; 68:1043-1051.

Received November 14, 2022; Revised December 6, 2022; Accepted December 25, 2022.

[§]These authors contributed equally to this work.
*Address correspondence to:
Peize Zhang and Guofang Deng, Division Two of Pulmonary Diseases Department, The Third People's Hospital of Shenzhen, National clinical research center for infectious disease, Southern University of Science and Technology, Shenzhen 518112, Guangdong, China.
E-mail: 82880246@qq.com (PZ); jxxk1035@yeah.net (GD)

Haibin Chen, Breax Laboratory, PCAB Research Center of Breath and Metabolism Beijing, 100074, China.
E-mail: haibinc@hotmail.com (HC)

Released online in J-STAGE as advance publication January 2, 2023.