Diagnostic value of the circulating tumor cells and circulating tumor-derived endothelial cells detection for non-small cell lung cancer

Jianzhu Xie[1], Zheng Ruan[2], Jian Zheng[2], Yanping Gong[1], Yulan Wang[1], Binjie Hu[1], Jin Cheng[1,*], Qian Huang[1,*].

[1]Cancer Center, [2]Department of Thoracic Surgery, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

***Corresponding Author:**

Drs. Qian Huang and Jin Cheng, Cancer Center, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, 650 Xinsongjiang Road, Songjiang District, Shanghai 201620, China. Tel: +86-21-37798906. Fax: +86-21-37798833. E-mail: qianh2011@126.com (QH), ajinch@163.com (JC).

Grant Support: This study was supported by the National Natural Science Foundation of China (grant 81972843 to Q. Huang and grant 81972887 to J. Cheng).

【Abstract】

Effective biomarkers are essential to the early diagnosis of non-small cell lung cancer (NSCLC). Herein, a retrospective study of 49 newly diagnosed and recurrent NSCLC patients, 31 patients with benign pulmonary disease and 24 healthy volunteers was conducted, to evaluate the diagnostic value of circulating rare cells for NSCLC. The expression of circulating tumor cells (CTCs) and circulating tumor-derived endothelial cells (CTECs) in peripheral blood were measured by subtraction enrichment-immunostaining-fluorescence in situ hybridization (SE-iFISH). The level of CTCs ($P<0.001$) and CTECs ($P<0.001$) was significantly higher in NSCLC group than that in benign pulmonary disease group. The proportion of small CTCs ($P<0.001$) and CTECs ($P<0.0001$) significantly increased from benign lung disease individuals to NSCLC patients. The AUC of ROC curves of total CTCs and CTECs were 0.815 (95%CI: 0.722~0.907), 0.739 (95%CI: 0.618~0.860), respectively. The cut-off values for discriminating NSCLC with benign lung disease patients were total CTCs 11.5 units/6ml and total CTECs 10.5 units/6ml, with sensitivity and specificity being 67.3% and 83.9%, 77.6% and 77.4%, respectively. When CTCs and CTECs were combined, predictive value significantly increased to 82.6% as measured by the area under the curve. Small CTCs and triploid CTCs had high positive predictive value (PPV) and positive likelihood ratio (LR+) of the diagnosis of NSCLC in early stage. CTCs and CTECs can not only be used as new biomarkers for the diagnosis of NSCLC, but can also improve diagnostic performance of the early stage NSCLC. Moreover, the combined examination of CTCs and CTECs is be superior to the single.

【Keywords】 Non-small cell lung cancer; Circulating tumor cells; Circulating tumor-derived

 doi:10.20944/preprints202101.0129.v1

endothelial cells; SE-iFISH; Biomarkers

1. Introduction

Lung cancer is one of the most rapidly increasing malignant tumors in the world. In China, lung cancer is also the most common cause of death from the neoplastic diseases, with an overall five-year survival rate of only 15% [1]. According to histopathological classification, lung cancer can be classified into small cell lung cancer and non-small cell lung cancer (NSCLC). NSCLC, which accounts for 85% of all lung cancers [2], can be mainly further classified into squamous cell carcinoma, adenocarcinoma, large cell carcinoma, etc. Although mortality and recurrence rates are high for lung cancer, the 5-year survival rate is 56% if detected at early stage(I-II) [3], compared with 5% for the advanced lung cancer [4]. Unfortunately, only 0.1-0.17% percent of lung cancer cases are currently diagnosed at early stage [5]. Therefore, early detection and early treatment of lung cancer has always been the main direction of the diagnosis and treatment of lung cancer.

In current clinical practice, screening of lung cancer mainly relies on tumor biomarkers such as cancer embryonic antigen (CEA), squamous cell carcinoma antigen (SCC), and cytokeratin-19 fragment (CYFRA 21-1) combined with low-dose computed tomography (LDCT) imaging. Whereas, the gold standard for lung cancer diagnosis depends on biopsy of tumor tissue. However, tumor biomarkers related to lung cancer could be non-specific since they can also increase in other diseases or inflammation. Imageogical examination are difficult to detect early micro-lesions, and invasive biopsy often results in poor patient compliance. The early diagnosis of lung cancer still remains a major problem to be solved.

In recent years, liquid biopsy techniques centering on patients' blood and body fluids have provided new insight for the early diagnosis of cancer and the early detection of metastasis and recurrence. Compare with tissue samples, blood and body fluid samples are easier to obtain and are preferred for real-time monitoring. Liquid biopsy mainly detects circulating rare cells (CRCs), which including circulating tumor cells (CTCs) and circulating tumor-derived endothelial cells (CTECs), circulating tumor DNA (ctDNA) released by tumor issue, exosomes, and platelets that were circulating in the peripheral blood of cancer patients [6, 7]. Several studies have shown the concentrations of tumor-related liquid biopsy markers increased in patients with early, recurrent and metastatic cancers such as breast, prostate, and colon cancer [8-10].

Circulating tumor cells (CTCs) are cancer cells shed from primary or metastatic solid tumors into peripheral blood, whereas circulating tumor-derived endothelial cells (CTECs) are derived from endothelial cells of tumor into circulation [11]. As the most important biomarker of liquid biopsy, CTCs has been recognized by the American Society of Clinical Oncology (ASCO) as an acceptable tumor marker since 2007 [12]. Many clinical studies have confirmed that CTCs have become a powerful tool for screening high-risk groups, early diagnosis, efficient monitoring cancer progression and metastasis and prognosis evaluation for lung cancer [13-16]. Furthermore, as an important factor of tumor angiogenesis [17, 18], aneuploid CD31+ CTECs have been confirmed existed among enriched non-hematopoietic CRCs in the peripheral blood [19, 20]. However, there is no clinical study on whether CTECs can be used as an important biomarker for the diagnosis of early stage lung cancer.

A wide variety of CRC detection techniques have been reported [21-24]. Most techniques are based on the positive capture of epithelial cell antigens or staining of cellular proteins such as epithelial cell adhesion molecule (EpCAM) and monoclonal antibody antihuman cytokeratin (CK).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

However, due to high heterogeneous and dynamic expression of EpCAM on tumor cells [25], as well as down-regulation of CK during epithelial-to-mesenchymal transition (EMT) [26], the clinical application of EpCAM or CK-dependent capture techniques are liable to miss detection. A comprehensive and stable detection technique should be applied in clinical practice.

In this study, subtraction enrichment-immunostaining fluorescence *in situ* hybridization (SE-iFISH) were performed to investigate the expression of circulating rare cells in the peripheral blood of NSCLC patients. With the use of magnetic beads coated with a variety of leukocyte surface antibodies, we performed a negative collection of circulating rare cells, which effectively avoided the heterogeneity of cell surface antigen expression and epithelial-mesenchymal transformation (EMT) interference effect on the results. At the same time, with combined chromosome 8 aneuploidy detection, we further found circulating rare aneuploidy cells, which provides a clue for the early detection and intervention of NSCLC.

2. Material and methods

2.1 Study subjects and specimen collection

A retrospective study was conducted on 49 patients with non-small cell lung cancer who were newly diagnosed and relapsed at the Department of Thoracic Surgery and Cancer Center in Shanghai General Hospital from August 2017 to April 2019, including 26 males and 23 females, and the average age was 61.69±11.34 years. The subject patients were diagnosed with NSCLC or benign lung diseases through clinical manifestations, medical history, and pathology. None of them received any anti-tumor treatment before detection. According to the 2015 World Health Organization (WHO) criteria for histological classification of lung tumors, there were 36 cases of lung adenocarcinoma, 8 cases of lung squamous carcinoma, 3 cases of lung adeno-squamous cell carcinoma, and 2 cases of other subtypes. According to International Association for the Study of Lung Cancer (IASLC 2009, TNM staging standard for lung cancer, 2009, 7th edition), 27 cases were stage I, 2 cases were stage II, 9 cases were stage III, and 11 cases were stage IV. Furthermore, there were 31 patients with benign lung diseases including 17 males and 14 females with average age (46 ± 11.7 years). As control，24 healthy volunteers were recruited，including 8 males and 16 females with average age (31.29±7.98 years). The study has been approved by the Ethics Committee of Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, China. All participants signed informed consent prior to blood collection. Peripheral venous blood was collected using ACD anticoagulant vacuum vessel collection tube. To avoid contamination of epithelial cells, the first 2 ml blood was discarded, 6 ml blood was then collected by ACD collection tube, and the blood were immediately mixed upside down for 8 times. It is best to store in room temperature and away from light for 24 hours, not more than 48 hours.

2.2 Subtraction Enrichment (SE)

The circulating tumor cells and circulating tumor endothelial cells in peripheral blood are firstly enriched according to the manufacture's updated protocol v5.0 (Cytelligen, San Diego, CA, USA). Blood samples were centrifuged at 200×*g* for 15 min at room temperature to separate plasma. Sedimented blood cells were mixed with 1×CRC buffer to 6 ml, followed by loading on the top of non-hematopoietic cell separation matrix in a 50 ml tube. Samples were subjected to centrifugation at 350×*g* for 6 min. Layers containing white blood cells (WBCs) and tumor cells above red blood cells (RBCs) were collected into a 50 ml tube, and subsequently incubated with 300 μl of immuno-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

magnetic beads coated with leukocyte surface antibodies at room temperature for 30 min. WBCs bound to immuno-beads were depleted using a 50 ml size magnetic separator (Cytelligen). The solution, free of magnetic beads, was collected into a 50 ml tube, followed by adding 1×CRC buffer to 50 ml. Samples were then spun at 500×*g* for 5 min at room temperature, followed by aspirating supernatants down to 100 µl. sedimented cells were gently resuspended, followed by multiple immunofluorescence staining and subsequent chromosome 8 centromere probe (CEP8) hybridization.

2.3 iFISH

Five color iFISH in this study was performed according to the manufacture's updated protocol v4.0 (Cytelligen). We chose 3 kinds of tumor-biomarker to tag different CTCs with various phenotypic, epithelial marker (EpCAM or CK18) and immunosuppressive marker: programmed death-ligand 1 (PD-L1). The resuspended sedimented cells were incubated with indicated post-fluorescence labeled monoclonal antibodies at 1: 200 dilution, including Alexa Flour 594-CD45 (red), Alexa Flour 488-CD31, CK18, EpCAM or PD-L1(green), Cy5-CD31, CK18, EpCAM (gold) at room temperature for 20 min in dark [27, 28]. Since the study applied a single tumor specific biomarker (CK18, EpCAM or PD-L1), different type of tumor biomarker used in each group was shown in Table 1. After washing, dry monolayer of cells mixed with the special fixative coated on the formatted CTC slides (Cytelligen). Furthermore, hybridization with Vysis chromosome 8 centromere probe (CEP8) Spectrum Orange for 4 hours was performed on the enriched cells by using a S500 ThermoBrite Slide Hybridization System. Finally, samples were mounted with mounting media containing DAPI (Vector Laboratories, Burlington, CA, USA), and subjected to automated CTC image scanning and analyses (Carl Zeiss, Oberkochen, Germany; MetaSystems, Altlussheim, Germany; and Cytelligen) [19].

2.4 Identification of lung cancer CTCs and CTECs

By high through-put scanning, obtaining and processing cell images , target cells were judged to be a circulating tumor cell or a circulating tumor endothelial cell, according to the expression of CD45, CD31, tumor markers (CK18, EpCAM, PD-L1) and karyotypic combined with the cell morphology and size [11]. Positive target cells are defined as follows: Morphological criteria: The cells are round or oval in shape, and the cell morphology and margins are intact under fluorescence. WBCs (5μm) were used as cutoff to define large or small cells. Immunofluorescence staining criteria: CTC: DAPI+/CD45-/tumor biomarker+/CD31-with chromosome 8 diploid or aneuploid; DAPI+/CD45-/tumor biomarker-/CD31- with chromosome 8 aneuploid. CTEC: DAPI+/CD45-/tumor biomarker +/CD31+. CTM: Cluster containing two or more CTCs.

2.5 Serum tumor markers detection and positive criteria

Three serum tumor markers: cancer embryonic antigen (CEA), squamous cell carcinoma antigen (SCC), and cytokeratin-19 fragment (CYFRA 21-1), associated with non-small cell lung cancer were selected for serological detection. CEA and CYFRA21-1 was detected by electrochemical luminescence immunoassay (Roche Cobas® C-601), and SCC was detected by chem-iluminescent microparticle immunoassay (Abbott Architect i2000 analyzer), respectively. The testing procedure was strictly in accordance with the kit instructions and the requirements of the testing instruments, and the normal reference values of the kits are recommended as follows: SCC< 2.5 ng/ml, CEA < 5

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

ng/ml, CYFRA21-1<3.3 ng/ml. Any one of the three above the normal value is considered positive.

2.6 Statistical analysis

All data were statistically analyzed by SPSS 25.0 software. Measurement data were compared between groups by Mann-Whitney U test (two groups) or Kruskal-Wallis test (multiple groups), and counting data were compared by Chi-Square test. The inspection level was $\alpha = 0.05$. ROC analysis was performed to determine the cut-off values, and valuate the predictive performance of the disease progression-related biomarker signatures. For ROC analysis of the CTCs combination with CTECs, *P* (probability of a patient sample) was calculated for in the logistic regression by the formula: $Z=\text{logit}(P) = \ln(P/1-P)=b0+b1X1+b2X2+b3X3+\ldots+bnXn$, where the bi terms were the ith regression coefficients by binary logistic regression, and the Xi terms were the relative expression levels of each biomarker. All *P* values were two-sided. $P < 0.05$ are statistically significant.

3. Results

3.1 Characterization of CTCs and CTECs in NSCLC patients by SE-iFISH

Five-channel double-marker-iFISH was applied to perform phenotypic and karyotypic characterization of CTCs and CTECs enriched by SE from study subjects. Several typical cells of CTCs (CD31-/CD45-) and CTECs (CD31+/CD45-) found in NSCLC patients has been showed in Figure 1.

In Fig.1A, CTCs and CTECs in NSCLC patients has been shown from 1A-a to 1A-d. Fig.1A–a, a small triploid CTC displayed a DAPI+/CD45-/CD31-/CK18+ phenotype (≤5 μm WBC). Displayed in Fig. 1A–b, a large multiploid CTC was DAPI+/CD45-/CD31-/CK18- phenotype. A CD31+ large multiploid CTEC and a CD31+ CTEC cluster were shown in Fig. 1A–c/d, respectively. In Fig.1B, circulating tumor microemboli (CTM) displayed DAPI+/CD45-/CD31-/CK18- phenotype. Furthermore, we also found multiploid CTECs in NSCLC, displayed DAPI+/CD45-/CD31+/PD-L1+.

3.2 Comprehensive analysis of CTCs and CTECs in peripheral blood between groups

The CTCs detection rate was 97.9% (48/49) in the NSCLC group, 83.9% (26/31) and 95.8% (23/24) in the benign lung disease group and the healthy lung disease group, respectively. In the NSCLC group, only two individuals were detected with CK18+ CTCs, with a tumor marker positive detection rate of 4.2% (2/48), and the rest were tumor marker negative CTCs, 95.8% (46/48). There were no tumor-marked positive circulating tumor cells in the control group. The detection rate of CTECs was 98.0% (48/49) in NSCLC group, 87.1% (27/31) and 87.5% (21/24) in the benign group and the healthy group, respectively. One CTM was found in each of the 4 patients in the NSCLC group, and a total of 4 CTM were found, while no CTM was found in the other groups.

The results showed that the number of total CTCs or CTECs in the peripheral blood of NSCLC group was significantly higher than that of the benign lung disease group ($P<0.001$, $P<0.01$, Fig. 2A and D). Similarly, the number of CTCs or CTECs in different size in patients with NSCLC was significantly increased compared with that in patients with benign lung disease ($_{L}$CTCs: $P<0.05$; $_{S}$CTCs: $P<0.001$; $_{L}$CTECs: $P<0.01$; $_{S}$CTECs: $P<0.001$, Fig. 2B, C, E and F, respectively). In addition, no significant differences of CTCs and CTECs were found between healthy individuals and patients with benign lung disease, except for the small CTECs ($_{S}$CTECs: $P<0.01$, Fig. 2E). Either the total CTCs and CTECs, or CTCs and CTECs in different sizes were significantly higher in all

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

pathological subtype groups than in the benign lung disease group, except for the large CTECs (Fig. 3), while there was no statistically significant difference between lung adenocarcinoma, lung squamous cell carcinoma, and other NSCLC (Fig. 3, Supplemental Table 1). I stage, III stage, and IV stage NSCLC patients had a significantly higher number of CTCs than benign lung disease group, CTECs in I and IV stage NSCLC patients is significantly higher than benign lung disease group (Fig. 4A and D). Similar results could be obtained, when compared CTCs and CTECs in different sizes with the benign lung disease group, but there was no significant difference between TNM stages (Fig. 4B, C, E and F, Supplemental Table 1). Furthermore ,when we compared the proportions of CTCs and CTECs in different groups, we found that the proportion of small CTCs and CTECs significantly increased from benign lung disease individuals to NSCLC patients (Fig. 5).

3.3 Analysis of aneuploid characteristics of CTCs and CTECs between groups

In order to further elucidate the differences of CTCs and CTECs with different chromosomal ploidy numbers among different groups in number and proportion. Three types of chromosome aneuploidy were identified: triploid, tetraploid and multiploid (CEP 8 signal ≥5). Data analysis and intergroup comparison were performed, respectively.

The results showed that CTCs and CTECs of all aneuploid subtypes in NSCLC group and NSCLC I stage group were significantly higher than those in benign lung disease group, except for the multiploid CTCs in NSCLC I stage group. Compared with the healthy group, the CTECs triploid and tetraploid were significantly higher in the benign lung disease group, while the number of CTCs polyploid was lower in the benign lung disease group than in the healthy group (Fig. 6A). In addition, the number of triploid and tetraploid CTECs in adenosquamous carcinoma was significantly higher than that in adenocarcinoma and squamous carcinoma, according to Mann-Whitney U test. Stage IV had significant higher number of CTECs than that in stage I (P=0.0276*, Mann-Whitney U test), shown in Supplemental Table 2. However, according to the Kruskal-Wallis test, no significant difference had been found in number of CTCs and CTECs among aneuploid subtypes and TNM stages. According to the proportion analysis, we found that the proportion of triploid/tetraploid CTCs and CTECs in benign lung disease group were significantly higher than those in healthy group. In addition, we found an increased trend that the proportion of all aneuploid subtypes of CTCs and CTECs from benign lung disease individuals to NSCLC patients (Fig. 6B).

3.4 ROC curve analysis to determine the cut-off value and assess the diagnostic performance of NSCLC

Compared with benign lung disease group, the ROC curves were established based on the results of total and different subtypes of CTCs and CTECs in NSCLC group (Fig.7 A-E). The area under ROC curve (AUC)>0.7: Total CTCs=$_S$CTCs (0.815) > triploid CTECs (0.809) > sCTECs (0.781) > Total CTECs (0.739) > triploid CTCs (0.716), other statistical parameters of ROC analysis were shown in Table 2. Furthermore, according to morphology and aneuploid subtypes, we chose 6 items to combined whose area under ROC curve (AUC)>0.7. ROC curves were combined through the logistic regression analysis results, the combined formulas were calculated as: Z(CTCs+CTECs)=-1.499+0.148*CTCs+0.020*CTECs, Z($_S$CTCs+$_S$CTECs)=-1.983+0.250*$_S$CTCs+0.321*$_S$CTECs, Z triploid(CTCs+CTECs)=-2.151+0.358*CTCs+0.342*CTECs. After combined, the area under the ROC curves increased, indicating that the combined items were more effective. The critical value corresponding to the maximum point of the Jordan index was used to determine the cut-off values,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

shown in Table 2. These cut-off values were further used as the standards to evaluate the sensitivity, specificity and other diagnostic efficiency indicators of these detection items. As a comparison, we selected three clinically common serum tumor markers associated with non-small cell lung cancer for serological detection [29], as well as CT or PET-CT imaging diagnostic results (Table 3). Results showed that the sensitivity and specificity of total CTCs and CTECs for NSCLC were 67.3% and 77.6%, 83.9% and 77.4%, respectively. Among the rest detection items shown in Table 3, small CTCs and triploid CTCs had high specificity in the diagnosis of NSCLC, which were 93.5% and 96.8% respectively. In addition, through the evaluation of the diagnostic efficiency of these items for patients with different pathological subtypes and TNM stages, the result indicated the sensitivity of total CTCs or CTECs was relatively high, which had been shown in supplemental Table 3. Especially for NSCLC patients in stage I, the sensitivity of total CTCs or CTECs was both 70.4%, which were apparently higher than those of multiple tumor markers 29.2% and imaging diagnostic results 37%, respectively. Similarly, small CTCs and triploid CTCs had high specificity in the diagnosis of NSCLC I stage, shown in Table 4. In particular, the positive predictive value (PPV) and positive likelihood ratio (LR+) of these two items were also high, which means that they can be used as ideal screening indicators for high-risk lung cancer population in early stage, such as smoking, chronic lung disease or genetic history of lung cancer.

4. Discussion

Since 1869, Thomas Ashworth first defined CTCs as free tumor cells shed from the primary cancer [30], a variety of technologies were developed around the world to detect them. Because the concentration of CTCs in the peripheral blood of healthy people is extremely low, and the probability of CTCs appearing is about 1 CTC per 10^9 blood cells [31], it requires high sensitivity and accuracy of the technology to detect. In recent decades, more and more researches focused on investigating the potential of CTCs in cancer diagnostic and prognostic in clinic practice.

The effectiveness of the separation and identification determines the sensitivity and specificity of CTC detection. At present, the techniques used to sort CTCs mainly include negative enrichment and positive enrichment. As a positive enrichment technology ,which were widely recognized and applied in the detection of lung cancer CTCs, the CellSearch® system isolates and enriches CTCs based on the EpCAM expression [24]. However, the biggest limitation of such methods is that they ignore the intrinsic biological characteristics of tumor cells-high heterogeneity of EpCAM expression. The expression level of EpCAM varies in different tumors. For instance, EpCAM expression was low in bladder cancer, pancreatic cancer, lung cancer, or even not expressed in melanoma [32]. Especially in NSCLC, by down-regulating their epithelial characteristics and up-regulating their stem-like characteristics, epithelial tumor cells enhance their mobility and invasiveness, thus entering the blood circulation and colonizing distant organs to form metastases [33]. Another way is based on the cell size and plasticity to sort CTCs, such as: isolation by size of epithelial tumor cells (ISET) (RareCell Diagnostics, France), Parsortix (Angle Plc, UK) and ClearCell1 (Clearbridge BioMedics Pte Ltd, Singapore). The advantage of these techniques is independent on circulating tumor cell surface antigen, to maintain the morphology and activity of circulating tumor cells. However, with low-purity of the sorting cells, and the circulating tumor cells with small size are easily lost [34]. Furthermore, due to the high expression of folic acid receptor on the surface of non-small cell lung cancer cells [35], more and more clinical laboratories in China have applied Folate receptor (FR) based polymerase chain reaction (PCR) method to detect CTCs

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

in peripheral blood of NSCLC patients [14, 36-38]. However, all above methods cannot detect aneuploid circulating tumor cells.

In our research, we applied subtraction enrichment-immunostaining fluorescence *in situ* hybridization (SE-iFISH) to systematically detect and differentiate CTCs and CTECs in patients with non-small cell lung cancer. By using a variety of anti-leukocyte antibody coated magnetic beads, SE-iFISH ensures maximum white blood cell removal and minimal nonspecific adherence to tumor cells, well overcoming technical limitations that rely on expression of tumor cell surface markers to screen circulating tumor cells. On the other hand, compared with the use of immunofluorescence staining or FISH method alone, SE-iFISH can effectively detect single abnormal or double abnormal CTCs of chromosome or tumor biomarker, thus greatly improving the sensitivity and specificity of CTCs detection. Furthermore, the ability to detect circulating tumor endothelial cell aneuploidy gives SE-iFISH an advantage over other methods. As tumor endothelial cells (TECs) shed from solid tumor into peripheral circulating blood, tumor endothelial cells (CTECs) often exist in the form of aneuploidy, which has extremely critical significance in the process of tumor angiogenesis and metastasis [39, 40].

Numerous studies have confirmed the presence of circulating tumor cells in the peripheral blood of patients with both early and late stage NSCLC [13, 14, 41, 42]. In this study, we compared CTCs and CTECs levels in peripheral blood of non-small cell patients, benign lung diseases, and healthy individuals. Statistical analysis showed that there was no significant difference in CTCs and CTECs levels between the lung benign disease group and the healthy group. CTCs and CTECs levels of NSCLC patients was significantly higher than that of benign lung diseases and healthy group, which is consistent with previous studies [13, 14, 43]. Among the patients with different pathological types of NSCLC in this study, there was no statistically significant difference between pathological types, which is consistent with previous studies [13, 44]. In NSCLC patients with different stages, there was no significant difference between TNM stages, which is different from results of previous researches. Their results showed that CTCs in IV stage lung cancer patients were significantly higher than I stage patients [14, 38]. There are two possible explanations for this discrepancy. On the one hand, because of the small number of cases of this study, we found only an increase trend of CTCs and CTECs between I and IV stage, but no significant difference. On the other hand, in the early stage of NSCLC patients, a considerable amount of tumor cells or tumor vascular endothelial cells may shed off into the peripheral blood circulation. These circulating rare cells are likely to be chromosomal aneuploidy that cannot be identified using folate receptor-based PCR, and in this study were detected by the more sensitive and specific SE-iFISH.

Chromosomal aneuploidy is a common biological feature of tumor cells. About 90% solid malignant tumors and 75% blood tumors have chromosome aneuploidy of tumor cells, especially in tumor cells of lung cancer, stomach cancer, colon cancer, liver cancer, esophageal cancer and other solid tumors [11]. In addition, there were also aneuploidy of chromosome 8 in CTCs and CTECs in peripheral blood of tumor patients [19, 20]. Among all CTCs detected in this study, only five tumor-marked positive CTCs were found in NSCLC patients, and the rest were all tumor-marked negative expression of chromosome 8 aneuploidy, indicating that most circulating tumor cells in peripheral blood were tumor-marked negative. In addition, these non-hematologic aneuploid cells were not only present in the peripheral blood of patients with NSCLC, but also in small amounts in healthy people, and the number of these abnormal cells in the peripheral blood of NSCLC patients was significantly higher than that of healthy people. Similarly, most of the CTECs

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

in the form of aneuploidy in the peripheral blood of NSCLC patients were also found in the peripheral blood of healthy people, and the level was significantly lower than that of patients with NSCLC, which was the same as the previous research results of P. Lin's team [19]. In order to characterize a full spectrum of aneuploid CTCs and CTECs in NSCLC patients, CTCs and CTECs were analyzed in more detail according to the chromosome ploidy number. Results showed that there were statistically significant differences in the number of triploids and tetraploids of CTCs, as well as in the number of triploids, tetraploids and multiploids of CTECs between the lung benign disease group and the stage I lung cancer group. However, compare to the amount analysis, no significant differences in the proportion of karyotype in CTCs or CTECs between the above two groups were revealed. Therefore, we believe that CTC aneuploids can be used as an auxiliary method for clinical differentiation of benign lung diseases and early lung cancer. Surprisingly, between healthy and lung benign disease group, significant differences were found not only in the number of cells but also in the proportion of karyotypes in CTCs or CTECs. As for the source of non-hematologic aneuploid cells in the peripheral blood in normal people, some research suggested that occasional CD31+ aneuploid circulating endothelial cells in the blood of normal people are mainly from stromal cell trans-differentiation or cell aging, which is unrelated to tumor and can be cleared by the body's immune system [43]. But why did these non-hematologic aneuploid cells in the peripheral blood increased, when suffered benign lung disease such as pulmonary nodules, remains further research.

ROC curve analysis showed that sensitivity and specificity of total CTCs and CTECs being 67.3% and 83.9%, 77.6% and 77.4%, respectively for differential diagnosis of NSCLC, which are obviously superior to serological tumor marker and imaging diagnosis. Furthermore, for the NSCLC I stage, the sensitivity of total CTCs and CTECs are both 70.4%, which means both of them can be used as an important basis for early diagnosis of NSCLC. Surprisingly, small CTCs and triploid CTCs had been found to have high positive predictive value (PPV) and positive likelihood ratio (LR+) in the diagnosis of early stage NSCLC, which has obvious diagnostic advantages in screening high-risk lung cancer population in early stage compared with other routine clinical examinations.

Many studies have confirmed that aneuploid CTECs and CTCs in peripheral blood has important clinical significance in the prognosis and treatment of cancer. For example, CTCs with high ploidy have been shown to be associated with tumor resistance and relapse [28]. Jingyao Li et al. found that patients in stage IIIA lung cancer with CTCs≥0.5 had significantly lower DFS than those with CTCs<0.5 [13]. In addition, with the concept of precision medicine emerging, more and more attention has been paid to the research of tumor immunotherapy. In these studies, PD1/PD-L1 is the most popular target in immunotherapy of lung cancer. Many studies have shown that the expression level of PD-1/PD-L1 in NSCLC tissues is significantly correlated with the prognosis of patients [45-48]. However, few studies on the expression of PD-L1 on circulating rare cells in peripheral blood and the significance for the prognosis of NSCLC immunotherapy have been reported. Recent study indicated that, PD-L1+ CTCs and CTECs were associated with poor prognosis [20, 49]. In our research, we found two PD-L1+ CTECs existed in the peripheral blood of a patient with stage I lung adenocarcinoma, which means PD-L1+ CTECs are not only exist in patients with advanced lung cancer, but also in early stage lung cancer. However, due to our small sample size, we were unable to obtain the positive rate of PD-L1+ CTECs in patients with early lung cancer, which needs further study.

 doi:10.20944/preprints202101.0129.v1

5. Conclusions

In conclusion, the liquid biopsy technique of SE-iFISH to detect CTCs and CTECs has good sensitivity and specificity for the diagnosis of NSCLC, which are ideal biomarkers for NSCLC. In the context of accurate diagnosis and individualized treatment, this highly specific, sensitive and non-invasive examination is of great significance for the early diagnosis of NSCLC, and will have greater application space in the early diagnosis and treatment of other cancers in the future.

6. References

[1] J. Ferlay, H.R. Shin, F. Bray, D. Forman, C. Mathers, D.M. Parkin, Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008, Int J Cancer, 127 (2010) 2893-2917.

[2] R.S. Herbst, J.V. Heymach, S.M. Lippman, Lung cancer, N Engl J Med, 359 (2008) 1367-1380.

[3] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2015, CA Cancer J Clin, 65 (2015) 5-29.

[4] E.F. Patz, Jr., M.J. Campa, E.B. Gottlin, I. Kusmartseva, X.R. Guan, J.E. Herndon, 2nd, Panel of serum biomarkers for the diagnosis of lung cancer, J Clin Oncol, 25 (2007) 5578-5583.

[5] L. Kasmann, L. Bolm, S. Janssen, D. Rades, Prognostic Factors and Treatment of Early-stage Small-cell Lung Cancer, Anticancer Res, 37 (2017) 1535-1537.

[6] A.E. Revelo, A. Martin, R. Velasquez, P.C. Kulandaisamy, J. Bustamante, S. Keshishyan, G. Otterson, Liquid biopsy for lung cancers: an update on recent developments, Ann Transl Med, 7 (2019) 349.

[7] N. Ramalingam, S.S. Jeffrey, Future of Liquid Biopsies With Growing Technological and Bioinformatics Studies: Opportunities and Challenges in Discovering Tumor Heterogeneity With Single-Cell Level Analysis, Cancer J, 24 (2018) 104-108.

[8] F.C. Bidard, C. Proudhon, J.Y. Pierga, Circulating tumor cells in breast cancer, Mol Oncol, 10 (2016) 418-430.

[9] T. Zhang, A.J. Armstrong, Clinical Utility of Circulating Tumor Cells in Advanced Prostate Cancer, Curr Oncol Rep, 18 (2016) 3.

[10] S. Hinz, J. Bockhorst, C. Roder, J.H. Egberts, C. Schafmayer, T. Kuchler, T. Becker, H. Kalthoff, Disseminated tumor cells in the bone marrow negatively influence survival after resection of colorectal liver metastases, Ann Surg Oncol, 19 (2012) 2539-2546.

[11] P.P. Lin, Aneuploid CTC and CEC, Diagnostics (Basel), 8 (2018).

[12] L. Harris, H. Fritsche, R. Mennel, L. Norton, P. Ravdin, S. Taube, M.R. Somerfield, D.F. Hayes, R.C. Bast, Jr., O. American Society of Clinical, American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer, J Clin Oncol, 25 (2007) 5287-5312.

[13] J. Li, Y. Liao, Y. Ran, G. Wang, W. Wu, Y. Qiu, J. Liu, N. Wen, T. Jing, H. Wang, S. Zhang, Evaluation of sensitivity and specificity of CanPatrol technology for detection of circulating tumor cells in patients with non-small cell lung cancer, BMC Pulm Med,

20 (2020) 274.

[14] Q.L. Guo Qiaomei, Wang Lin,Lou Jiatao, Diagnostic value of human circulating tumor cell detection for non-small cell lung cancer, Chinese Journal of Laboratory Medicine 39 (2016) 589-594.

[15] S.Q. Wang, Z.F. Shuai, X.J. Zhang, T. Wu, H.Y. Dong, T. Liu, Q.T. Wen, X.W. Yu, Detection of CTCs and CSCs in the staging and metastasis of non-small cell lung cancer based on microfluidic chip and the diagnostic significance, Eur Rev Med Pharmacol Sci, 24 (2020) 9487-9496.

[16] J. Zhou, F. Dong, F. Cui, R. Xu, X. Tang, The role of circulating tumor cells in evaluation of prognosis and treatment response in advanced non-small-cell lung cancer, Cancer Chemother Pharmacol, 79 (2017) 825-833.

[17] F. Bertolini, Y. Shaked, P. Mancuso, R.S. Kerbel, The multifaceted circulating endothelial cell in cancer: towards marker and target identification, Nat Rev Cancer, 6 (2006) 835-845.

[18] P.P. Lin, Aneuploid Circulating Tumor-Derived Endothelial Cell (CTEC): A Novel Versatile Player in Tumor Neovascularization and Cancer Metastasis, Cells, 9 (2020).

[19] P.P. Lin, O. Gires, D.D. Wang, L. Li, H. Wang, Comprehensive in situ co-detection of aneuploid circulating endothelial and tumor cells, Sci Rep, 7 (2017) 9789.

[20] L. Zhang, X. Zhang, Y. Liu, T. Zhang, Z. Wang, M. Gu, Y. Li, D.D. Wang, W. Li, P.P. Lin, PD-L1(+) aneuploid circulating tumor endothelial cells (CTECs) exhibit resistance to the checkpoint blockade immunotherapy in advanced NSCLC patients, Cancer Lett, 469 (2020) 355-366.

[21] M. Yu, S. Stott, M. Toner, S. Maheswaran, D.A. Haber, Circulating tumor cells: approaches to isolation and characterization, J Cell Biol, 192 (2011) 373-382.

[22] P.K. Grover, A.G. Cummins, T.J. Price, I.C. Roberts-Thomson, J.E. Hardingham, Circulating tumour cells: the evolving concept and the inadequacy of their enrichment by EpCAM-based methodology for basic and clinical cancer research, Ann Oncol, 25 (2014) 1506-1516.

[23] S.D. Mikolajczyk, L.S. Millar, P. Tsinberg, S.M. Coutts, M. Zomorrodi, T. Pham, F.Z. Bischoff, T.J. Pircher, Detection of EpCAM-Negative and Cytokeratin-Negative Circulating Tumor Cells in Peripheral Blood, J Oncol, 2011 (2011) 252361.

[24] A. Tartarone, E. Rossi, R. Lerose, G. Mambella, G. Calderone, R. Zamarchi, M. Aieta, Possible applications of circulating tumor cells in patients with non small cell lung cancer, Lung Cancer, 107 (2017) 59-64.

[25] P.T. Went, A. Lugli, S. Meier, M. Bundi, M. Mirlacher, G. Sauter, S. Dirnhofer, Frequent EpCam protein expression in human carcinomas, Hum Pathol, 35 (2004) 122-128.

[26] R. Kalluri, EMT: when epithelial cells decide to become mesenchymal-like cells, J Clin Invest, 119 (2009) 1417-1419.

[27] Y. Li, X. Zhang, J. Gong, Q. Zhang, J. Gao, Y. Cao, D.D. Wang, P.P. Lin, L. Shen, Aneuploidy of chromosome 8 in circulating tumor cells correlates with prognosis in patients with advanced gastric cancer, Chin J Cancer Res, 28 (2016) 579-588.

[28] J. Zhang, H. Shi, T. Jiang, Z. Liu, P.P. Lin, N. Chen, Circulating tumor cells with karyotyping as a novel biomarker for diagnosis and treatment of nasopharyngeal carcinoma, BMC Cancer, 18 (2018) 1133.

[29] H. Wu, Q. Wang, Q. Liu, Q. Zhang, Q. Huang, Z. Yu, The Serum Tumor Markers in Combination for Clinical Diagnosis of Lung Cancer, Clin Lab, 66 (2020).

[30] A.T. R, A case of cancer in which cells similar to those in the tumors were seen in the blood after death[J], Australian Med J, (1869) 146-149.

[31] H.R. Williams A, Zheng S, Albanese C, Schlegel R, Tai Y-C, Ram H. Datar, Richard J, Perspectives on the functional characterization and in vitro maintenance of circulating tumor cells[J], Circulating Tumor Cell, 4 (2016) 215-231.

[32] F. Ge, H. Zhang, D.D. Wang, L. Li, P.P. Lin, Enhanced detection and comprehensive in situ phenotypic characterization of circulating and disseminated heteroploid epithelial and glioma tumor cells, Oncotarget, 6 (2015) 27049-27064.

[33] N. Bednarz-Knoll, C. Alix-Panabieres, K. Pantel, Plasticity of disseminating cancer cells in patients with epithelial malignancies, Cancer Metastasis Rev, 31 (2012) 673-687.

[34] M. Alunni-Fabbroni, M.T. Sandri, Circulating tumour cells in clinical practice: Methods of detection and possible characterization, Methods, 50 (2010) 289-297.

[35] L.E. Kelemen, The role of folate receptor alpha in cancer development, progression and treatment: cause, consequence or innocent bystander?, Int J Cancer, 119 (2006) 243-250.

[36] Q. Zhou, Q. Geng, L. Wang, J. Huang, M. Liao, Y. Li, Z. Ding, S. Yang, H. Zhao, Q. Shen, C. Pan, J. Lou, S. Lu, C. Chen, Q. Luo, Value of folate receptor-positive circulating tumour cells in the clinical management of indeterminate lung nodules: A non-invasive biomarker for predicting malignancy and tumour invasiveness, EBioMedicine, 41 (2019) 236-243.

[37] J. Shen, J. Zhao, T. Jiang, X. Li, C. Zhao, C. Su, C. Zhou, Predictive and prognostic value of folate receptor-positive circulating tumor cells in small cell lung cancer patients treated with first-line chemotherapy, Oncotarget, 8 (2017) 49044-49052.

[38] X. Chen, F. Zhou, X. Li, G. Yang, L. Zhang, S. Ren, C. Zhao, Q. Deng, W. Li, G. Gao, A. Li, C. Zhou, Folate Receptor-Positive Circulating Tumor Cell Detected by LT-PCR-Based Method as a Diagnostic Biomarker for Non-Small-Cell Lung Cancer, J Thorac Oncol, 10 (2015) 1163-1171.

[39] A. Dart, Metastasis: RIP endothelial cells, Nat Rev Cancer, 16 (2016) 551.

[40] V. Platel, S. Faure, I. Corre, N. Clere, Endothelial-to-Mesenchymal Transition (EndoMT): Roles in Tumorigenesis, Metastatic Extravasation and Therapy Resistance, J Oncol, 2019 (2019) 8361945.

[41] Q. Chen, F. Ge, W. Cui, F. Wang, Z. Yang, Y. Guo, L. Li, R.M. Bremner, P.P. Lin, Lung cancer circulating tumor cells isolated by the EpCAM-independent enrichment strategy correlate with Cytokeratin 19-derived CYFRA21-1 and pathological staging, Clin Chim Acta, 419 (2013) 57-61.

[42] C. Wu, H. Hao, L. Li, X. Zhou, Z. Guo, L. Zhang, X. Zhang, W. Zhong, H. Guo, R.M. Bremner, P. Lin, Preliminary investigation of the clinical significance of detecting circulating tumor cells enriched from lung cancer patients, J Thorac Oncol, 4 (2009) 30-36.

[43] Y. Lei, N. Sun, G. Zhang, C. Liu, Z. Lu, J. Huang, C. Zhang, R. Zang, Y. Che, S. Mao, L. Fang, X. Wang, S. Zheng, J. He, Combined detection of aneuploid circulating tumor-

 doi:10.20944/preprints202101.0129.v1

derived endothelial cells and circulating tumor cells may improve diagnosis of early stage non-small-cell lung cancer, Clin Transl Med, 10 (2020) e128.

[44] V. Murlidhar, R.M. Reddy, S. Fouladdel, L. Zhao, M.K. Ishikawa, S. Grabauskiene, Z. Zhang, J. Lin, A.C. Chang, P. Carrott, W.R. Lynch, M.B. Orringer, C. Kumar-Sinha, N. Palanisamy, D.G. Beer, M.S. Wicha, N. Ramnath, E. Azizi, S. Nagrath, Poor Prognosis Indicated by Venous Circulating Tumor Cell Clusters in Early-Stage Lung Cancers, Cancer Res, 77 (2017) 5194-5206.

[45] W.A. Cooper, T. Tran, R.E. Vilain, J. Madore, C.I. Selinger, M. Kohonen-Corish, P. Yip, B. Yu, S.A. O'Toole, B.C. McCaughan, J.H. Yearley, L.G. Horvath, S. Kao, M. Boyer, R.A. Scolyer, PD-L1 expression is a favorable prognostic factor in early stage non-small cell carcinoma, Lung Cancer, 89 (2015) 181-188.

[46] H. Li, Y. Xu, B. Wan, Y. Song, P. Zhan, Y. Hu, Q. Zhang, F. Zhang, H. Liu, T. Li, H. Sugimura, F. Cappuzzo, D. Lin, T. Lv, A.M.E.L.C.C.G. written on behalf of, The clinicopathological and prognostic significance of PD-L1 expression assessed by immunohistochemistry in lung cancer: a meta-analysis of 50 studies with 11,383 patients, Transl Lung Cancer Res, 8 (2019) 429-449.

[47] S. Igawa, Y. Sato, S. Ryuge, M. Ichinoe, K. Katono, Y. Hiyoshi, S. Otani, R. Nagashio, H. Nakashima, M. Katagiri, J. Sasaki, Y. Murakumo, Y. Satoh, N. Masuda, Impact of PD-L1 Expression in Patients with Surgically Resected Non-Small-Cell Lung Cancer, Oncology, 92 (2017) 283-290.

[48] M.D. Keller, C. Neppl, Y. Irmak, S.R. Hall, R.A. Schmid, R. Langer, S. Berezowska, Adverse prognostic value of PD-L1 expression in primary resected pulmonary squamous cell carcinomas and paired mediastinal lymph node metastases, Mod Pathol, 31 (2018) 101-110.

[49] N. Guibert, M. Delaunay, A. Lusque, N. Boubekeur, I. Rouquette, E. Clermont, J. Mourlanette, S. Gouin, I. Dormoy, G. Favre, J. Mazieres, A. Pradines, PD-L1 expression in circulating tumor cells of advanced non-small cell lung cancer patients treated with nivolumab, Lung Cancer, 120 (2018) 108-112.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

Figure legend

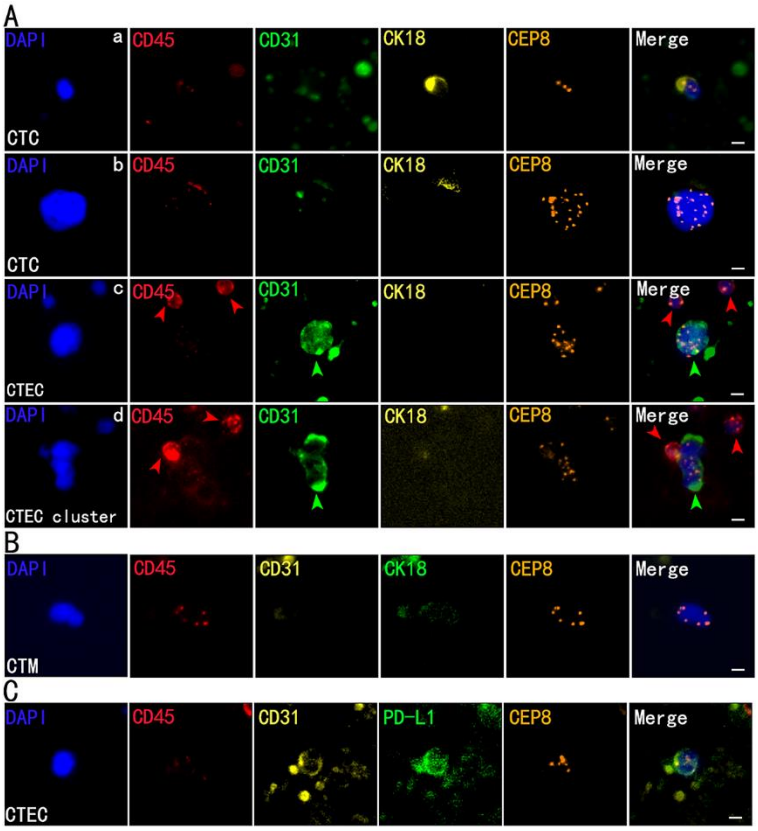

Fig1. Multi-Fluorescence of CTCs and CTECs: In situ single-tumor-biomarker SE-iFISH (five channel)：nucleus (blue) , CD45(red), CD31 (green/yellow), CK18 (green/yellow) and CEP8 (orange) **A.** CTCs and CTECs in NSCLC patients. **A-a**. DAPI+/CD45-/CD31-/CK18+/CEP8 triploid CTC. **A-b**. DAPI+/CD45-/CD31-/CK18-/CEP8 multiploid CTC. **A-c.** DAPI+/CD45-/CD31+/CK18-/CEP8 multiploid CTEC (green arrow) with adjacent CD45+ WBCs (red arrow). **A-d.** DAPI+/CD45-/CD31+/CK18-/CEP8 multiploid CTECs cluster (green arrow) with adjacent CD45+ WBCs (red arrow). **B.** DAPI+/CD45-/CD31-/CK18- CTM. **C.** DAPI+/CD45-/CD31+/PD-L1+/CEP8 multiploid CTEC. Bars: 5 μm.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

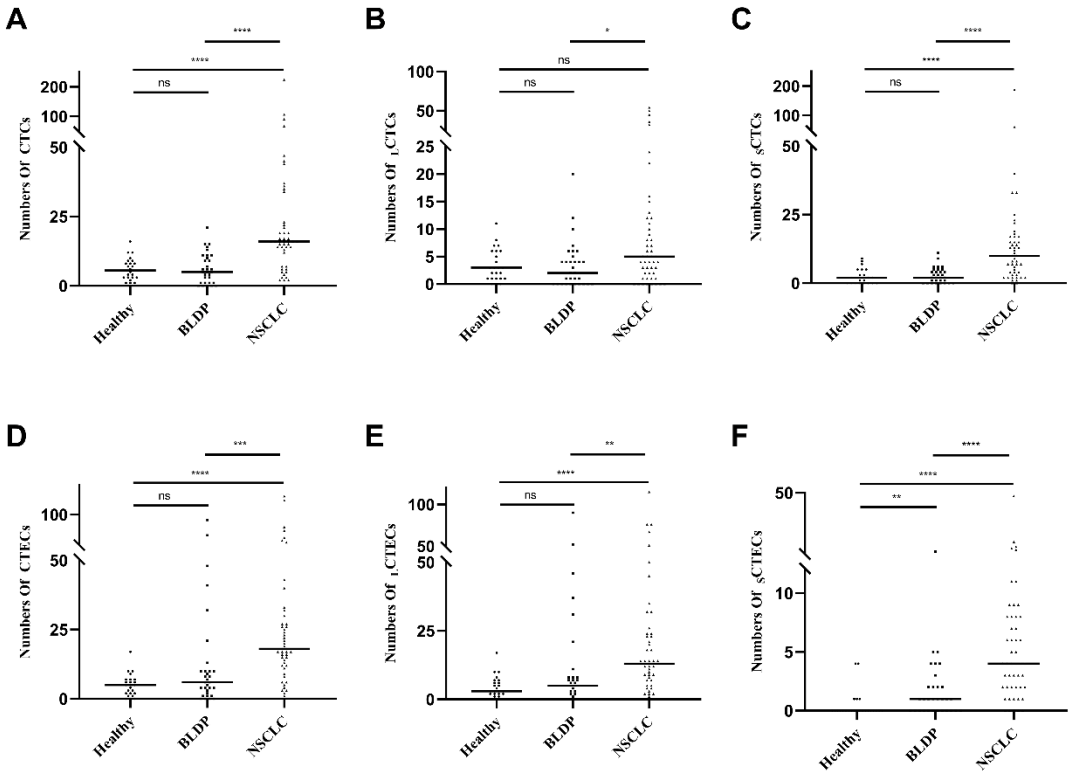

Fig 2. Comparison of CTCs and CTECs counts among healthy, benign lung disease and NSCLC groups. **A.** Comparison of total CTCs counts among healthy, benign lung disease and NSCLC group. **B.** Comparison of large CTCs (LCTCs) counts among healthy, benign lung disease and NSCLC group. **C.** Comparison of small CTCs (SCTCs) counts among healthy, benign lung disease and NSCLC group. **D.** Comparison of total CTECs counts among healthy, benign lung disease and NSCLC group. **E.** Comparison of LCTECs counts among healthy, benign lung disease and NSCLC group. **F.** Comparison of SCTECs counts among healthy, benign lung disease and NSCLC group. Count data are presented as the median, **P*<0.05, ***P*<0.01, ****P*<0.001, *****P*<0.0001, Mann-Whitney *U* test.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

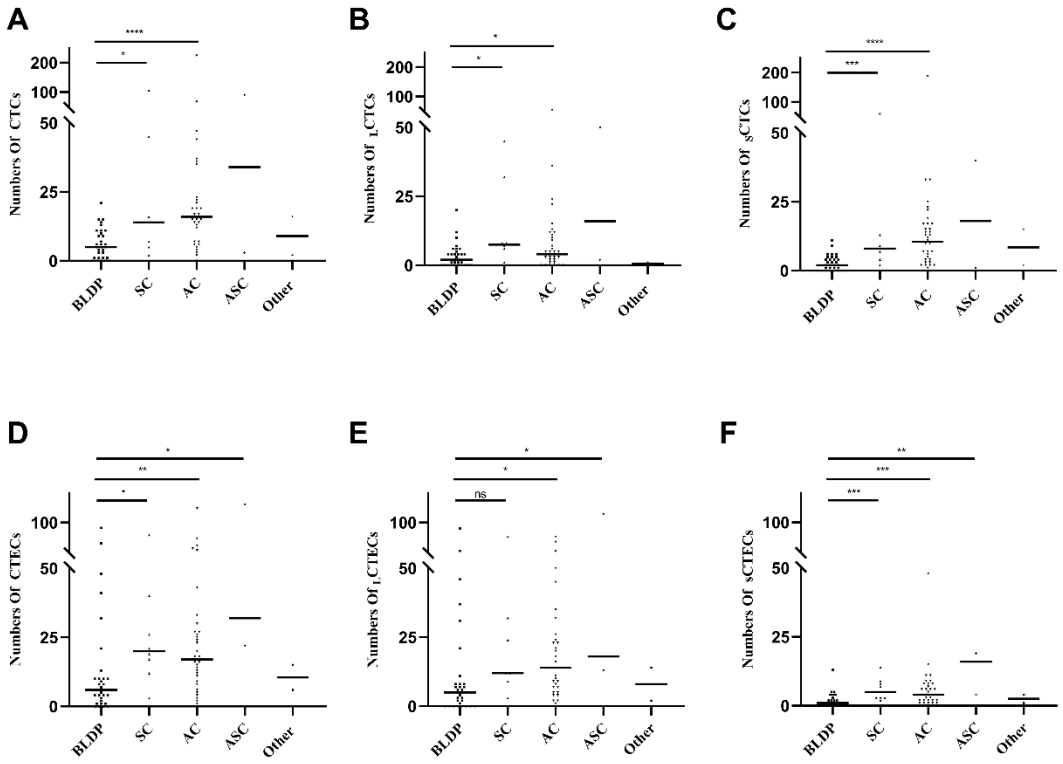

Fig 3. Comparison of CTCs and CTECs counts among patients of different pathological types. **A.** Comparison of total CTCs counts among benign lung disease patients (BLDP), squamous carcinoma (SC), adenocarcinoma(AC) , adenosquamous carcinoma (ASC) and other types. **B.** Comparison of large CTCs ($_L$CTCs) counts among BLDP, SC, AC, ASC and other types. **C.** Comparison of small CTCs ($_s$CTCs) counts among BLDP, SC, AC, ASC and other types. **D.** Comparison of total CTECs counts among BLDP, SC, AC, ASC and other types. **E.** Comparison of $_L$CTECs counts among BLDP, SC, AC, ASC and other types. **F.** Comparison of $_s$CTECs counts among BLDP, SC, AC, ASC and other types. Count data are presented as the median, **P*<0.05, ***P*<0.01, ****P*<0.001, *****P*<0.0001, Mann-Whitney *U* test.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

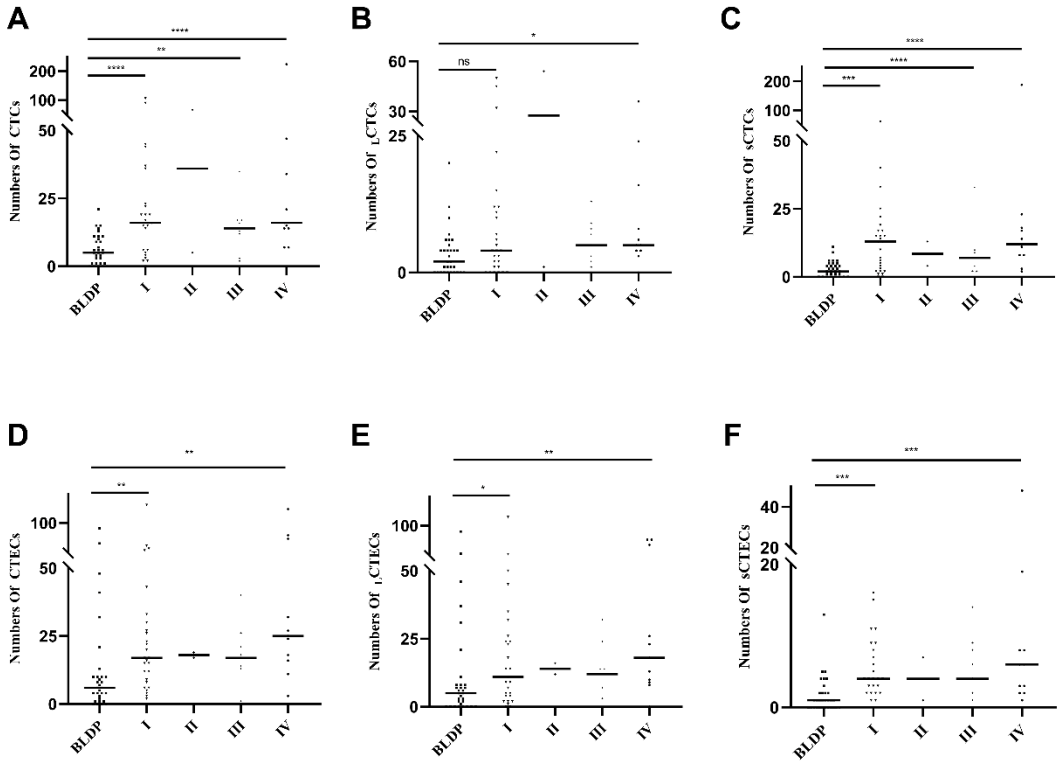

Fig 4. Comparison of CTCs and CTECs counts among patients in different stages. **A.** Comparison of total CTCs counts among benign lung disease patients (BLDP) and stage I-IV. **B.** Comparison of large CTCs ($_{L}$CTCs) counts among BLDP and stage I-IV. **C.** Comparison of small CTCs ($_{S}$CTCs) counts among BLDP and stage I-IV. **D.** Comparison of total CTECs counts between BLDP and stage I-IV. **E.** Comparison of $_{L}$CTECs counts among BLDP and stage I-IV. **F.** Comparison of $_{S}$CTECs counts among BLDP and stage I-IV. Count data are presented as the median, *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001, Mann-Whitney U test.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

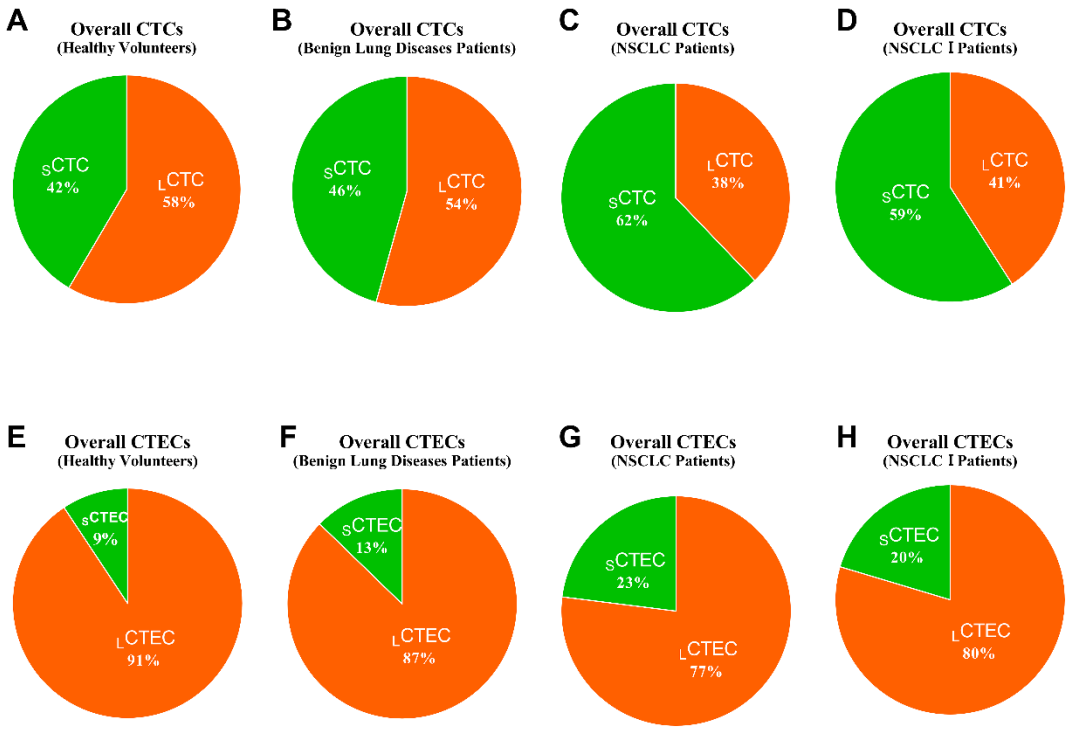

Fig 5. Compositional analysis of CTCs and CTECs in different sizes among groups (healthy, benign lung diseases, total NSCLC, NSCLC I stage). **A-C**. increased proportions of sCTCs from 42% to 62% (green) (****P <0.0001, Chi-Square test). **A-D**. increased proportions of sCTCs from 42% to 59% (green) (***P <0.001, Chi-Square test). **B-C**. increased proportions of sCTCs from 46% to 62% (green) (****P <0.0001, Chi-Square test). **B-D**. increased proportions of sCTCs from 46% to 59% (green) (**P=0.001, Chi-Square test). **E-G**. increased proportions of sCTECs from 9% to 23% (green) (***P <0.001, Chi-Square test). **E-H**. increased proportions of sCTECs from 9% to 20% (green) (**P=0.005, Chi-Square test).**F-G**. increased proportions of sCTECs from 13% to 23% (green) (****P <0.0001, Chi-Square test). **F-H**. increased proportions of sCTECs from 13% to 20% (green) (**P=0.001, Chi-Square test).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

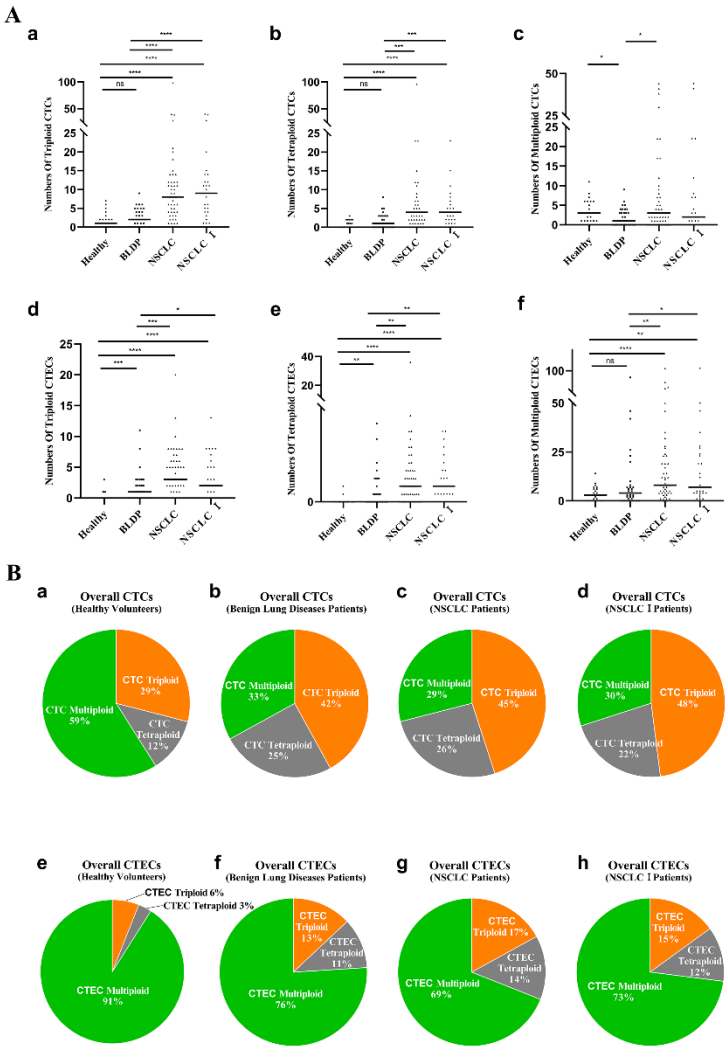

Fig6. Comprehensive analysis of aneuploid CTCs and CTECs changes among groups. **A(a-f)**. Quantification analysis of chromosomal ploidy in CTCs and CTECs among groups. (healthy, benign lung diseases, total NSCLC, NSCLC I stage). **B(a-h)**. Compositional analysis of chromosomal ploidy in CTCs and CTECs in different groups. **B(a-b):** increased proportions of triploid CTCs from 29% to 42% (orange) (**P*=0.016, Chi-Square test); increased proportions of tetraploid CTCs from 12% to 25% (grey) (***P*=0.0022, Chi-Square test); decreased proportions of multiploid CTCs from 59% to 33% (green) (****P*=0.0001, Chi-Square test); **B(e-f):** increased proportions of triploid CTECs from 6% to13% (orange) (**P*=0.023, Chi-Square test); increased proportions of tetraploid CTECs from 3% to 11% (grey) (**P*=0.015, Chi-Square test); decreased proportions of multiploid CTECs from 91% to 76% (green) (***P*=0.0014, Chi-Square test); **B(f-g):** decreased proportions of multiploid CTECs from 76% to 69% (green) (***P*=0.0079, Chi-Square test). Count data are presented as the mean ± SD, **P*<0.05, ***P*<0.01, ****P*<0.001, *****P*<0.0001. Blank means no statistical significance.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

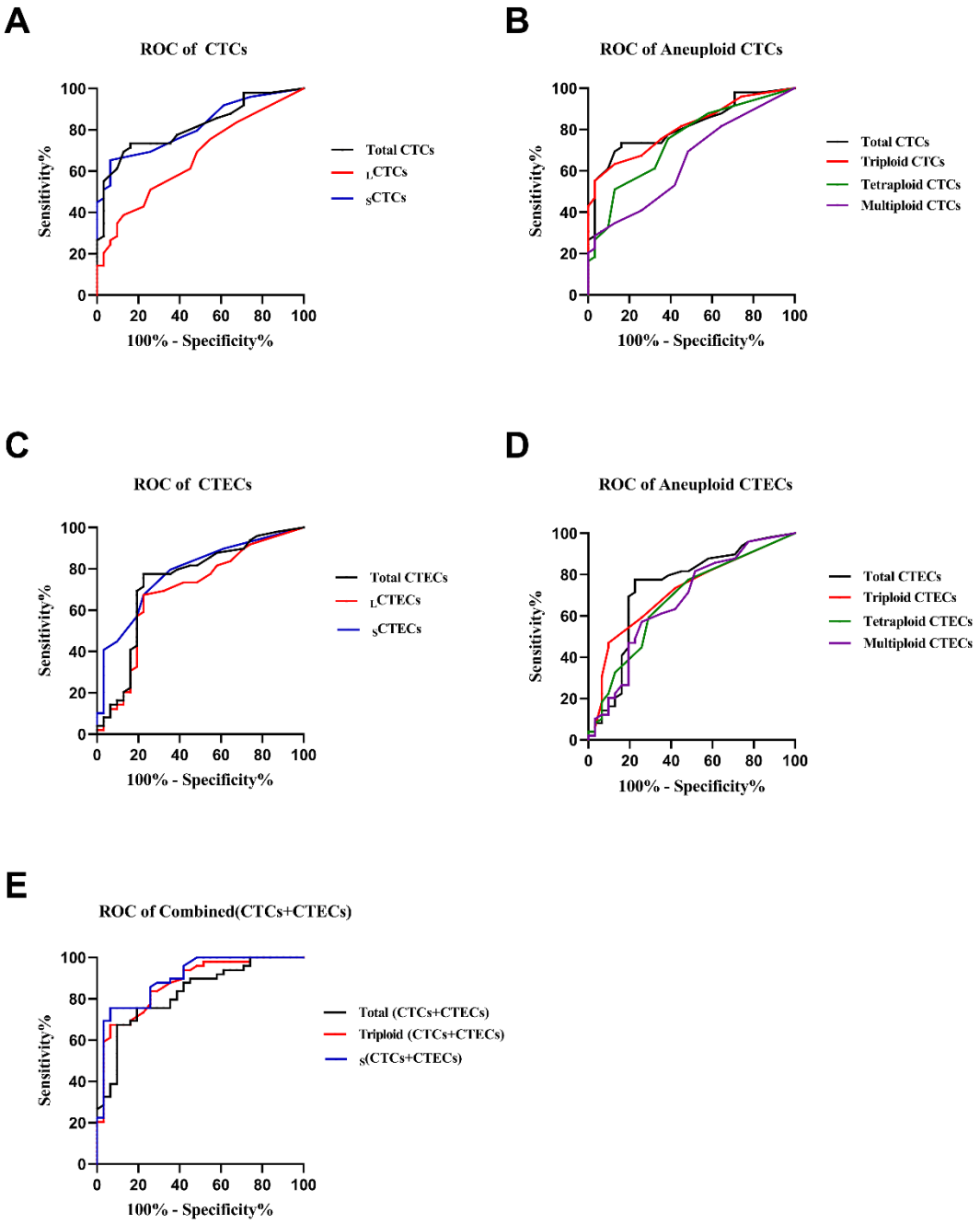

Fig 7. ROC analysis of CTCs and CTECs for NSCLC. **A.**ROC curves of total CTCs, large CTCs (LCTCs) and small CTCs (SCTCs) for NSCLC **B.** ROC curves of total CTECs, LCTECs and SCTECs for NSCLC **C.** ROC curves of total CTCs, triploid CTCs, tetraploid CTCs and multiploid CTCs for NSCLC **D.** ROC curves of total CTECs, triploid CTECs, tetraploid CTECs and multiploid CTECs for NSCLC **E.** ROC curves of combined detection: total (CTCs+ CTECs), SCTCs + SCTECs and triploid (CTCs + CTECs) for NSCLC.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

Table1. CK18 EpCAM or PD-L1 distribution among groups

	Total	CK18	EpCAM	PD-L1
NSCLC	49	37	9	3
Benign lung diseases	31	18	10	3
Health	24	17	4	3

Table 2. Statistical parameters for ROC analysis of CTCs and CTECs for NSCLC

Test item	AUC	Cutoff value	Std. Error	P	95% CI
Total CTCs	0.815	11.5	0.047	<0.0001	0.722-0.907
$_{L}$CTCs	0.663	6.5	0.061	0.015	0.544-0.782
$_{S}$CTCs	0.815	6.5	0.046	<0.0001	0.724-0.906
triploid CTCs	0.809	6.5	0.047	<0.0001	0.717-0.901
tetraploid CTCs	0.698	4.5	0.056	<0.001	0.634-0.852
multiploid CTCs	0.646	6.5	0.062	0.029	0.525-0.766
Total CTECs	0.739	10.5	0.062	<0.001	0.618-0.860
$_{L}$CTECs	0.691	8.5	0.064	0.004	0.566-0.815
$_{S}$CTECs	0.781	2.5	0.053	<0.0001	0.679-0.884
triploid CTECs	0.716	3.5	0.058	0.001	0.602-0.830
tetraploid CTECs	0.677	2.5	0.063	0.008	0.555-0.799
multiploid CTECs	0.675	7.5	0.063	0.009	0.550-0.799
Total CTCs+ CTECs	0.826	—	0.047	<0.0001	0.734-0.917
$_{S}$CTCs+ $_{S}$CTECs	0.898	—	0.036	<0.0001	0.828-0.968
triploid(CTCs+CTECs)	0.872	—	0.040	<0.0001	0.794-0.950

Abbreviations: Large CTCs, $_{L}$CTCs, small CTCs, $_{S}$CTCs, Large CTECs, $_{L}$CTECs, small CTECs, $_{S}$CTECs

Table3. Evaluation of the diagnostic efficiency of single and combined tests for NSCLC[%(n/n)]

Test item	SEN%	SPE%	PPV%	NPV%	AC%	LR+
Total CTCs	67.3	83.9	86.8	61.9	73.8	4.2
$_{L}$CTCs	38.8	67.7	65.5	41.2	50.0	1.2
$_{S}$CTCs	65.3	93.5	94.1	63.0	76.3	10.1
triploid CTCs	57.1	96.8	96.6	58.8	72.5	17.7
tetraploid CTCs	51.0	87.1	86.2	52.9	65.0	4.0
multiploid CTCs	28.6	96.8	93.3	46.2	55.0	8.9
Total CTECs	77.6	77.4	84.4	68.6	77.5	3.4
$_{L}$CTECs	67.3	67.7	76.7	56.8	67.5	2.1
$_{S}$CTECs	67.3	77.4	82.5	60.0	71.3	3.0
triploid CTECs	49.0	90.3	88.9	52.8	65.0	5.1
tetraploid CTECs	36.7	71.0	66.7	41.5	50.0	1.3
multiploid CTECs	59.2	74.2	78.4	53.5	65.0	2.3
SCC+CEA+CYFRA 21-1	52.2	75.8	77.4	50.0	61.3	2.17
CT/PET-CT Diagnosis	57.1	77.4	80.0	53.3	65.0	2.48

Abbreviations: Large CTCs, $_{L}$CTCs, small CTCs, $_{S}$CTCs, Large CTECs, $_{L}$CTECs, small CTECs, $_{S}$CTECs, SEN, Sensitivity, SPE, Specificity, PPV, Positive predictive value, NPV, Negative predictive value, LR+ , Positive likelihood ratio

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 January 2021 doi:10.20944/preprints202101.0129.v1

Table 4. Evaluation of the diagnostic efficiency of single and combined tests for NSCLC I stage [%(n/n)]

Test item	SEN%	SPE%	PPV%	NPV%	AC%	LR+
Total CTCs	70.4	87.1	82.6	77.1	79.3	5.5
$_{L}$CTCs	37.0	61.3	45.5	52.8	50.0	1.0
$_{S}$CTCs	59.3	93.5	88.9	72.5	77.6	9.2
triploid CTCs	59.3	96.8	94.1	73.2	79.3	18.4
tetraploid CTCs	29.6	87.1	66.7	58.7	60.3	2.3
multiploid CTCs	37.0	61.3	45.5	52.8	50.0	1.0
Total CTECs	70.4	77.4	73.1	75.0	74.1	3.1
$_{L}$CTECs	59.3	77.4	69.6	68.6	69.0	2.6
$_{S}$CTECs	66.7	77.4	72.0	72.7	72.4	3.0
triploid CTECs	40.7	90.3	78.6	63.6	67.2	4.2
tetraploid CTECs	59.3	71.0	64.0	66.7	65.5	2.0
multiploid CTECs	48.1	74.2	61.9	62.2	62.1	1.9
SCC+CEA+CYFRA 21-1	29.2	75.0	50.0	56.4	54.7	1.2
CT/PET-CT Diagnosis	37.0	77.4	59.1	58.5	58.6	1.6

Abbreviations: Large CTCs, $_{L}$CTCs, small CTCs, $_{S}$CTCs, Large CTECs, $_{L}$CTECs, small CTECs, $_{S}$CTECs, SEN, Sensitivity, SPE, Specificity, PPV, Positive predictive value, NPV, Negative predictive value, LR+ , Positive likelihood ratio