Color Fluorescence Ratio for Detection of Bronchial Dysplasia and Carcinoma In situ

Pyng Lee,1 Remco M. van den Berg,1 Stephen Lam,2 Adi F. Gazdar,3 Katrien Grunberg,1 Annette McWilliams,2 Jean LeRiche,2 Pieter E. Postmus,1 and Tom G. Sutedja1

Abstract

Background: Autofluorescence bronchoscopy is more sensitive than conventional bronchoscopy for detecting early airway mucosal lesions. Decreased specificity can lead to excessive biopsy and increased procedural time. Onco-LIFE, a device that combines fluorescence and reflectance imaging, allows numeric representation by expressing red-to-green ratio (R/Gratio) within the region of interest. The aim of the study was to determine if color fluorescence ratio (R/Gratio) added to autofluorescence bronchoscopy could provide an objective means to guide biopsy.

Methods: Subjects at risk for lung cancer were recruited at two centers: VU University Medical Centre (Amsterdam) and BC Cancer Agency (Canada). R/Gratio for each site appearing normal or abnormal was measured before biopsy. R/Gratio were correlated with pathology, and a receiver operating characteristic curve of R/Gratio for high-grade and moderate dysplasia was done. Following analysis of the training data set obtained from two centers, a prospective validation study was done.

Results: Three thousand three hundred sixty-two adequate biopsies from 738 subjects with their corresponding R/Gratio were analyzed. R/Gratio 0.54 conferred 85% sensitivity and 80% specificity for the detection of high-grade and moderate dysplasia, area under the curve was 0.90, and 95% confidence interval was 0.88 to 0.92. In another 70 different sites that were assessed, κ measurements of agreement of R/Gratio with visual scores and pathology were 0.66 (P < 0.0001) and 0.61 (P < 0.0001), respectively. R/Gratio combined with visual score improved specificity to 88% (95% confidence interval, 0.73-0.96) for high-grade and moderate dysplasia.

Conclusion: Color fluorescence ratio can objectively guide the bronchoscopist in selecting sites for biopsy with good pathologic correlation.

Lung cancer causes more deaths than cancers of the prostate, breast, and colon combined, and prognosis is strongly dependent on the stage of disease at presentation (1). In a recently published computed tomography screening trial, one could expect 10-year survival in excess of 80% for clinically stage I parenchymal tumors if intervened early (2). However, computed tomography yield for early central airway cancers remains poor (3, 4).

Based on necropsy study of smokers, Auerbach et al. postulated that squamous cell carcinomas arose from preinvasive lesions that affected the central airways (5, 6). Because preinvasive lesions are rarely detectable by white-light bronchoscopy (WLB; ref. 7), the hypothesis that squamous cell carcinogenesis is a multistep process, which develops from squamous metaplasia through mild, moderate, and severe dysplasia, carcinoma in situ, and finally invasive carcinoma, is supported by sputum cytology studies where 10% of moderate dysplasia and 40% to 83% of severe dysplasia are reported to progress to invasive lung cancer (8–10).

Autofluorescence bronchoscopy (AFB), which exploits differences in fluorescence properties of normal and abnormal bronchial mucosa, facilitates the detection of preinvasive neoplasia, which is otherwise invisible on WLB (11–15). Although AFB is highly sensitive, difficulties in distinguishing airway inflammation from preinvasive lesions as well as in the accurate prediction of pathologic diagnosis based on visual grading of tissue fluorescence can lead to excessive biopsy, longer procedural time, higher incidence of procedure-related complications, and greater health costs (16). In fact, studies have shown that as high as one third of lesions with abnormal fluorescence represented false positives when correlated with pathology (12, 13).

Several devices have been developed for commercial use, which include the LIFE-Lung (Xilix Technologies), Storz D-light (Karl Storz), SAFE 1000 and 3000 (Pentax), AFI (Olympus), and Onco-LIFE (Onco-LIFE Endoscopic Light Source and Video...
Translational Relevance

Autofluorescence bronchoscopy, which exploits differences in fluorescence properties of normal and abnormal bronchial mucosa, facilitates the detection of preinvasive neoplasia otherwise invisible on conventional bronchoscopy. Autofluorescence bronchoscopy is highly sensitive for airway dysplasia; however, decreased specificity leads to excessive biopsy and increased procedural time that may compromise patient safety. By representing the red-to-green fluorescence ratio of the target site and correlating to pathology, a receiver operating characteristic curve is generated. The derived color fluorescence ratio that is predictive of moderate dysplasia or worse at risk of progression to carcinoma in situ is prospectively validated and can serve as an objective means to guide selection of sites for biopsy.

Materials and Methods

The study was conducted at two centers: VU University Medical Center (Amsterdam) and BC Cancer Agency (Canada). The protocol was approved by respective institutional review boards, and written informed consent was obtained from all participants.

Study population. Indications for participants undergoing AFB include one of the following: (a) current or former smokers with known or suspected lung cancer scheduled for bronchoscopy, (b) abnormal sputum cytology and normal radiograph, and (c) surveillance after curative surgery for stage I lung cancer. Patients who had received photosensitizing agents or chemopreventive drugs such as retinoids within 3 months and radiotherapy to chest or cytotoxic chemotherapy within 6 months of bronchoscopic procedure were excluded. Other exclusion criteria were pneumonia, acute bronchitis, poorly controlled hypertension, unstable angina, bleeding disorders, pregnancy, and adverse reactions to topical lidocaine.

Equipment. Light hitting the bronchial surface can be absorbed, reflected, back-scattered, or induce fluorescence. These optical properties are used to determine structural features, biochemical composition, as well as functional changes in normal and abnormal bronchial tissues. WLB is an example of reflectance imaging where structural features of the epithelial surfaces are defined to allow discrimination of normal from abnormal bronchial tissues. Fluorescence imaging provides information about the biochemical composition and metabolic state of bronchial tissues, which in turn depend on the concentration of fluorophores in collagen, elastin, and those involved in cellular metabolism such as nicotinamide adenine dinucleotide, flavins, aromatic amino acids, porphyrins, and lipopigments as well as the distinct excitation and emission spectrum of each fluorophore. When normal bronchial epithelium is illuminated by blue light, it fluoresces in green, and as it transforms through different grades of dysplasia, carcinoma in situ to invasive cancer, a progressive decrease in green fluorescence due to increased epithelial thickness and vascularization occurs, making these abnormal areas appear red.

The newly developed Onco-LIFE device uses both reflectance and fluorescence light for imaging. Blue light (395-445 nm) and small amount of red light (675-720 nm) from a filtered mercury arc lamp are used for illumination. A red reflectance image is captured together with the green autofluorescence image by non-image-intensified charge-coupled devices to enhance the contrast among premalignant, malignant, and normal tissues. By using reflected infrared red light as a reference, it has the theoretical advantage over reflected blue or green light, as it is less absorbed by hemoglobin and therefore less influenced by changes in vascularity associated with inflammation (17). Onco-LIFE also allows quantitative analysis of the fluorescence image by providing a numeric representation (R/G ratio) of the combined colors in the central portion of the displayed image (Supplementary Fig. S1). R/G ratio of the 16 x 16 pixel square target defined within the displayed brackets is calculated by dividing the average red reflectance with green fluorescence signals captured by the camera (18, 19). The ability to quantify color changes can serve as an objective means for comparison against normal tissue as well as an aide to reduce interobserver variation.

Procedure. The procedure was carried out in two stages using the Onco-LIFE device under local anesthesia with or without sedation. The first stage was done with fiber-optic WLB. During this stage, sites of interest were visually classified according to Table 1 and no change in classification was allowed once WLB had been completed. Under the visual classification system, areas without any visual abnormality were classified as class I (normal); areas with the appearance of inflammation, trauma, anatomical abnormalities, metaplasia, or mild dysplasia were classified as class II (abnormal); and areas suggestive of moderate dysplasia, severe dysplasia, carcinoma in situ, or invasive tumor were class III (suspicious).

After WLB, the device was switched to AFB mode and another examination of the airways was done. Areas of interest were visually classified using the same classification scheme as detailed in Table 1.
Sites classified as III with AFB but not classified using WLB were automatically defaulted to class I under WLB. Still, images of all sites of interest were taken under WLB and AFB. During capture of the image under AFB mode, the endoscopist also centered a small target over the most abnormal part of the lesion, and red-to-green light intensity (R/G) ratio was recorded automatically by the computer without display on the screen. The R/G ratios were subsequently used for comparison against the pathology results of the biopsies after the trial had been completed (Supplementary Fig. S2).

After AFB examination, all areas classified as III by either imaging modality were sampled as well as at least one class I/II site as control. Biopsy was acquired under AFB unless the lesion was visible only under WLB. The clinical center’s pathologist, who was blinded to the endoscopic findings, first evaluated the biopsy slide followed by a reference pathologist who was blinded to the previous results. All biopsy specimens had to have sufficient bronchial epithelium for evaluation, and specimens were graded in accordance with the International Histological Classification of Tumors published by the WHO (20). Pathologists coded the biopsy slides based on a nine-point scheme: (a) normal, (b) inflammation/bronchitis, (c) hyperplasia, (d) squamous metaplasia, (e) mild dysplasia, (f) moderate dysplasia, (g) severe dysplasia, (h) carcinoma in situ, or (i) invasive carcinoma. High-grade preinvasive lesions (HGD) were defined as severe dysplasia and carcinoma in situ, whereas low-grade lesions (LGD) include hyperplasia, squamous metaplasia, and mild dysplasia. In case of disagreement between center and reference pathologist, slides were re-read by the reference pathologist, and these results were taken as final.

Following analysis of the training data sets obtained from two centers, a prospective study using similar protocol was done at VU University Medical Centre for validation. All procedures in the validation study were conducted by T.G.S. who had 15 years of experience in AFB where each site was first visually classified under WLB and AFB before R/G ratio determination and biopsy.

Statistics. For the purpose of statistical analysis, visual classification was converted to a two-point visual score where classes 1 and 2 became “negative” and class 3 became “positive.” If the lesion was graded class 3 by AFB and class 2 on WLB, classification by AFB took precedence. Final pathologic diagnosis was also converted to a two-point scale where codes 1 to 5 were labeled as “negative” and codes 6 to 8 as “positive.” A receiver operating characteristic (ROC) was constructed, and a cutoff R/G ratio was selected from coordinates of ROC that conferred good sensitivity and specificity for the detection of HGD and moderate dysplasia. A hierarchical model with multilevel mixed-effects logistic regression was used to assess the effects of the following variables on HGD and moderate dysplasia: age group (≤59, ≥60 years), gender, smoking status (current or former smoker), indication for recruitment, and R/G ratio (<0.54, ≥0.54). Heterogeneity effects due to center and within- and between-patient variation arising from multiple biopsy samples per patient were factored into the model. Samples coming from the same patient were nested within that patient who in turn was nested within the center. STATA version 10.1 was used to carry out the analysis.

We prospectively validated the cutoff value by performing κ measurements of agreement between the R/G ratios of respective sites with visual scores and final pathologic diagnoses. Values were expressed as median (interquartile range) and mean (SD). Comparisons between groups were done with χ² and Mann Whitney U tests. P < 0.05 was considered statistically significant.

**Results**

Seven hundred thirty-eight individuals underwent WLB and AFB at two centers. There were 492 males (67%), and the

---

**Table 2. Distribution of tissue histology and R/G ratios for two participating centers**

<table>
<thead>
<tr>
<th>Centers</th>
<th>Biopsy</th>
<th>Mean (SD) R/G ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC Cancer Agency</td>
<td>2,601</td>
<td></td>
</tr>
<tr>
<td>Histology of tissue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carcinoma in situ</td>
<td>3</td>
<td>1.40 (0.38)</td>
</tr>
<tr>
<td>Severe dysplasia</td>
<td>62</td>
<td>1.13 (0.43)</td>
</tr>
<tr>
<td>Moderate dysplasia</td>
<td>44</td>
<td>0.80 (0.27)</td>
</tr>
<tr>
<td>Mild dysplasia or less</td>
<td>2,492</td>
<td>0.42 (0.16)</td>
</tr>
<tr>
<td>VU University Medical Centre</td>
<td>691</td>
<td></td>
</tr>
<tr>
<td>Histology of tissue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carcinoma in situ</td>
<td>89</td>
<td>1.45 (0.54)</td>
</tr>
<tr>
<td>Severe dysplasia</td>
<td>37</td>
<td>1.04 (0.59)</td>
</tr>
<tr>
<td>Moderate dysplasia</td>
<td>43</td>
<td>0.70 (0.31)</td>
</tr>
<tr>
<td>Mild dysplasia or less</td>
<td>522</td>
<td>0.42 (0.25)</td>
</tr>
</tbody>
</table>

---

Fig. 1. Curve of ROC for R/G ratio in the detection of HGD and moderate dysplasia based on training data.

Fig. 2. Curve of ROC for R/G ratio in the detection of HGD (severe dysplasia and carcinoma in situ) based on training data.
median age was 61 years (range, 54-68). All were current (33%) and former (67%) smokers of median 43 pack-years (range, 36-53). Subjects at BC Cancer Agency were recruited for known or suspected lung cancer based on symptoms and/or abnormal sputum cytology while at VU University Medical Centre; the majority (64%) was included in the surveillance program following curative surgery for lung cancer and the rest (36%) was referred for suspected lung cancer and/or abnormal sputum cytology.

A total of 3,292 biopsies was assessed adequate for evaluation with the corresponding sites analyzed for R/G ratios. Table 2 and Supplementary Fig. S3 show the distribution of biopsy specimens obtained at both centers categorized according to histology and their respective R/G ratios.

Two hundred seventy-eight biopsies (8%) were graded as HGD and moderate dysplasia, whereas the rest represented LGD. Median R/G ratios for HGD and moderate dysplasia and for LGD were 1.06 (range, 0.69-1.40) and 0.39 (range, 0.29-0.52), respectively (P < 0.001). R/G ratio 0.54 conferred sensitivity of 85% and specificity of 80% for the detection of HGD and moderate dysplasia. Area under the ROC curve (Fig. 1) was 0.90 with 95% confidence interval (95% CI) 0.88 to 0.92. When only HGD was analyzed, R/G ratio 0.60 gave sensitivity of 86% and specificity of 85% and area under ROC curve (Fig. 2) was 0.94 with 95% CI 0.92 to 0.96. Positive and negative predictive values using 0.54 as the threshold for HGD and moderate dysplasia were 27% (95% CI, 0.24-0.30) and 98% (95% CI, 0.98-0.99), respectively.

The mixed-effects logistic regression analysis showed that the female gender who underwent AFB surveillance after curative surgery for lung cancer was 1.6 times more at risk of moderate to HGD (odds ratio, 1.6; 95% CI, 1.2-2.2). Former smokers as opposed to current smokers were 1.4 times more at risk (odds ratio, 1.4; 95% CI, 1.1-2.2) of moderate to HGD. High R/G ratios of ≥0.54 were 19.2 times more likely to have moderate dysplasia and HGD (odds ratio, 19.2; 95% CI, 13.3-27.9). Age and indication for recruitment did not show statistically significant association with moderate to HGD.

We prospectively validated our findings in another 30 patients (25 males) comparable in age and tobacco exposure where 70 sites were first visually scored before determination of R/G ratio and biopsy. Thirty sites diagnosed as HGD and moderate dysplasia (3 carcinomas in situ, 13 severe dysplasias, and 14 moderate dysplasias) had median R/G ratio 1.26 (range, 0.90-1.57) compared with the remaining 40 showing LGD that had R/G ratio 0.40 (range, 0.23-0.56; P < 0.001). Validation data showed that R/G ratio 0.54 had 93% sensitivity and 70% specificity for the detection of HGD and moderate dysplasia. Area under the curve was 0.96 with 95% CI 0.92 to 1.00 (Fig. 3). When R/G ratio 0.54 was used as the cutoff value, \( \kappa \) measurements of agreement between R/G ratios and visual scores as well as pathology were 0.66 (\( P < 0.0001 \)) and 0.61 (\( P < 0.0001 \)), respectively (Table 3).

In our validation study, two sites showing moderate dysplasia had R/G ratios <0.54 (Table 3); however, we did not find HGD with R/G ratio <0.54. When R/G ratio was combined with visual score, it resulted in 80% sensitivity (95% CI, 0.61-0.92) and an improved specificity of 88% (95% CI, 0.73-0.96; Table 4).

**Discussion**

Since its inception, AFB has made inroads in facilitating the detection of early-stage central airway cancers that are missed on computed tomography in a population who has significant tobacco exposure and likely to suffer from smoking related comorbidities such as chronic obstructive pulmonary disease and coronary heart disease. In fact, as high as 20% of these patients may harbor synchronous cancers or develop multifocal disease due to field cancerization (7). Although surgery achieves the best results for cure, it carries a significant

Table 3. Validation of derived R/G ratios with visual score and pathology

<table>
<thead>
<tr>
<th>R/G ratio</th>
<th>( \kappa(P) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥0.54</td>
<td>&lt;0.54</td>
</tr>
<tr>
<td>Visual score by expert (n = 70 sites)</td>
<td></td>
</tr>
<tr>
<td>Positive (score 1; class 3/suspicious)</td>
<td>29</td>
</tr>
<tr>
<td>Negative (score 2; class 1/normal and 2/abnormal)</td>
<td>11</td>
</tr>
<tr>
<td>Pathology (n = 70 specimens)</td>
<td></td>
</tr>
<tr>
<td>Positive (HGD and moderate dysplasia)</td>
<td>28</td>
</tr>
<tr>
<td>Negative (LGD)</td>
<td>12 (1 normal, 2 inflammation, 5 squamous metaplasia, and 4 mild dysplasia)</td>
</tr>
</tbody>
</table>

*\( P < 0.05. \)
perioperative risk (21) and may not be feasible for patients with poor cardiorespiratory reserve or in those with multicentric cancers. Moreover, it has been shown that patients with preinvasive lesions are predisposed to developing invasive lung cancers elsewhere (22, 23), and prior lung surgery for preinvasive neoplasm might render them unfit for subsequent curative surgery (24). By detecting these lesions early when they are still confined within the airway wall, endobronchial treatments such as photodynamic therapy, electrocautery, and high-dose brachytherapy can be considered as effective alternatives while preserving lung tissue and quality of life at the same time (25–28).

Many investigators have independently or collaboratively shown that, when AFB is combined with WLB, AFB increases the detection of preinvasive lesions by 1.5- to 6.3-fold (11–15, 29–32). However, its lower specificity is a cause for concern as this may lead to extensive biopsy, prolonged procedural time, increased procedure-related complications such as bronchitis and bleeding, as well as greater health costs. Of particular clinical relevance would be in the surveillance of patients who have received endobronchial therapy for severe dysplasia or carcinoma in situ, where previous biopsy sites and consequent airway fibrosis could cause abnormal fluorescence that might persist for months to years, thereby making interpretation a challenge even to experienced bronchoscopists.

Undoubtedly scoring each site as normal, abnormal, or suspicious as well as deciding if a lesion with borderline fluorescence requires biopsy or not comes with experience and a steep learning curve (16). Although various investigators have shown higher R/G ratios with increasing grade of bronchial dysplasia (18, 19), our study is the first to derive from ROC analysis a threshold R/G ratio that is sensitive and specific for the detection of moderate and HGD. This was further investigated where lesions showing moderate and HGD were more likely to have R/G ratios >0.54. Similarly, sites with R/G ratios <0.54 were more likely to show LGD.

Through quantification of red-to-green image intensities, one can reduce interindividual variation that may occur during visual scoring and be prompted to perform biopsy of the site under evaluation when it registers R/G ratio ≥0.54. Notwithstanding that an important prerequisite to assure accurate calculation of R/G ratio depends on centering the target on the most abnormal part of the lesion, incorporating R/G ratio to visual scoring of abnormal airway lesions further aids the experienced bronchoscopist in deciding which sites for sampling with enhanced specificity.

Anecdotally, only performing biopsy of AFB abnormal sites with R/G ratio ≥0.54 has decreased extent of sampling and procedural time at both centers. A cutoff value that shows good sensitivity for moderate and HGD is useful to prompt biopsy, but the converse can be argued where sites showing R/G ratios <0.54 are likely LGD and therefore obviates the need for tissue confirmation. An outcome study incorporating color fluorescence ratio to determine its effect on procedural time, interindividual variability, as well as watchful observation of airway lesions left alone due to low R/G ratio will be important particularly if any progresses to carcinoma.

### Table 4. Combining visual score and derived R/G ratio

<table>
<thead>
<tr>
<th>Pathology</th>
<th>Visual score and derived R/G ratio*</th>
<th>κ (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive (HGD and moderate dysplasia)</td>
<td>≥0.54 (Positive)</td>
<td>0.68</td>
</tr>
<tr>
<td>Negative (LGD)</td>
<td>&lt;0.54 (Negative)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6 (moderate dysplasia)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 (1 inflammation,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 squamous metaplasia,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 mild dysplasia)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>35</td>
<td></td>
</tr>
</tbody>
</table>

*Biopsy sites graded as visual score 1 (class 3/suspicious) and had R/G ratios ≥0.54 were grouped as positive, whereas those graded as visual score 2 (class 1/normal and 2/abnormal) and had R/G ratios <0.54 were grouped as negative. 

† P < 0.05.

### Conclusion

Color fluorescence ratio when incorporated to AFB examination not only confers more precise localization of preinvasive lesions but also can be used to guide biopsy with good histologic correlation. The concept of color fluorescence ratio is not device specific, as it is derived by dividing average red reflectance with green fluorescence signals captured by the camera (17–19). Thus, it can be integrated into any reflectance-fluorescence imaging system other than Onco-LIFE.

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

### Acknowledgments

We thank Stephanie Fook Chong (Biostatistics, Singapore General Hospital) and Dr. Khin Lay Wai (Research Epidemiologist, National Heart Center) for statistical support.

### References


