Study Protocol and Statistical Analysis Plan

Official Title: Assessment of an Exhaled Breath Test Using High-Pressure Photon Ionization Time-of-Flight Mass Spectrometry to Detect Bronchiectasis

NCT number: NCT05293314

Document date: July 17, 2023

Study Description

To investigate whether the breath test is able to detect bronchiectasis using breathomics. This study was conducted with a prospective specimen collection, evaluator-blinded, case-controlled clinical study designed to evaluate the accuracy of breathomics to diagnosis of bronchiectasis in adults.

Study Design

Study Population:

The age of all subjects were all older than 18. The study population was recruited from Shanghai pulmonary hospital. Diagnosis of bronchiectasis was performed using chest HRCT scans in suspected patients with coughing and expectoration, or long durations of haemoptysis. Healthy participants were as control. Participants with Interstitial lung disease or Sarcoidosis were as unhealthy control.

Criteria:

Inclusion Criteria:

Patients were recruited according to the following inclusion criteria: (1) Age>18 years; (2) the diagnosis of BE was according to European Respiratory Society guidelines for the management of adult bronchiectasis; (3) Willing to join in and sign the informed consent form.

Healthy control subjects were recruited according to the following inclusion criteria:(1) Age>18 years; (2) No history of any lung disease (according to pulmonary imaging and physical examinations); (3) Willing to join in and sign the informed consent form.

Unhealthy control subjects were recruited according to the following inclusion criteria: (1) Age>18 years; (2) The diagnosis of ILD was according to HRCT and clinical symptoms by two experts' consensus; (3) The diagnosis of Sarcoidosis (SA) was according to American Thoracic Society Clinical Practice Guideline;(4) Willing to join in and sign the informed consent form.

Exclusion Criteria:

(1) Patients combined with serious comorbidities (chronic renal failure, hepatic disease, etc.); (2) Patients who are diagnosed with asthma, pneumonia, malignant tumor, and COPD; (3) Women who are pregnant or preparing for pregnancy or breastfeeding; (4) Participated in other clinical trials within three months; (5) Refused to join in and sign the informed consent form.

Exhaled Breath collection:

All exhaled breath samples were collected by trained investigators following the same protocol. All participants were asked to fast for eight hours and rinsed their mouths with clean water before collecting breath samples. To reduce potential confounding factors, all participants were asked to not ingest spicy food, alcohol, or coffee the night before exhaled breath collection. Participants were asked to perform a single deep nasal inhalation followed by complete exhalation via their mouth into an airbag with a volume of 1.2L made of polyether-ether-ketone (PEEK) bags until the bags were full of exhaled air. For each day of sample collection, ambient air was collected to make sure the environment was not polluted. The PEEK airbags were carried out to the laboratory within one week and analyzed by HPPI-

TOFMS analysis.

Clinical Data Collection:

The clinical information of hospitalized patients was collected through a medical record system and questionnaire survey, including chronic medical history, smoking history, chest HRCT imaging data, annual acute attack frequency, disease severity (BSI score), etc.

The information of health examiners is obtained through a survey questionnaire, including age, chronic medical history, smoking history, family history of tumors, and chest HRCT.

The imaging data were obtained from the Shanghai Shanghai pulmonary hospital. And all clinical data were extracted from the medical system by two individual researchers.

VOCs detection:

On the HPPI-TOFMS detection platform, soft ionization was achieved by a commercial VUV-Kr lamp with a photon energy of 10.6 eV to predominantly produce radical cations (M+) which are formed as $M + h\gamma \rightarrow M+ + e$, which can ionize most VOCs in the breath sample. Meanwhile, some protonated ions are also produced. The VOCs with mass-charge ratio(m/z) in the range of [20, 320] were detected on HPPI-TOFMS. In this study, the area of the strongest peak in the range of [x-0.5, x+0.5) was calculated, normalized, and regarded as the feature of VOC with m/z of x. Thus, three hundred kinds of VOC ions features were extracted via breath detection in HPPI-TOFMS.

VOCs analysis:

To select valuable breathomics features and avoid model over-fitting, a combined feature selection method was executed for breathomics features. Firstly, the breathomics features with no significant differences ($p > 0.05$) between BE and controls were excluded. Then, the VOC ions with low intensity but highly correlated with other selected VOC ions (correlation coefficient >0.9) among all discovery samples were excluded. At last, a random forest (RF) model was constructed on the discovery data set, and the ten most important VOC ions were selected based on the feature coefficient in RF model.

In this study, we constructed BE detection models via six popular ML methods, Random Forests (RF), Logistic regression (LR), Xtreme Gradient Boosting (XGB), K-Nearest Neighbor (KNN), Decision tree (DT) and the soft ensemble algorithm of RF, XGB, and LR, in discovery data set. To evaluate the performance of VOCs based BE detection models, the model detection results were evaluated and compared with the HRCT confirmed diagnosis results in the blinded test set.

Statistical analysis

In this study, descriptive statistics were described as frequencies (percentages) for categorical variables or medians (interquartile range [IQR]) for continuous variables. The Mann-Whitney U test for continuous variables and the chi-square test for categorical variables were calculated among the different groups. Sensitivity (SEN), specificity (SPE), and accuracy (ACC) were calculated to evaluate the BE detection performance. A receiver operating characteristic (ROC) curve was plotted, and the area under the ROC curve (AUC) was calculated to assess the overall performance of the VOCs based BE detection model. All statistical analyses were

performed by SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) and the Origin software (version 2018). A p-value<0.05 was considered statistically significant. All tests were two tailed.

Study plan

To construct an easy, fast, and non-invasive VOCs-based bronchiectasis detection mode

This study plans to recruit 200 Bronchiectasis patients diagnosed by HRCT from the Department of Respiratory and Critical Care Medicine of Shanghai Pulmonary Hospital and 200 healthy participants from July,2022 to March,2023.

The exhaled gas of admitted patients was collected on the day of admission, that of outpatients on the day of outpatient service, and that of examinees on the day of physical examination. After the exhaled gas was collected, VOCs were detected by mass spectrometry according to the standard processing flow, and the full spectrum of exhaled VOCs of Bronchiectasis patients and healthy people were analyzed to explore the characteristics of exhaled VOCs of Bronchiectasis patients and healthy people, Analyze the correlation between VOCs and clinical disease severity and prognosis (number of acute exacerbations) characteristics. The data of 400 cases were randomly assigned in a 5:2:3 ratio to the discovery cohort, internal validation, and blinded test cohorts. The discovery and internal validation sets were used for potential VOC biomarkers selection and diagnostic model construction.

To evaluate the VOCs-based bronchiectasis detection mode

Participants were recruited as mentioned above. The exhaled gas was collected and analysed as previously described. The data of 400 cases were randomly assigned in a 5:2:3 ratio to the discovery cohort, internal validation, and blinded test cohorts. The blinded test set was used for model evaluation.