Evaluating the diagnostic accuracy of a ctDNA methylation classifier for incidental lung nodules: protocol for a prospective, observational, and multicenter clinical trial of 10,560 cases

Wenhua Liang[1#], Dan Liu[2#], Min Li[3#], Wei Wang[4#], Zheng Qin[5], Jian Zhang[6], Yong Zhang[7], Yang Hu[8], Hairong Bao[9], Yi Xiang[10], Bo Wang[11], Jing Wu[11], Jianyu Sun[11], Chengping Hu[3], Xianwei Ye[12], Xiangyan Zhang[12], Wei Xiao[13], Chunmei Yun[14], Dejun Sun[14], Wei Wang[5], Ning Chang[6], Yunhui Zhang[15], Jianping Zhao[16], Xin Zhang[7], Jinfu Xu[8], Di Wu[17], Xiaoju Liu[9], Yubiao Guo[18], Qichuan Zhang[19], Wei Zhang[20], Lan Yang[21], Zhanqing Li[22], Xiaoju Zhang[23], Baohui Han[24], Zhaohui Tong[25], Jianxing He[1], Jieming Qu[10], Jian-Bing Fan[11,26], Nanshan Zhong[27]

[1]Department of Thoracic Surgery/Oncology, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou Institute of Respiratory Disease & Health, China State Key Laboratory and National Clinical Research Center for Respiratory Disease, Guangzhou, China; [2]Department of Respiratory Medicine, West China Hospital of Sichuan University, Chengdu, China; [3]Department of Respiratory Medicine, Xiangya Cancer Center, Xiangya Hospital, Central South University, Changsha, China; [4]School of Medicine, Shandong University, Jinan, China; [5]Department of Respiratory Medicine, The First Hospital of China Medical University, Shenyang, China; [6]Department of Pulmonary Medicine, Xijing Hospital, Air Force Medical University of PLA, Xi'an, China; [7]Department of Pulmonary Medicine, Zhongshan Hospital, Fudan University, Shanghai, China; [8]Department of Respiratory Medicine, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, China; [9]Department of Gerontol Respiratory Medicine, The Frist Hospital of Lanzhou University, Lanzhou, China; [10]Department of Pulmonary & Critical Care Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [11]AnchorDx Medical Co., Ltd., Guangzhou, China; [12]Department of Respiratory and Critical Care Medicine, Guizhou Provincial People's Hospital, Guiyang, China; [13]Department of Respiratory Medicine, QILU Hospital, Shandong University, Jinan, China; [14]Department of Pulmonary and Critical Care Medicine, Inner Mongolia Autonomous Region People's Hospital, Hohhot, China; [15]Department of Respiratory Medicine, The First People's Hospital of Yunnan Province, Kunming, China; [16]Department of Respiratory and Critical Care Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China; [17]Department of Respiratory Medicine, Shenzhen People's Hospital, Shenzhen, China; [18]Department of Pulmonary Medicine, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China; [19]Department of Respiratory Medicine, Shantou Central Hospital, Shantou, China; [20]Department of Respiration, The First Affiliated Hospital of Nanchang University, Nanchang, China; [21]Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital, Xi'an Jiaotong University, Xi'an, China; [22]Department of Cardiothoracic Surgery, The Second Affiliated Hospital of Xiamen Medical College, Xiamen, China; [23]Department of Respiratory Medicine, Henan Provincial People's Hospital, Zhengzhou, China; [24]Department of Pulmonary Medicine, Shanghai Chest Hospital, Shanghai Jiaotong University, Shanghai, China; [25]Department of Respiratory and Critical Care Medicine, Beijing Institute of Respiratory Medicine, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China; [26]Department of Pathology, School of Basic Medical Science, Southern Medical University, Guangzhou, China; [27]National Clinical Research Center for Respiratory Disease, State Key Laboratory of Respiratory Diseases, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

Contributions: (I) Conception and design: N Zhong, JB Fan, W Liang; (II) Administrative support: B Wang, J Wu, J Sun; (III) Provision of study materials or patients: W Liang, D Liu, M Li, W Wang, Z Qin, J Zhang, Y Zhang, Y Hu, H Bao, Y Xiang, C Hu, X Ye, X Zhang, W Xiao, C Yun, D Sun, W Wang, N Chang, Y Zhang, J Zhao, X Zhang, J Xu, D Wu, X Liu, Y Guo, Q Zhang, W Zhang, L Yang, Z Li, X Zhang, B Han, Z Tong, J He, J Qu; (IV) Collection and assembly of data: W Liang, D Liu, M Li, W Wang, Z Qin, J Zhang, Y Zhang, Y Hu, H Bao, Y Xiang, B Wang, J Wu, C Hu, X Ye, X Zhang, W Xiao, C Yun, D Sun, W Wang, N Chang, Y Zhang, J Zhao, X Zhang, J Xu, D Wu, X Liu, Y Guo, Q Zhang, W Zhang, L Yang, Z Li, X Zhang, B Han, Z Tong, J He, J Qu; (V) Data analysis and interpretation: W Liang, B Wang, JB Fan; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work.

Correspondence to: Jian-Bing Fan. Department of Pathology, School of Basic Medical Science, Southern Medical University, Guangzhou 510515, China. Email: jianbingfan1115@smu.edu.cn; Nanshan Zhong. National Clinical Research Center for Respiratory Disease, State Key Laboratory of Respiratory Diseases, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China. Email: nanshan@vip.163.com.

Background: Lung nodules are a diagnostic challenge. Current clinical management of lung nodule patients is inefficient and therefore causes patient misclassification, which increases healthcare expenses. However, a precise and robust lung nodule classifier to minimize discomfort for patients and healthcare costs is still lacking. The aim of the present protocol is to evaluate the effectiveness of using a liquid biopsy classifier to diagnose nodules compared to physician estimates and whether the classifier can reduce the number of unnecessary biopsies in benign cases.

Methods: A prospective cohort of 10,560 patients enrolled at 23 clinical centers in China with non-calcified pulmonary nodules, ranging from 0.5 to 3 cm in diameter, indicated by LDCT or CT will be included. After signed consent forms, the participants' pulmonary nodules will be assessed using three evaluation tools: (I) physician cancer probability estimates (II) validated lung nodule risk models, including Mayo Clinic and Veteran's Affairs models (III) ctDNA methylation classifier previously established. Each patient will undergo LDCT/CT follow-ups for 2 to 3 years and their information and one blood sample will be collected at baseline, 3, 6, 12, 24 and 36 months. The primary study outcomes will be the diagnostic accuracy of the methylation classifier in the cohort. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) will be used to compare the diagnostic value of each testing tool in differentiating benign and malignant pulmonary nodules.

Discussion: We are conducting an observational study to explore the accuracy of using a ctDNA methylation classifier for incidental lung nodules diagnosis

Trial registration: Clinicaltrials.gov NCT03651986.

Keywords: Pulmonary nodule; lung cancer early diagnosis; liquid biopsy

Introduction

Lung cancer is the leading cause of cancer deaths (1) and has been a research focus as a major public health problem worldwide. Even though the integrated application of surgery, chemotherapy, radiotherapy and targeted therapy has significantly improved survival rates, prognosis of patients with lung cancer is still relatively poor compared with other cancers. One major contributing factor is that a high proportion of lung cancer is diagnosed at advanced stage, related to the lack of prevalent early screening for lung cancer.

Now early screening methods for lung cancer mainly rely on imaging examination to detect pulmonary nodules, and then use surgery, biopsy and other means to determine the pathological type of pulmonary nodules. Relevant literatures suggest that the size of solitary pulmonary nodules is associated with the probability of malignant lesions (2,3). The larger the pulmonary nodules, the higher the risk of malignancy. The probability distribution of malignant nodules whose diameter is less than 5 mm, 5 to 10 mm or larger than 2 cm is 1%, 6% to 28%, and 64% to

82%, respectively. American College of Chest Physicians (ACCP) suggests that when the diameter of solitary pulmonary nodules is greater than 8 mm, the probability of nodules being malignant lesions increases significantly, and the possibility of malignancy should be alert.

The imaging examination of pulmonary nodules includes chest X-ray and CT. Although chest X-ray examination has high penetration, highly convenient usage and low radiation dose, it also has low resolution and difficulties in detecting lung lesions of concealed sites and minimal lesions. Low-dose chest CT (LDCT) screening is currently internationally recognized as the most effective method for the detection of pulmonary nodules and the early diagnosis of lung cancer (4). However, many non-neoplastic pulmonary nodules are also detected due to its high sensitivity; these nodules require long-term follow-up, repeated CT examinations, and even invasive operations such as puncture or surgery to make a definitive diagnosis. In order to increase the specificity of lung cancer screening, various methods are comprehensively used in clinic to differentiate pulmonary nodules found by LDCT. These mainly include high-risk factors assessment tools (such as Mayo Clinic and Veteran's Affairs models),

imaging assessment, serological examination and pathological examinations of tissue biopsy (5,6). The clinical high-risk factor assessment of pulmonary nodules only helps with clinical subsequent plan-making of follow-up, and it cannot reliably differentiate benign nodules from malignant ones (7). Presently, there is still no specific lung cancer markers that can be applied in the early diagnosis of lung cancer in the serological detection of lung cancer. Pathological tissue biopsy is the gold standard for the identification of benign and malignant pulmonary nodules, however it also can cause complications such as pneumothorax, hemorrhage (8) and temporary pulmonary disfunction. While many efforts have been made to reduce the false positive rate by combining LDCT with PET-VT, CTC, tumor markers or other examinations, the ideal specificity has not yet been achieved.

The purpose of diagnosis of benign and malignant pulmonary nodules is to improve the long-term survival rate of patients with lung cancer, which requires not only the early diagnosis and treatment of pulmonary nodules, but also the non-invasive detection technology to reduce side effects. With the rapid development of high-throughput sequencing technology, a variety of molecular markers have been found and applied in early diagnosis, prognosis assessment, efficacy monitoring and recurrence prediction in cancer management. DNA methylation is one of the major epigenetic modifications in diverse biological processes and diseases, especially in carcinogenesis (9). Studies have found that there are differences in DNA methylation positions and levels of methylation at different tumor stages, and abnormal DNA methylation is one of the characteristics of tumors (10). Methylation has previously been reported as an early molecular event in the occurrence of lung adenocarcinoma. In recent years, with the breakthrough of DNA methylation analysis technology and clinical translational research, DNA methylation detection can be used as a new technology for early screening and diagnosis of tumors.

Liquid biopsy assists in clinical diagnosis, prognosis and treatment choice for tumor diseases by detecting components of blood and other body fluids including circulating tumor cells (CTCs), circulating tumor DNA (circulating tumor DNA, ctDNA), microRNA and exosome (11-13). Because DNA is far more stable than RNA, proteins or small molecule metabolites, and ctDNA is present in blood stream at the early stage of neoplastic process while CTCs are not, ctDNA detection can be

accurately used for early diagnosis of tumors (14). A study suggested that mutations of oncogenes detected in the plasma and serum of tumor patients was consistent with the primary tumors (15). ctDNA, released by tumor cells into the blood stream, is an ideal biomarker for tumor diagnosis, treatment evaluation, and monitoring of recurrence. Compared with traditional tissue biopsy, ctDNA detection has many advantages particularly as a rapid, convenient, and non-invasive method.

DNA methylation levels in peripheral blood of patients with a tumor were significantly different compared to normal controls. Sun *et al*. compared the status of DNA genome methylation in plasma of patients with liver cancer and healthy individuals, and established a diagnostic model based on ctDNA methylation with sensitivity and specificity up to 80% and 80%, respectively (16). Meanwhile, investigators were able to analyze the source of ctDNA in tumor patients with the predictive accuracy of 80% by means of detecting methylation position of tumor specific DNA in plasma (17). In the preliminary study, high-throughput DNA methylation sequencing of malignant and benign tissue samples was performed to obtain specific markers for differentiation. Together with mining the methylation data from two databases, The Cancer Genome Atlas (TCGA) and The International Cancer Genome Consortium (ICGC), specific methylation markers for malignant tissues (*Figure 1*) were identified and applied to ctDNA methylation detection in plasma samples. In view of the extremely low fraction of ctDNA in plasma, we have developed a highly sensitive method, AnchorIRIS™, for detecting cell-free DNA (cfDNA) as low as 1 ng (18). We demonstrated the sensitivity and specificity of this test were 82.5% and 83.3%, respectively in a retrospective cohort of 66 patients with benign nodules and 63 with malignant nodules whose histological and pathological diagnosis were in place (*Table 1*); especially the sensitivity of ctDNA methylation test reached up to 81.5% for 27 patients with lung cancer in stage Ia. However, the accuracy of the classifier in a larger, perspective cohort remains unknown. This study aims to evaluate the effectiveness of using a liquid biopsy classifier to diagnose nodules compared to physician estimates and whether the classifier can reduce the number of unnecessary biopsies in benign cases.

In this paper, we present the study protocol according to the SPIRIT protocol guidelines (19) (available at http://dx.doi.org/10.21037/tlcr-20-701).
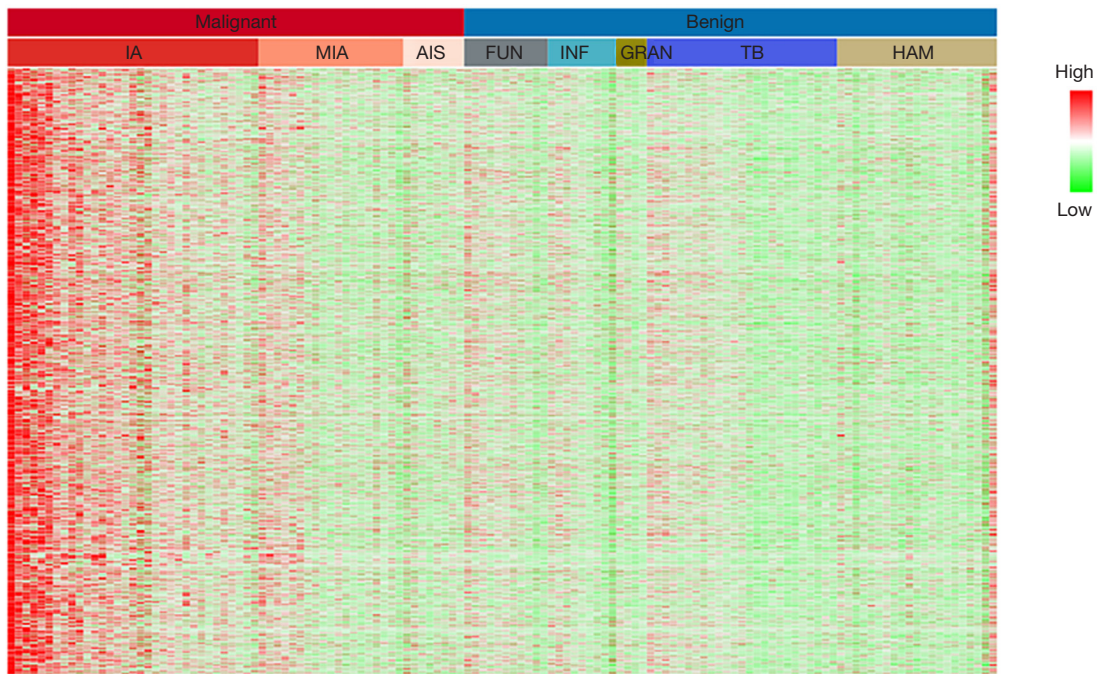
Figure 1 Different areas of DNA methylation in malignant and benign pulmonary nodules.

Table 1 Pathological diagnosis & ctDNA methylation detection in patients with pulmonary nodules

|	Positive	Negative	Total	Sensitivity
Tumor pathological subtype				
Invasive lung adenocarcinoma (IA)	33	9	42	78.60%
Minimal invasive lung adenocarcinoma (MIA)	10	2	12	83.30%
Other	9	0	9	100%
Total	52	11	63	82.50%
Benign lesions subtype				
Hamartoma (HAM)	3	2	5	
Tuberculosis granuloma (TB)	0	6	6	
Fungal infection (FUN)	0	1	1	
Inflammation (INF)	2	2	4	
Other	0	1	1	
Normal (NOR)	6	43	49	
Total	11	55	66	83.30%
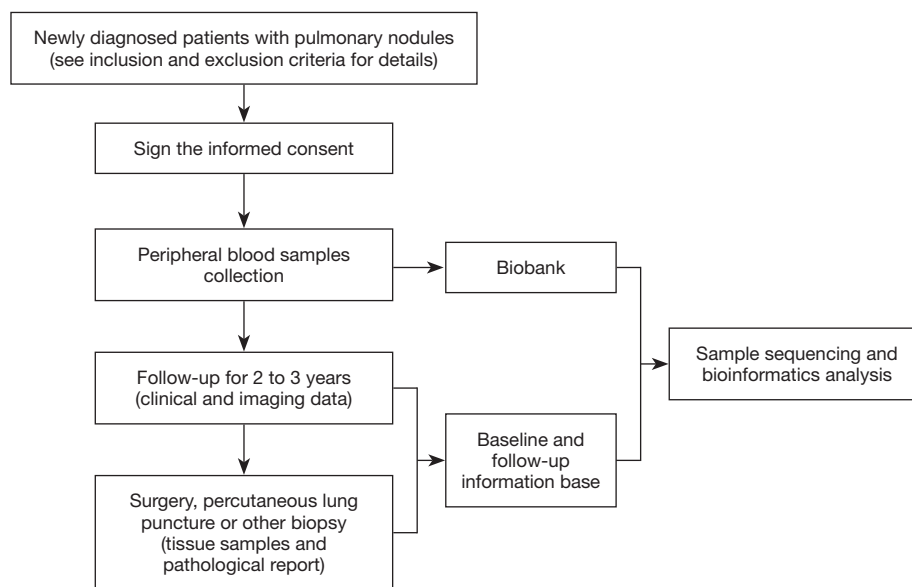
Figure 2 Summary of research procedures.

Methods

Trial design

This is a prospective, observational and multi-center clinical trial for patients with positive pulmonary nodules, followed up by LDCT or CT for 2 to 3 years, with the collection of blood samples, pathology reports of patients undergoing surgery, and fresh or archived tissue samples (e.g., FFPE blocks) during follow-ups. At the end of follow-up, genetic testing and bioinformatics analysis will be performed on the collected samples and expect that combined with targeted high-throughput DNA methylation sequencing of ctDNA, the performance of differential diagnosis for patients with positive pulmonary nodules will be verified and improved.

This trial will be conducted at The First Affiliated Hospital of Guangzhou Medical University in addition to 22 other hospitals in China. Each participant will visit the trial institutions at most 6 times for 3 years. If subjects are diagnosed by surgery, percutaneous lung puncture or other biopsies during visits, they will also be considered as having completed the study. At baseline visit (V0), general data of participants will be collected and those met the eligibility criteria will be enrolled, sign the informed consent, have 10 ml of venous blood collected, and finish *The Subjects' Lung History Questionnaire*. At each following visit, third month (V1), sixth month (V2), first year (V3), second year (V4) and third year (V5) after baseline visit, participants will have 10 ml of venous blood collected and, complete CT/LDCT

examination as well. Meanwhile, investigators should finish *The Clinical Assessment Form*. Substantial follow-up arrangements after each visit (V1, V2, V3, V4, V5 or Vx) for each subject will be determined by investigators based on baseline and/or current LDCT/CT imaging diagnosis. When in an unexpected visit (Vx), if surgery, percutaneous lung puncture or other biopsies is operated within 3 years after the baseline visit, 10 mL of preoperative venous blood and fresh or archived tissue samples (e.g., FFPE blocks) shall be collected and a pathological report shall be obtained. After obtaining blood, fresh or archived tissue samples, the research center will send the samples to the testing center (AnchorDx Medical Co., Ltd.) for subsequent testings. Meanwhile, the participants' pulmonary nodules will be assessed using two current clinical evaluation tools: (I) validated lung nodule risk models, including Mayo Clinic and Veteran's Affairs models (II) Physician cancer probability estimates. Adverse events during the trial should be recorded. The clinical trial flow chart and schedule of events are shown in *Figure 2* and *Table 2*.

Study objectives

The primary objective of this trial is to evaluate the diagnostic accuracy of ctDNA methylation classifier in differentiating benign and malignant pulmonary nodules. AUC-ROC, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), of the

Table 2 Clinical trial flow chart

Assessment/ Procedures	Enrolment	Follow-up					
	V0	V1*	V2*	V3*	V4*	V5*	Vx*
	−2~0 mo.	3 mo. ± 2 wk.	6 mo. ± 2 wk.	12 mo. ± 2 wk.	24 mo. ± 1 mo.	36 mo. ± 1 mo.	Surgery/percutaneous lung puncture/the other biopsies
General Data	×						
Inclusion and Exclusion criteria	×						
Informed Consent	×						
The Subjects' Lung History Questionnaire	×						
Blood Sample Collection	×	×	×	×	×	×	×
CT/LDCT	×	×	×	×	×	×	×
The Clinical Assessment Form	×	×	×	×	×	×	
Fresh Tissue or FFPE Samples							×
Pathological Report							×
Adverse Events during the trial		×	×	×	×	×	×

*, substantial follow-up arrangements after each visit (V1, V2, V3, V4, V5 or Vx) for each subject will be determined by the investigator based on baseline and/or current CT/LDCT imaging diagnosis.

classifier as well as the clinically applied lung nodule risk models and physician cancer probability estimates will also be determined and compared in current cohort study, which will be used to calculate the potential clinical value of using the liquid biopsy tool in routine diagnostic tests for diagnosis of lung nodules, esp. reduction of unnecessary biopsies in benign cases.

Expected duration and study population

Ten thousand five hundred and sixty patients eligible for inclusion in the study will be recruited in 23 centers throughout China. Study recruitment was planned to be completed within 12 months, with follow-up observation for 24 to 36 months and finished within 48 months expectedly.

Inclusion criteria for participation in the study defined

Subjects must meet all the following criteria to be eligible for this study:

(I) Either sex, age ≥18 years;

(II) Patients with positive pulmonary nodules detected by standard-dose or low-dose CT screening:

    (i) The diameter of positive pulmonary non-calcified nodules is 5 to 30 mm;

    (ii) Three imaging types of pulmonary nodules will be included, solid nodules, part-solid nodules (mixed ground-glass nodules) and pure ground-glass nodules.

(III) Patients with newly positive pulmonary nodules detected by baseline CT were initially diagnosed or received within 60 days before enrollment;

(IV) Participants who are willing to fill in *The Subjects' Lung History Questionnaire*;

(V) Patients accept follow-up for 2 to 3 years and cooperate with the relevant imaging, serological or surgery and other examinations;

(VI) Ability to fully understand the informed consent, agree to participate in the study and sign the informed consent.

Exclusion criteria defined

(I)   Pregnant or lactating females;
(II)   Patients underwent any diagnostic puncture therapy, such as percutaneous lung biopsy, transbronchial biopsy or surgery prior to enrollment;
(III)   Receive any blood transfusion therapy within 30 days prior to enrollment;
(IV)   Patients with cancer confirmed pathologically within 2 years prior to enrollment (except for nonmelanoma skin cancer);
(V)   Inability to understand or obtain informed written consent.

Discontinuing study interventions and patient withdrawal

Subjects may discontinue the study intervention prematurely for the following reasons:

(I)   Subjects request to leave the trial i.e., patients withdraw informed consent and are free to leave the study at any time;
(II)   Patient dies;
(III)   Subjects are lost to follow-up;
(IV)   The researcher believes that patients do not follow instructions in protocol;
(V)   Individuals who are judged inappropriate for the study by investigator.

If a subject discontinues the intervention or chooses to discontinue the intervention prematurely, an early withdrawal visit will be completed without delay but no later than one week after discontinuation. The reason(s) for premature discontinuation will be documented accordingly. Every effort will be made to follow-up subjects who terminate prematurely to determine the final clinical outcome.

Expected outcome

Based on the results of previous studies and preliminary studies, the accuracy of the ctDNA methylation classifier was expected to be around 0.83. The AUC score for physician assignment of nodules was expected to reach up to 0.80, the Mayo and VA models up to 0.75, the ctDNA methylation classifier up to 0.80 or above.

Patient and public involvement statement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

Risks and benefits

This study is an observational study and will not interfere with the pathological diagnosis and treatment of patients. Therefore, participation in this study will not lead to the occurrence of disease progression and other risks for patients. Venous blood collection will be performed strictly in accordance with requirements of sterilization by professionals. There may be only very small risks associated with venous blood collection including temporary pain, local ecchymosis, mild dizziness in a few people, or extremely rare needle infections. Subjects enrolled in this study will be followed up on time according to the follow-up recommendations of investigators, and will receive free CT/LDCT examination in every follow-up site. The results of follow-up imaging examination will be helpful for the investigators to make further targeted follow-up or therapeutic schedule. During follow-up, for subjects diagnosed with lung cancer or highly suspected lung cancer by CT/LDCT, investigators will give priority to their following treatment in this center or assist them to seek medical treatment in the nearby sub-center participating in this trial as required.

Adverse events observation, recording and disposal

Pain and bruising are self-limiting and patients can be relieved by a moderate hot compress. In the event of severe malaise or infection, investigators would use their judgement to provide appropriate medical treatment.

Statistical analysis

Sample size calculation

This study predicts that both the sensitivity and specificity of ctDNA methylation detection in the diagnosis of malignant pulmonary nodules, based on the pre-experimental results, will be 80%. At the tolerance of 0.05 (α=0.05), at least 264 patients with benign pulmonary nodules and 264 patients with malignant pulmonary nodules should be included according to the calculation formula of diagnostic test sample size. Based on previous relevant studies, 3% of people with positive pulmonary nodules were

diagnosed with lung cancer; considering a drop-out rate of 20%, the required sample size is N=10,560. Thus, 10,560 participants will need to be enrolled for this trial.

Assessment of diagnostic performance

Definition of benign and malignant pulmonary nodules

The definition of benign pulmonary nodules: (any of the following conditions is satisfied):

(I)   Definite pathological diagnosis;

(II) Pulmonary nodules shrank and disappeared after anti-inflammatory treatment;

(III) Based on CT/LDCT follow-up for 2 years, the size of the solid pulmonary nodules was stable;

(IV) For the pure ground-glass nodules or part-solid nodules (mixed ground-glass nodules), if the size and density of the pulmonary nodules were stable after CT/LDCT follow-up for 3 years, nodules are considered as benign.

The definition of malignant pulmonary nodules: those malignant are diagnosed by tissue pathology.

Assessment indicators of diagnostic performance

All statistical analysis will be performed using the statistical software package SPSS 20.0. According to the data type and distribution characteristics, the data will be statistically described by means of mean, standard deviation (SD), percentage. Differences will be compared between groups of patients with benign and malignant pulmonary nodules using two independent sample *t*-tests or rank-sum tests for continuous measures and χ^2 test or Fisher's exact test for categorical measures. The accuracy of ctDNA methylation detection in the diagnosis of benign and malignant pulmonary nodules will be assessed by sensitivity, specificity, PPV, NPV, positive likelihood ratio, negative likelihood ratio and accuracy. At the same time, the sensitivity and specificity of ctDNA methylation detection in the diagnosis of malignant pulmonary nodules in patients with different pathological subtypes will be further evaluated, including invasive lung adenocarcinoma (IA), minimally invasive lung adenocarcinoma (MIA), and adenocarcinoma in situ (AIS).

Diagnostic assessment indicators are defined as follows:

(I)   Sensitivity: the probability of ctDNA methylation detection model being positive in patients with pathological diagnosis of malignant pulmonary nodules;

(II) Specificity: the probability of ctDNA methylation test model being negative in patients with pathological diagnosis of benign pulmonary nodules;

(III) Accuracy: the proportion of patients with benign and malignant pulmonary nodules correctly determined by ctDNA methylation detection model;

(IV) PPV: the proportion of patients with positive pulmonary nodules determined by ctDNA methylation test model who were pathologically diagnosed as malignant pulmonary nodules;

(V)   NPV: the proportion of patients with negative pulmonary nodules determined by ctDNA methylation test model who were pathologically diagnosed as benign pulmonary nodules;

(VI) Positive likelihood ratio: sensitivity/(1–specificity);

(VII) Negative likelihood ratio: (1–sensitivity)/specificity.

As for the assessment of potential clinical impact, invasive procedure utilization following initial nodule detection by different methods will be tabulated according to the final diagnosis. Then the number and percentages of invasive testing that could have been avoided had the classifier test been available and used for nodule management will be calculated.

Trial management

Safety assessments

This clinical trial will be conducted in strict compliance with regulations, protocol and Standard Operating Procedure (SOP) requirements. Adverse events (AE) and Serious Adverse Events (SAE) will be recorded. Several factors affecting the trial population suggest that we would expect to observe a larger than normal incidence of episodes of ill-health due to both the age and co-morbidities of the study population. AEs (as defined) will be recorded as soon as they are known from any of the study subjects.

Ethical considerations

All research centers should obtain approval documents from the corresponding ethics committee or institutional review committee before the study is conducted in accordance with the Declaration of Helsinki, ICH-GCP (Good clinical practice provided by the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use), Chinese laws and regulations and the requirements of relevant organizations. Any extensions, amendments or updates in the protocol must be reported to and approved by the ethics committee for all necessary modifications.

Protocol amendments

Protocol modifications are not expected. However, if they are necessary, any changes in the protocol, which may affect the conduct of the study, potential benefits to the patients or patient safety, including changes to the study design, should be proposed by the sponsor, and reported to and approved by the ethics committee before implementation.

Data management

Data processing & *data entry*

The data collected in this clinical trial includes informed consents, basic information recording patients' diagnosis and treatment records (including medical records, imaging data and reports, blood test reports and pathological reports, etc.), biological sample collection forms and records. All research data will be included into the Information Management System by the personnel of participating units and the paper and digital versions will be submitted to the project leading entity step by step on schedule after being checked and qualified by quality controllers. The biological sample management center will designate professionals to input the biological sample information including detailed records made for the entry and exit of biological samples, timely submission to the information management center for entry into the data platform of biological sample information and report to the trial leading unit.

Data collection & *statistics analysis*

The data management center is responsible for integrating basic information of the population, diagnosis and treatment and follow-up information, processing and analyzing relevant data according to the trial design. After verification of statistical analysis results, the statistical analysis summary report and statistical analysis summary report of each subcenter are produced and reported to the project leading unit. Of note, data monitoring committee (DMC) will be formed to monitor patient safety and treatment efficacy data during the clinical trial. There are five members of the Data Monitoring committee, Professor Nanshan ZHONG, as lead principal investigator of this trial; Dr. Wenhua LIANG, as sub-investigator of this trial; Professor Pingyan Chen, as chief statistician of this trial; Professor Jian-Bing FAN, the CEO of AnchorDx, as chief scientist of this trials; and Mr. Bo WANG, the Medical Director of AnchorDx, as project manager of this trial.

Instructions & informed consent for subjects

Core information and informed consent of the trial will be provided for all subjects. Prior to study initiation, investigators must obtain the approval or consent of the ethics committee or institutional review committee for the written consent and other written information to be provided to the subjects. Written approval documents from the ethics committee or institutional review committee and the informed consents must be filed together in the trial documents.

Confidentiality

All records of the subjects' identity will be kept in a secure storage area or in password-protected electronic documents with access limited to authorized study staff only. Clinical information will not be released without the written permission of the participant, except as necessary for monitoring and auditing by the Sponsor, its designee or Regulatory Authorities. Only the subject number and initials are recorded in the case report form, and If the subject's name appears in any other document (e.g., pathological report sheet), it must be hidden before a copy of the document is provided to the sponsor and the research reports stored on computers must comply with local laws on data protection. Published results will not contain any personal data that could allow identification of individual participants and investigators will keep a list of subjects to check on their records.

Dissemination

All information and results obtained from this study, as well as all intellectual property rights, such as the publication of articles, patent application or application for scientific and technological awards, will be subject to the approval by sponsors, National Clinical Research Center for Respiratory Disease in The First Affiliated Hospital of Guangzhou Medical University and AnchorDx Medical Co., Ltd. Data obtained from this study may be used solely by investigators for scientific purposes, while investigators must obtain the written consent in consultation with the sponsor prior to publication. The sponsor recognizes that the investigators have the right to publish the results of the trial after the completion of the study. Individual participants may not publish data concerning their patients that are directly relevant to questions posed by the trial. Data from all centers will be analyzed together and published as soon as possible and the findings of this trial will be submitted for publication in a peer-reviewed journal in respiratory disease. Abstracts will be submitted to the relevant national and international conferences.

Documents file

As required by relevant regulations, investigators will properly keep original records of this clinical trial. All clinical trial data must be kept for 5 years according to the requirements of good clinical practice (GCP). The sponsor shall contact the trial unit to discuss the subsequent storage of materials within 30 days prior to the end of the storage period. If the sponsor fails to contact the research unit after the expiration of the storage period, the research unit will have the right to dispose of the research materials by itself in a reasonable way.

Study sponsorship and oversight

National Clinical Research Center for Respiratory Disease, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong, China will be overseeing the overall conduct of the study, including adverse event monitoring and data management. Interim analyses will be conducted upon the completion of 50% enrollment, 100% enrollment, 50% subject out and the last subject out of the study. Adverse events will be reported in the patient file. They will be assessed and managed according to GCP. Adverse events will be listed in the final study report. The independent statistical analysis will be performed by National Clinical Research Center for Respiratory Disease, The First Affiliated Hospital of Guangzhou Medical University.

Timeline

The study started recruitment in September 2018, and completion of data collection is anticipated by March 2023.

Discussion

The primary aim of this study was to evaluate the evaluate the effectiveness of using a liquid biopsy classifier based on ctDNA methylation analysis with NGS to diagnose nodules compared to physician estimation and whether the classifier can reduce the number of unnecessary biopsies in benign cases. The major strength of this study, to our knowledge, is that it is the first and largest trial worldwide to evaluate the performance and clinical impact of a novel liquid biopsy based ctDNA methylation analysis tool, as well as comparison to pathologic diagnosis in diagnosing and following up of pulmonary nodules. Some similar studies have been initiated, but the sample sizes are smaller, and no significant achievements have been released yet.

This study has some potential limitations. The major one may be that it only includes Chinese population, and this tool has not been applied extensively to patients in different cohorts from various races and regions. It needs to be externally validated before they can be used in clinical practice worldwide. Furthermore, not all cases in this study will finally have a pathological diagnosis, especially those with clinically indicated benign diseases or stable malignancies, that is another possible limitation of our study.

Footnote

Reporting Checklist: The authors have completed the SPIRIT reporting checklist. Available at http://dx.doi.org/10.21037/tlcr-20-701

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tlcr-20-701). WL serves as an unpaid editorial board member of *Translational Lung Cancer Research* from Apr 2018 to Apr 2021. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study protocol and informed consent documents were approved by the ethical committees of the

participating institutions, Ethics Committee of The First Affiliated Hospital of Guangzhou Medical University (No. 2018-81), and informed consent will be obtained from all patients.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. CA Cancer J Clin 2016;66:115-32.
2. Shankar A, Saini D, Dubey A, et al. Feasibility of lung cancer screening in developing countries: challenges, opportunities and way forward. Transl Lung Cancer Res 2019;8:S106-21.
3. Takashima S, Sone S, Li F, et al. Small solitary pulmonary nodules (< or =1 cm) detected at population-based CT screening for lung cancer: Reliable high-resolution CT features of benign lesions. AJR Am J Roentgenol 2003;180:955-64.
4. Ferrari A, Bertolaccini L, Solli P, et al. Digital chest tomosynthesis: the 2017 updated review of an emerging application. Ann Transl Med 2018;6:91.
5. Ettinger DS, Wood DE, Aisner DL, et al. Non-Small Cell Lung Cancer, Version 5.2017, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2017;15:504-35.
6. Wender R, Fontham ET, Barrera E Jr, et al. American Cancer Society lung cancer screening guidelines. CA Cancer J Clin 2013;63:107-17.
7. Wu J, Zan X, Gao L, et al. A Machine Learning Method for Identifying Lung Cancer Based on Routine Blood Indices: Qualitative Feasibility Study. JMIR Med Inform 2019;7:e13476.
8. Bai C, Choi CM, Chu CM, et al. Evaluation of Pulmonary Nodules: Clinical Practice Consensus Guidelines for Asia. Chest 2016;150:877-93.
9. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 2012;13:484-92.
10. Hollstein M, Sidransky D, Vogelstein B, et al. p53 mutations in human cancers. Science 1991;253:49-53.
11. Crowley E, Di Nicolantonio F, Loupakis F, et al. Liquid biopsy: monitoring cancer-genetics in the blood. Nat Rev Clin Oncol 2013;10:472-84.
12. Diaz LA Jr, Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. J Clin Oncol 2014;32:579-86.
13. Cai X, Janku F, Zhan Q, et al. Accessing Genetic Information with Liquid Biopsies. Trends Genet 2015;31:564-75.
14. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014;6:224ra24.
15. Forshew T, Murtaza M, Parkinson C, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 2012;4:136ra68.
16. Sun K, Jiang P, Chan KC, et al. Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. Proc Natl Acad Sci U S A 2015;112: E5503-12.
17. Guo S, Diep D, Plongthongkum N, et al. Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. Nat Genet 2017;49:635-42.
18. Liang W, Zhao Y, Huang W, et al. Non-invasive diagnosis of early-stage lung cancer using high-throughput targeted DNA methylation sequencing of circulating tumor DNA (ctDNA). Theranostics 2019;9:2056.
19 Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013;158:200-7.