Comparison of autofluorescence and white-light bronchoscopies performed with the Evis Lucera Spectrum for the detection of bronchial cancers: a meta-analysis

Shuangshuang Sun¹,², Yang Yang², Meizi Chen³, Li Wang², Hangcheng Pan², Xiuwei Zhang², Georges Wagnieres⁴, Yousser Mohammad⁵,⁶, Esther Barreiro⁷, Giovanni Pirozzolo⁸, P. James Villeneuve⁹, Ping Zhan¹⁰,¹¹, Bing Wan²; written on behalf of the AME Lung Cancer Collaborative Group

¹Department of Respiratory Medicine, The Affiliated Hospital of Jiangsu University, Zhenjiang 212001, China; ²Department of Respiratory and Critical Care Medicine, the Affiliated Jiangning Hospital of Nanjing Medical University, Nanjing 210002, China; ³Department of Respiratory and Critical Care Medicine, Chenzhou No.1 People's Hospital, Chenzhou 423000, China; ⁴Laboratory for Functional and Metabolic Imaging, Institute of Physics, Swiss Federal Institute of Technology in Lausanne (EPFL), Station 6, Bâtiment de Chimie, Lausanne, Switzerland; ⁵National Center for Research on Chronic Respiratory Diseases, Tishreen University, Tishreen Hospital, Latakia, Syria; ⁶Department of Internal Medicine, 6-Syrian Private University, Damascus, Syria; ⁷Pulmonology Department-Muscle Wasting & Cachexia in Chronic Respiratory Diseases & Lung Cancer Research Group, IMIM-Hospital del Mar, Parc de Salut Mar, Barcelona, Spain; Centro de Investigación en Red de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III (ISCIII), Barcelona, Spain; ⁸General Surgery Department, dell’Angelo Hospital, Venice, Italy; ⁹Cancer Therapeutics Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada; ¹⁰Department of Respiratory Medicine, Jinling Hospital, Nanjing University of Medicine, Nanjing 210002, China; ¹¹Nanjing University Institute of Respiratory Medicine, Nanjing 210002, China

Contributions: (I) Conception and design: P Zhan, B Wan; (II) Administrative support: B Wan; (III) Provision of study materials or patients: S Sun, Y Yang, M Chen; (IV) Collection and assembly of data: L Wang, H Pan, X Zhang; (V) Data analysis and interpretation: S Sun, Y Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Bing Wan, MD. Department of Respiratory and Critical Care Medicine, the Affiliated Jiangning Hospital of Nanjing Medical University, Nanjing 210002, China Email: bingwan76@163.com; Ping Zhan, MD. Department of Respiratory Medicine, Jinling Hospital, Nanjing University of Medicine and Nanjing University Institute of Respiratory Medicine, Nanjing 210002, China. Email: zhanping207@163.com.

Background: Many recent studies have reported that autofluorescence bronchoscopy (AFB) has a superior sensitivity and decreased specificity in the diagnosis of bronchial cancers when compared with white-light bronchoscopy (WLB). We specifically analyzed the diagnostic performances of autofluorescence imaging video bronchoscopy (AFI) performed with the Evis Lucera Spectrum from Olympus, which is a relatively novel approach in detecting and delineating bronchial cancers, and compared it to the older WLB method.

Methods: We searched the PubMed, Embase, Web of Science, and CNKI databases from inception to July 12th, 2018 for trials in which patients were diagnosed with lung cancer via concurrent or combined use of AFI and WLB. The included studies were required to have a histologic diagnosis as the gold standard comparison, and a sufficient amount of data was extracted to assess the diagnostic capacity. A 2×2 table was constructed, and the area under the receiver-operating characteristic curve (AUC) of AFI and WLB was estimated by using a stochastic model for diagnostic meta-analysis using STATA software.

Results: A total of 10 articles were eligible for the meta analysis, comprising 1,830 patients with complete data included in the analysis. AFI showed a superior sensitivity of 0.92 (95% CI, 0.88–0.95) over WLB’s 0.70 (95% CI, 0.58–0.80) with P<0.01, and a comparable specificity of 0.67 (95% CI, 0.51–0.80) compared with WLB’s 0.78 (95% CI, 0.68–0.86) with P=0.056. Egger’s test P value (0.225) demonstrated that there was no publication bias.

Conclusions: Our research showed that in the evaluation of bronchial cancers, AFI was superior to conventional WLB. With its higher sensitivity, AFI could be valuable for avoiding misdiagnosis.
Introduction

Lung cancer is a major cause of cancer-related death worldwide (1,2). It was reported that the 5-year survival rate for those patients with a stage 0 cancer is more than 90% (3), whereas the rate for patients with stage IA disease is 73%, and the rate for stage II to IV ranges from 9% to 46% (4). Therefore, it is essential to clinically detect early lung cancers by using more sensitive methods, as the discovery and treatment of early stage lung cancer not only enhances the survival rate, but also the patients’ quality of life (5-7). This goal can only be achieved with the development of more sensitive methods that have an acceptable specificity for these early stages. Interestingly, the development of these newer technologies has also provided useful information to better understand tumor transformation and the carcinogenic mechanisms of bronchial (pre) cancerous lesions (8).

A number of studies have compared the sensitivity and specificity of autofluorescence bronchoscopy (AFB) and white-light bronchoscopy (WLB) (3,9,10). The results and conclusions were discordant, a situation which, unfortunately, contributed to the limited acceptance of AFB (10-16). In order to solve this problem, Chen et al. used a meta-analysis which demonstrated that AFB had a higher sensitivity and lower specificity, and that the combined utilization of AFB and WLB was superior to WLB alone and had a higher sensitivity; there was, however, no analysis comparing the specificity of these methods performed in this study (17).

Autofluorescence imaging (AFI) is a term adopted by the Olympus Medical Systems Corporation to describe AFB based on its technology. Briefly, the AFI system developed by this company (Evis Lucera Spectrum) consists of three main parts (for more details, please visit: https://www.olympus-global.com/en/news/2006a/nr060515evise.html): a xenon light source, an autofluorescence video bronchoscope (BF-F260), and a video processor unit (CV-260SL). Images produced by AFI technology can be displayed in both the traditional (white light) and autofluorescence modes on the same monitor via a switch. The system transmits 3 wavelengths: excitation blue light (395–445 nm, to induce autofluorescence), 550 nm (red reflected light), and 610 nm (blue reflected light). Normal mucosa appears green, inflammation appears blue (because of a high concentration of hemoglobin which can absorb the green and red wavelengths), and cancers and precancerous lesions appear magenta (because they can mix red/blue signals and shorten the green autofluorescence) (18,19).

In the present study, we explored the reported performance of AFI compared to WLB in the diagnosis of bronchial cancerous lesions.

Methods

Search strategy

We conducted a systematic search of the PubMed, Embase, Web of Science and CNKI databases, from inception to July 12th, 2018; we restricted our search to English language publications to avoid sources of local/national articles which are frequently of low quality. The following keywords were used as search terms: (“optical imaging”[MeSH Terms] OR (“optical”[All Fields] AND “imaging”[All Fields]) OR “optical imaging”[All Fields] OR (“autofluorescence”[All Fields] AND “imaging”[All Fields]) OR “autofluorescence imaging”[All Fields] AND videobronchoscopy[All Fields] and white-light[All Fields] AND (“bronchoscopy”[MeSH Terms] OR “bronchoscopy”[All Fields]).

Inclusion and exclusion criteria

Our inclusion criteria for studies were as follows:

(I) Involved patients who were suspected of having bronchial cancer;

(II) Compared the use of an AFI system with WLB bronchoscopy;

(III) Used histological analysis of biopsies as the golden standard for diagnosing bronchial cancer, with the status of positive results for “moderate dysplasia or worse” or “mild dysplasia or worse” or “tumor” in different studies. The detailed characteristics of the
Studies were excluded if they were duplicate studies or in vitro studies, involved animal experiments, lacked a control group to compare the capabilities of WLB with AFI, or if the identified study was a meeting abstract.

Quality assessment and data extraction

Data were extracted independently by two investigators, and differences were resolved by consensus. Quality assessment of these studies was performed using Cochrane Collaboration’s risk-of-bias tool which considers the following criteria: reporting of randomization method, allocation concealment, blinding of outcome assessment, completeness of follow-up, and bias of selective reporting (29).

Statistical analysis

The random model for the diagnostic meta-analysis was used to obtain pooled sensitivities and specificities (30), with pooled sensitivity and specificity of AFI and WLB being estimated as diagnostic capability, which were displayed in a forest plot. The positive predictive value, negative predictive value, and the area under the curve (AUC) were analyzed simultaneously. P<0.05 was used to identify significant differences.

The two investigators constructed 2×2 tables for each study. The contents of the four table cells were as follows: true positive (TP), false positive (FP), false negative (FN), and true negative (TN). We used STATA 14.0 (StataCorp, College Station, TX, USA), in particular the MIDAS Command Language, for all statistical analysis. The sensitivity was identified as the percentage of the disease which was diagnosed correctly according to the criteria of the screening method. The specificities were identified as the percentage of the actual disease which was not diagnosed according to the method.

Results

Literature search results and population characteristics

Using the methods described above, we identified 189 publications which were selected by our filtration criteria. Of these, 33 duplicates and 76 other articles were excluded.

<table>
<thead>
<tr>
<th>Studies</th>
<th>Positive results</th>
<th>Biopsy specimens</th>
<th>Biopsy specimens</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AFI</td>
<td>WLB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TP</td>
<td>FP</td>
<td>FN</td>
</tr>
<tr>
<td>Chiyo (18)</td>
<td>Dysplasia or worse</td>
<td>62</td>
<td>26</td>
</tr>
<tr>
<td>Ueno (20)</td>
<td>Severe dysplasia or worse</td>
<td>64</td>
<td>18</td>
</tr>
<tr>
<td>Li (21)</td>
<td>Severe dysplasia or worse</td>
<td>241</td>
<td>72</td>
</tr>
<tr>
<td>Zaric (22)</td>
<td>Carcinoma</td>
<td>108</td>
<td>36</td>
</tr>
<tr>
<td>Herth (23)</td>
<td>Moderate or severe dysplasia or CIS</td>
<td>57</td>
<td>11</td>
</tr>
<tr>
<td>Cetti (24)</td>
<td>Moderate dysplasia or worse</td>
<td>81</td>
<td>14</td>
</tr>
<tr>
<td>Zaric (25)</td>
<td>Carcinoma</td>
<td>624</td>
<td>286</td>
</tr>
<tr>
<td>Ikeda (26)</td>
<td>CIS or severe dysplasia</td>
<td>177</td>
<td>78</td>
</tr>
<tr>
<td>Zheng (27)</td>
<td>Malignant lesion</td>
<td>218</td>
<td>151</td>
</tr>
<tr>
<td>Zhu (28)</td>
<td>Invasive cancer or severe dysplasia</td>
<td>198</td>
<td>156</td>
</tr>
</tbody>
</table>

AFI, autofluorescence imaging; WLB, white-light bronchoscopy; TP, true positive; FP, false positive; FN, false negative; TN, true negative.
(animal experiments, in vitro studies, multiple subjects, and meeting abstracts), leaving 80 articles. A further 46 studies were excluded after a careful review of the titles and abstracts revealed that they were not comparisons or were not relevant to the present study.

After screening for articles of high-quality that also met our specific inclusion criteria, a total of 10 articles (18,20-28) were eligible for the final meta-analysis. A flow chart of our meta-analysis is presented in Figure 1, while detailed features of the included studies are presented in Table 2.

Quality assessment

As shown in Figure 2, the quality of the studies included in the present investigation was assessed by the Cochrane risk-of-bias tool. Some studies failed to provide a clear method of blinding (including the blinding of participants, personnel, and outcome assessment), while a few studies had limitations in sample size.

Diagnostic accuracy indices

As shown in Figure 3, the sensitivity of AFI ranged from 0.65 to 0.98, with an I² of 72.29 (range, 54.59–90.00), while the specificity varied between 0.21 and 0.92, with an I² of 95.10 (range, 93.17–97.02). As shown in Figure 4, the sensitivity of WLB ranged from 0.18 to 0.94, with an I² of 87.37 (range, 80.82–93.92), and the specificity ranged from 0.50 to 0.92, with an I² of 88.56 (range, 82.79–94.33). The pooled sensitivity and specificity of AFI were 0.92 (95% confidence interval, 0.88–0.95) and 0.67 (95% confidence interval, 0.51–0.80), respectively (Figure 3). The pooled sensitivity and specificity of WLB were 0.70 (95% confidence interval, 0.58–0.80) and 0.78 (95% confidence interval, 0.68–0.86), respectively (Figure 4). The positive predictive value of AFI vs. WLB was 85.0% vs. 76.7% respectively (P<0.05), and the negative predictive value of AFI vs. WLB was 67.6% vs. 70.5% (P=0.06), respectively. Our study showed that the AUCs of AFI and WLB were 0.92 (range, 0.89–0.94).
and 0.81 (range, 0.77–0.84), respectively (Figures 5 and 6 for AFI and WLB, respectively). Chi-square test was used to compare the difference between the two rates, and the specificity of AFI vs. WLB showed no difference (P=0.056). There was no difference of the negative predictive value between AFI and WLB (P=0.06) (data available in Table 3).

### Publication bias and stability of the results

The Egger's test used to assess publication bias resulted in a P value of 0.225 (Table 4), which suggested there was no or little publication bias. Sensitivity analysis (Figure 7) was used to assess the stability of the results; the 95% confidence intervals of each trial overlapped with one another, proving our eligible stability.

### Discussion

After passing this review’s strict selection criteria, 10 articles were eligible for the final meta-analysis. A total of 1,830 patients were included in the analysis. We found a better relative sensitivity and a comparable specificity of AFI versus WLB by comparing various indicators of diagnostic effectiveness. Subsequent objective evaluation of

---

**Table 2 Detailed features of the 10 studies included**

<table>
<thead>
<tr>
<th>Studies</th>
<th>Randomized</th>
<th>Controlled</th>
<th>Multi-center</th>
<th>Prospective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chiyo 2005 (18)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ueno 2007 (20)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Li 2010 (21)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Zaric 2009 (22)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Herth 2009(23)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Cetti 2010 (24)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Zaric 2010 (25)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ikeda 2006 (26)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Zheng 2017 (27)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Zhu 2012 (28)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Figure 2** Quality assessment of the studies. Quality assessment of these studies was performed using the Cochrane Collaboration's risk-of-bias tool by RevMan 5.3.3 software.
Figure 3  Forest plot showing study-specific (right-axis) and mean sensitivity along with specificity of AFI, with corresponding heterogeneity statistics. The sensitivity of AFI ranged from 0.65 to 0.98, with an $I^2$ of 72.29 (range, 54.59–90), while the specificity varied ranged 0.21 to 0.92, with an $I^2$ of 95.10 (range, 93.17–97.02), as determined by the STATA 14 software. The pooled sensitivity and specificity of AFI were 0.92 (95% confidence interval, 0.88–0.95) and 0.67 (95% confidence interval, 0.51–0.80). AFI, autofluorescence imaging video bronchoscopy.

the publication bias and the stability of our results indicate that our results are likely to be credible.

Zheng et al. (27) have previously described the factors explaining AFI’s poorer specificity when compared to WLB: the friction damage of the airway wall caused by bronchoscopy in the process, airway mucosal inflammation, oral anticoagulation, ingestion of photosensitizing drugs within 3 months, cytotoxic chemotherapy conducted within 6 months, and many nonneoplastic diseases leading to false-positive results in AFI. This study is an impetus to research more effective image analysis methods and to optimize the spectral design of AFI. Indeed, it should be noted that the results reported above are specific for the Evis Lucera Spectrum from Olympus, whereas other manufacturers are commercial systems based on significantly different spectral designs to perform AFB. Sutedja (31) showed that by using AFI and combining methods of forceps biopsy, brush biopsy, needle aspiration, and douche to acquire samples, the comprehensive positive diagnosis rate of lung cancer was clearly improved, demonstrating that AFI has more significant clinical value than WLB for the diagnosis of bronchial cancer. Other studies and meta-analyses have concluded that AFB has a higher sensitivity and lower specificity than WLB (20-22,28), and some studies proved that the usage of AFI does not increase the adverse effects of bronchoscopy (23,25).

Some limitations existed in our research. First, there was an absence of relevant, large randomized controlled trials. Second, we did not investigate the different pathological types of lung cancers, such as mild-to-moderate, moderate-to-severe, and severe types. Instead, we roughly divided the pathological types into a normal group and a malignant cancer group. Third, the level of training (learning curve) of the bronchoscopists was not taken into account. Fourth,
Figure 4 Forest plot showing study-specific (right-axis) and mean sensitivity along with the specificity of WLB, with corresponding heterogeneity statistics. The sensitivity of WLB ranged from 0.18 to 0.94, with an I² of 87.37 (range, 80.82–93.92), and the specificity ranged from 0.50 to 0.92, with an I² of 88.56 (range, 82.79–94.33). The pooled sensitivity and specificity of WLB were 0.70 (95% confidence interval, 0.58–0.80) and 0.78 (95% confidence interval, 0.68–0.86). WLB, white-light bronchoscopy.

Figure 5 Summary ROC curve with confidence and prediction contour around mean operating sensitivity and specificity point of AFI. The AUC of AFI was 0.92 (range, 0.89–0.94). AFI, autofluorescence imaging video bronchoscopy; ROC, receiver operating characteristic; AUC, the area under the ROC curve.

Figure 6 Summary ROC curve with confidence and prediction contour around the mean operating sensitivity and specificity point of WLB. The AUC of WLB was 0.81 (range, 0.77–0.84). WLB, white-light bronchoscopy; ROC, receiver operating characteristic; AUC, the area under the ROC curve.
since virtually all biopsies were taken under white light observation, the spatial precision of the tissue uptake was probably much better for WLB than AFI, a bias which precludes the assessment of the performances of the latter method. Furthermore, it was not possible to completely rule out sources of publication bias such like incomplete data and inconsistent positive results.

Table 3 The diagnostic evaluation of WLB and AFI

<table>
<thead>
<tr>
<th>Projects</th>
<th>AFI (95% CI)</th>
<th>WLB (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.92 (0.88–0.95)</td>
<td>0.70 (0.58–0.80)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.67 (0.51–0.80)</td>
<td>0.78 (0.68–0.86)</td>
<td>0.056</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>0.85 (0.74–0.92)</td>
<td>0.767 (0.7–0.82)</td>
<td>0.03</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>0.676 (0.51–0.74)</td>
<td>0.705 (0.67–0.7)</td>
<td>0.06</td>
</tr>
</tbody>
</table>

AFI, autofluorescence imaging; WLB, white-light bronchoscopy.

Table 4 The Egger’s test used to assess publication bias

| Egger’s test | Coef.       | Std. Err. | t   | P>|t| | 95% CI       |
|--------------|-------------|-----------|-----|-----|----------------|
| Slope        | 4.850776    | 1.181086  | 4.11| 0.003| 2.127188 to 7.574365 |
| Bias         | −2.96829    | 2.257466  | −1.31| 0.225| −8.174017 to 2.237436 |

Figure 7 Sensitivity analysis was used to assess the stability of the results; the 95% confidence intervals of each trial overlap with each other, proving our eligible stability.

More refined and extended versions of such analysis would also provide interesting information for the medical, industrial, and scientific communities active in the field of AFB. In particular, assessing the performance of AFI used in combination with WLB and/or narrow band imaging (NBI) would be interesting. Also, a comparison of the performances achieved with different commercially available systems for AFB would enable us to identify which generations and, importantly, which spectral designs (excitation and detection wavelengths, combined detection of autofluorescence, and backscattered light at specific wavelengths) are optimal for the detection and/or demarcation of bronchial cancers. Indeed, these spectral designs vary significantly across these systems, with some of them being particularly optimized to...
minimizing false positives, probably without affecting their sensitivities (32-34).

Finally, one interesting general consensus revealed from the analysis of the articles considered in our meta-analysis is that AFI should be used for detection purposes in patients in whom pre-invasive lesions (dysplastic, carcinoma in situ) have been detected but who have showed no evidence of invasive cancer. In addition, AFI should be used for demarcation purposes in patients with early invasive lung cancers for whom endobronchial therapy is indicated.

Acknowledgments

**Funding:** This work was supported by grants from the National Natural Science Foundation of China (grant number 81401903, 81572937 and 81572273); the 16th batch “Summit of the Six Top Talents” Program of Jiangsu Province (grant number WSN-154); China Postdoctoral Science Foundation 12th batch Special fund (Postdoctoral number: 45786); China Postdoctoral Science Foundation 64th batch (Postdoctoral number: 45786); Jiangsu Provincial Postdoctoral Science Foundation (grant number 2018K049A); the Natural Science Foundation of Jiangsu province (grant number BK20180139 and BK20161386 and BK20191351); Jiangsu Provincial Medical Youth Talent (grant number QNRC2016125), and the Nanjing Medical Science and Technology Development Project (No. ZKX17044), and Jiangsu Planned Projects for Postdoctoral Research Funds (2019k178).

Footnote

**Conflicts of Interest:** The 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.

References


© Translational lung cancer research. All rights reserved.