Advanced endoscopic imaging of esophageal neoplasia; old looks and new visions

Boerwinkel, D.F.

Citation for published version (APA):
ENDOSCOPIC TRIMODAL IMAGING AND BIOMARKERS FOR NEOPLASIA CONJOINED; A FEASIBILITY STUDY IN BARRETT’S ESOPHAGUS

D.F. Boerwinkel; M. Di Pietro; Xinxue Liu; M.K. Shariff; P. Lao-Sirieix; C.E. Walker; M. Visser; M. O’Donovan; P. Kaye; J.J.G.H.M. Bergman; R.C. Fitzgerald

Published in: Diseases of the Esophagus, 2012 Oct
ABSTRACT

INTRODUCTION: In Barrett’s esophagus (BE), the normal squamous lining of the esophagus is replaced by specialised columnar epithelium. Endoscopic surveillance with autofluorescence imaging (AFI) and molecular biomarkers have been studied separately to detect early neoplasia (EN) in BE. The combination of advanced imaging modalities and biomarkers has not been investigated; AFI may help detecting biomarkers as a risk-stratification tool. We retrospectively evaluated a cohort of patients undergoing endoscopy for EN in BE with AFI and correlated 5 biomarkers (HPP1, RUNX3, p16, Cyclin A and p53) in tissue samples with AFI and dysplasia status.

METHODS: 58 samples from a previous prospective study were selected: 15 true-positive (TP: AFI-positive, EN), 21 false-positive (FP: AFI-positive, no EN), 12 true-negative (TN1; AFI-negative, no EN in sample), 10 true-negative (TN2: AFI-negative, no EN in esophagus). Methylation-specific RT-PCR was performed for HPPI, RUNX3, p16, and immunohistochemistry for Cyclin A, p53. p < 0.05 was considered statistical significant. Bonferroni correction was used for multiple comparisons.

RESULTS: P16, Cyclin A, p53 correlated with dysplasia (p<0.01, p=0.003, p<0.001 respectively). Increased p16 methylation was observed between TP vs TN2 (p=0.003) and TN1 vs TN2 (p=0.04) subgroups, suggesting a field defect. Only p53 correlated with AFI-status (p=0.003). After exclusion of EN-samples, significance was lost.

CONCLUSION: Although correlation with dysplasia status was confirmed for p16, Cyclin A and p53, underlining the importance of these biomarkers as an early event in neoplastic progression, none of the investigated biomarkers correlated with AFI status. A larger prospective study is needed to assess the combination of AFI and a larger panel of
INTRODUCTION

In patients with chronic gastroesophageal reflux disease, a condition called Barrett’s esophagus can arise. In Barrett’s esophagus (BE), the normal squamous lining of the esophagus is replaced by specialised columnar epithelium, which resembles the epithelium found in the small intestine. This so-called intestinal metaplasia gives rise to an increased risk of neoplastic progression through subsequent stages of premalignant dysplasia: low-grade intraepithelial neoplasia (LGIN), followed by high-grade intraepithelial neoplasia (HGIN) and intramucosal carcinoma (IMC). To detect these premalignant lesions at an early stage, regular endoscopic surveillance of BE patients is recommended. However, lesions containing HGIN/IMC can be inconspicuous upon standard endoscopy and therefore can be easily missed. Random biopsies according to the Seattle protocol – taken every centimeter at four quadrants of the BE segment – only sample approximately 5% of the BE. Sampling error can thus easily underestimate the presence of neoplasia in BE.

To increase the detection rate of early neoplasia during endoscopic surveillance of Barrett’s esophagus, different imaging techniques have been studied. A promising approach is endoscopic inspection using an endoscopic tri-modal imaging (ETMI) system, which incorporates high-resolution white-light endoscopy (WLE), autofluorescence imaging (AFI) and narrow-band imaging (NBI). AFI is a technique based on the principle that certain endogenous substances, called fluorophores, emit fluorescent light when excited with short wavelengths of light. It has been demonstrated that early neoplasia has a different fluorescence spectrum than non-neoplastic Barrett’s tissue. In recent international multicenter studies with the ETMI system, AFI was shown to increase the sensitivity for detecting early neoplasia with high-resolution endoscopy from 53-68% to 87-90%. However, this increased detection rate suffers from a significant false positive rate, which can be as high as 71-81%. NBI is a technique that utilizes short wavelength light to enhance the superficial imaging of the mucosa. Detailed inspection with NBI has been used to reduce the high false positive rate of AFI. Although false positive rates could be reduced from 81 to 26% and from 71 to 48% in the two aforementioned multicenter studies, NBI also misclassified a significant number of lesions with HGIN/IMC. NBI, therefore, appears to have only limited value in further classifying lesions that are detected with WLE or AFI.

Despite technological advancement, the cost-effectiveness of endoscopic surveillance in BE is still a highly controversial issue due to the relatively low proportion of BE progress to cancer (0.1-0.5% every year), the high costs generated by the histopathological processing of multiple biopsies taken and the subjective nature of the diagnosis of dysplasia. To overcome these problems research has focused on the identification of biological markers that can predict neoplastic progression in patients with BE. This would allow concentration of endoscopic surveillance on patients at higher risk of progression. Several markers have been proven to correlate with malignant progression, such as loss of heterozygosity at 17p (p53 gene) and 9p (CDKN2A or p16 gene), chromosomal abnormalities (aneuploidy and tetraploidy) and cyclin A, p53 and MCM2 protein expression. Gene silencing by promoter hypermethylation is another common event in carcinogenesis and p16, RUNX3 and HPPI have been investigated previously and showed to correlate with dysplasia and rate of malignant progression. Despite these efforts, biomarkers have not been yet introduced into clinical practice. One of the drawbacks
of molecular biomarkers in terms of clinical applicability is related to the presumed molecular heterogeneity of BE. It has been shown that BE can be highly heterogeneous, with adjacent crypts demonstrating hallmarks of separate clonal origin\(^{(16)}\) \(^{(17,18)}\). In the presence of a heterogeneous disease the reliability of biomarkers assessed on random biopsies might fall significantly short, since random quadrantic biopsies are subjected to sampling error during standard endoscopy\(^{(2,19)}\). So far, few attempts have been carried out to combine novel endoscopic imaging technologies and molecular features, with the aim to increase detection of molecular biomarkers by means of advanced imaging modalities. It is known that the severity of dysplasia correlates with the morphologic appearance during endoscopy\(^{(20,21)}\) as well as with the number of molecular changes\(^{(22,23)}\). Furthermore, it is possible the AFI false positivity may relate to early cellular changes based on the accumulation of molecular abnormalities prior to the development of dysplasia. We hypothesized that the ETMI imaging patterns may correlate with the degree of molecular changes and that the non-dysplastic AFI positive areas may in fact already show increased biological markers for neoplasia compared to non dysplastic, AFI negative (true negative) areas.

Aims of this study were: (i) To correlate the number of molecular changes in tissue samples of BE with the AFI appearance of the corresponding endoscopic areas (ii) To externally validate the relationship between previously published biomarkers and the dysplasia status in BE in an independent cohort of patients.

**MATERIALS AND METHODS**

**Patient selection**

Between January 2007 and September 2009, a prospective study was conducted at the Academic Medical Center (AMC), Amsterdam, to compare standard white light endoscopy with ETMI for the detection of early neoplasia in BE. For this study, patients with a circumferential BE segment >2 cm, without severe oesophagitis (LA grade <B) and biopsy proven HGIN/EC were recruited. Results of this study are described in detail elsewhere\(^{(5)}\). From this cohort of patients, a total of 48 lesions were selected with the following AFI/histology features: 15 AFI positive lesions with HGIN/EC (true positive, TP), 21 AFI positive lesions without dysplasia (false positive, FP) and 12 AFI negative areas without dysplasia from patients with prevalent dysplasia elsewhere in the esophagus (true negative 1, TN1). As a control, 10 AFI negative areas without dysplasia were included (true negative 2, TN2). These areas were selected from Barrett’s patients, enrolled in a surveillance program, who never had any dysplasia during a minimum of 5 years surveillance endoscopy.

**Endoscopic procedure**

Endoscopic work-up was performed using the ETMI system, consisting of a high resolution white light endoscope with optical zoom (XQIF-Q260FZ, Olympus Inc., Tokyo, Japan), equipped with AFI and NBI, coupled to a lightsource and video processor (Lucera, Olympus Inc., Tokyo, Japan). Patients were sedated with intravenous midazolam (2.5 to 15 mg) supplemented with fentanyl (0.1-0.2 mg) if necessary. The esophagus was first examined with high-resolution white light endoscopy. The length of the Barrett’s segment was recorded according to the Prague C&M-
classification\textsuperscript{(24)}. The Barrett’s segment was carefully inspected for the presence of visible abnormalities and for all detected abnormalities the location was recorded on a case record form. Subsequently, the Barrett’s segment was inspected with AFI for additional abnormalities, followed by detailed inspection of all visible abnormalities with NBI. For each modality, overview and detailed (zoom) images were obtained and stored in a standardized way. Two targeted biopsies were obtained from all abnormalities detected during ETMI. As a control, endoscopically unsuspicious areas with a normal (green) AFI appearance were selected within the BE segment at least 2 cm from any AFI suspicious area and subsequently imaged and sampled. Finally, 4-quadrant random biopsy samples were obtained from distal to proximal for every 2-cm of the Barrett’s segment, avoiding taking biopsies from areas with suspicious lesions. Biopsy sampling was performed using standard disposable biopsy forceps.

\textbf{Histological evaluation}

Biopsies were fixed in formalin, embedded in paraffin and routinely cut and stained with haematoxylin and eosin (H&E). Histopathological assessment of the biopsies during the prospective ETMI study was performed by local junior and senior pathologists and revised by an expert GI-pathologist, who recorded the presence of intestinal metaplasia and neoplasia according to the WHO-classification: no-dysplasia, indefinite for dysplasia, LGIN, HGIN or invasive cancer\textsuperscript{(25)}. Samples were evaluated on representative sections throughout the biopsy. For the purpose of this study all original biopsies were revised by a second expert GI-pathologist (MV). Only concordant cases were included for this study and dysplastic samples had to show >50% dysplasia in the epithelial compartment, throughout the biopsy. The latter inclusion criterion was used to avoid dilutional effect on molecular abnormalities in case of very focal dysplasia, since microdissection was not used to isolate dysplastic crypts. Subsequently, all cyclin A and p53 slides were evaluated independently by three reviewers.

\textbf{Immunohistochemistry}

Five 7-micron slides, collected at fixed intervals throughout the biopsy to ensure adequate representation of the mucosal compartment, were cut and mounted on glass slides. Staining was performed with BOND-Max autostainer (Leica). For both p53 (dilution 1:50, clone DO7, Dakocytomation) and cyclin A (dilution 1:40, Dakocytomation), the H\textsubscript{1} antigen retrieval program was used according to the manufacturers’ instructions. A negative control was done by omission of the primary antibody.

The scoring of immunostaining was done by three independent investigators who had no prior knowledge of the clinical diagnosis. In paraffin embedded biopsies, the epithelial surface was defined as the most superficial layer of columnar cells for the glandular tissues. All the surface cells per biopsy were counted up to a maximum of 600 to determine the frequency of cyclin A expression. When biopsies presented areas of intestinal metaplasia and dysplasia, both were scored to generate a single value. Only cells with diffuse nuclear staining were considered cyclin-A positive. The immunopositive epithelial cells were expressed as a percentage of the total number of epithelial cells counted. The cut-off for Cyclin A positivity was 1\%\textsuperscript{(12,26)}.
PS3 expression was scored as either positive, negative or not representative in areas of histological abnormality (i.e. dysplasia). PS3 positivity was defined as high intensity p53 immunoreactivity in histological abnormal areas, relative to background staining and normal histological areas (i.e. non-dysplastic), which served as a internal control\(^{27}\).

**Methyilight**

Ten 15-micron sections from FFPE endoscopic biopsies were collected at fixed intervals throughout the biopsy to ensure adequate representation of the mucosal compartment. An H&E slide was obtained right adjacent to every section to evaluate the epithelial compartment. Sections were deparaffinised in xylene and the excess of xylene was removed with 100% molecular grade ethanol. FFPE sections were lysed with proteinase K in ALT buffer (Qiagen) for 48h at 56°C, with proteinase K freshly added twice daily. DNA was extracted with DNA Blood & Tissue Kit (Qiagen) according to protocol. Bisulfite treatment was performed with EpiTect bisulfite kit (Qiagen) according to manufacturer’s instruction. Bisulfite modified DNA was then subjected to qPCR amplification with methyilight technique for HPP1, p16, RUNX3 genes.

Methylation of genes was quantified by Methyilight\(^{28}\). β-actin was used as internal control. Primers and probes sequences have been previously published\(^{15,29,30}\). Methyilight amplification was performed with the Lightcycler 480 (Roche Diagnostics ltd., UK). Amplification conditions were as follows: 95 °C for 10 min followed by 50 cycles of denaturation at 95 °C C for 15 sec and annealing-extension at 60 °C for 1 min. CpG genome universal methylated DNA (Millipore) in four different dilutions (1, 1:10, 1:100, 1:1000) was used to derive a standard curve for quantification of methylation levels. The degree of methylation of each gene was expressed as MSP value. MSP values were calculated according to the following formula:

\[
\text{MSP value} = \frac{\text{Gene X-S/gene X-T}}{\beta\text{-Actin-S/\beta\text{-Actin-T}}}
\]

Gene X-S is the level of amplification of the gene of interest for the sample, gene X-T is the level of amplification of the gene of interest in the fully methylated control DNA, β-Actin-S is the level of amplification of β-Actin in the sample and β-Actin-T is the levels of amplification of β-Actin in the fully methylated control DNA\(^{30}\).

Previously published cut-off values were applied to the raw MSP data to calculate the frequency of hypermethylation for the three genes: HPP1: 0.05; RUNX3: 0.02; p16: 0.02\(^{15}\).

**Ethical and statistical considerations**

The aforementioned study was approved by the Medical Ethics Committee at the Academic Medical Center, Amsterdam. All patients signed informed consent, which also included permission for future analysis of biopsy material for biomarker research.

MSP values after application of the cut-offs were compared with Mann-Whitney or Chi-Square tests and expressed as median plus 95% CI\(^{15}\). All statistical analyses were performed using a statistical software package (Statistical Package for the Social Sciences 18.0.1; SPSS Inc, Chicago, Ill). A p < 0.05 was considered statistical significant. Bonferroni correction was used for multiple comparisons.
RESULTS

Sample characteristics

The original database from which the biopsies for this retrospective assessment were selected consisted of 187 endoscopic areas from 77 patients. 129 areas were excluded: 21 because of lack of consensus on histology, gastric metaplasia or cardia carcinoma and 42 because of histology of LGIN or indefinite for dysplasia. 25 areas with IM and 41 with HGIN/EC were excluded because of missing imaging data, inconclusive histology, advanced cancers or insufficient material. 58 areas (TP: 15, FP: 21, TN1: 12, TN2: 10) were included for analysis, out of 44 patients (patient characteristics; table 1, example of AFI and WLE areas; figure 1A-D)).

Methylation biomarkers

We used methylight to assess the methylation of three genes (HPP1, RUNX3 and p16) in this sample cohort. Figure 2a shows the raw MSP values for these three genes in samples categorized according to their dysplasia status. We found no significant difference for HPP1 and RUNX3, whereas p16 had significantly higher methylation levels in biopsies with HGIN/EC compared to non dysplastic samples (p=0.01). When samples were divided according to their AFI status, no significant difference in the raw MSP values was found for the three genes analysed (Figure 2b).

In a previously published study, methylight for these three genes has been performed with DNA extracted from FFPE biopsies with a protocol identical to the one used in this study15. In that study, cut-off values for a positive methylation call were generated (HPP1: 0.05; RUNX3: 0.02; p16: 0.02). We performed a subgroup analysis in the 4 sample groups applying these validated cut-off values (table 2). We detected methylation for HPP1 and RUNX3 in approximately 90% and 50% of all the samples tested, without significant difference across the subgroups. However, p16 showed a

Table 1. Patient characteristics of 44 patients referred for BE surveillance or work-up or treatment for early neoplasia in BE; 58 biopsy samples with different AFI and histological characteristics were selected for analysis.

<table>
<thead>
<tr>
<th>Patient characteristics (N=44)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Male : female</td>
<td>32 : 12</td>
</tr>
<tr>
<td>Age at endoscopy (years)</td>
<td>66 (SD: 11)</td>
</tr>
<tr>
<td>BE length (cm)*</td>
<td></td>
</tr>
<tr>
<td>circular maximum</td>
<td>7 (IQR: 3-9) 8 (IQR: 4-10)</td>
</tr>
<tr>
<td>Position of lesions relative to the GOJ (cm)</td>
<td>1 (IQR: 0-5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample characteristics (N=58)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AFI positive, HGIN/EC (true positive)</td>
<td>15</td>
</tr>
<tr>
<td>AFI positive, NDBE (false positive),</td>
<td>21</td>
</tr>
<tr>
<td>AFI negative, NDBE; HGIN/EC elsewhere (true negative 1)</td>
<td>12</td>
</tr>
<tr>
<td>AFI negative, NDBE; no HGIN/EC elsewhere (true negative 2)</td>
<td>10</td>
</tr>
</tbody>
</table>

*Prague classification. C: circular extent; M: maximum extent.
SD: standard deviation; IQR: interquartile range; GOJ: gastro-oesophageal junction
significant difference in methylation among all subgroups \((p<0.01)\). In the subgroup analysis TP had significantly more p16 methylation than FP \((p=0.03)\) and TN2 \((p<0.003)\), although significance for TP vs FP was lost after correction for multiple testing. Interestingly, p16 methylation levels in TN1 (HGIN/EC elsewhere) were higher than TN2 (no HGIN/EC elsewhere), suggesting a possible p16 field defect in non dysplastic Barrett’s epithelium from patients with dysplasia elsewhere \((p=0.04)\). However, significance was lost after Bonferroni correction.

AFI status was not associated with a significant change in MSP for all the tested markers as showed by the subgroup analyses FP vs TN1 and FP vs TN2 (table 2).

Figure 1A-D. Top: AFI (A) and white light (B) image of a false positive area in a BE. Bright pink areas in the AFI image represent the islands of squamous mucosa. Dark purple AFI appearance marks an area suspicious for dysplasia. Biopsies showed intestinal metaplasia without dysplasia. Bottom: AFI (C) and white light (D) image of a true-positive area: the flat lesion is hardly visible on the white light image, and showed high grade intraepithelial neoplasia in biopsies. The similarity with a false positive area (A,B) can be easily appreciated.
Fig 2a. Raw MSP values for HPP1, RUNX3 and p16, categorized according to dysplasia status. P16 showed a significant increase of methylation in the dysplastic samples (p=0.01). A non-significant trend was seen for RUNX3.

Fig 2b. Raw MSP values for HPP1, RUNX3 and p16, categorized according to AFI status. No significant difference in MSP values between AFI positive and AFI negative samples was observed.

Cyclin A

Cyclin A expression rate (figure 3A) in dysplastic vs. non-dysplastic, AFI positive vs. AFI negative (including HGIN/EC) and False Positive versus True Negative samples is represented in table 3a-c. As shown in table 3a, Cyclin A was significantly more expressed in dysplastic tissue compared to non-dysplastic samples (p=0.003). No difference was observed between AFI positive and AFI negative groups, independent of the presence (table 3b) or absence (table 3c) of dysplasia (p=0.73 and p=0.24 respectively).
P53

P53 expression (figure 3B) in dysplastic vs. non-dysplastic, AFI positive vs. AFI negative (including HGIN/EC) and False Positive versus True Negative samples is represented in table 3d-f. As shown in table 3d, p53 was significantly more expressed in dysplastic tissue compared to non-dysplastic samples (p<0.001). In addition, AFI positive samples showed significantly stronger p53 staining than AFI negative samples (p=0.003, table 3e). However, after exclusion of samples with dysplasia, the correlation between AFI status and p53 positivity was lost (table 3f).

DISCUSSION

In our study we retrospectively evaluated a cohort of patients undergoing surveillance for BE or work-up for early neoplasia who were investigated with AFI. We aimed to investigate the correlation between a panel of biomarkers in tissue samples with AFI- and dysplasia status. We have looked at biomarkers that have been previously shown to correlate with both dysplasia and risk of malignant progression in patients with BE, such as methylation at HPP1, RUNX3 and p16 genes, as well as protein expression of Cyclin A and p53. P16 methylation strongly

Table 2. MSP values of HPP1, RUNX3 and p16 in True Positive (TP), False Positive (FP), True Negative, dysplasia elsewhere (TN1) and True Negative, never dysplasia (TN2) subgroups. A significant difference in p16 MSP was observed for Dysplasia status and TP vs. FP vs. TN1 vs. TN2 (p<0.05), TP vs. FP (p=0.03), TP vs. TN2 (p=0.003) and TN1 vs. TN2 (p=0.04).

<table>
<thead>
<tr>
<th></th>
<th>HPP1</th>
<th>RUNX3</th>
<th>p16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dysplasia</td>
<td>NS</td>
<td>NS</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>AFI status</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>TP vs FP vs TN1 vs TN2</td>
<td>NS</td>
<td>NS</td>
<td>&lt;0.05*</td>
</tr>
<tr>
<td>TP vs FP</td>
<td>NS</td>
<td>NS</td>
<td>0.03*</td>
</tr>
<tr>
<td>TP vs TN1</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>TP vs TN2</td>
<td>NS</td>
<td>NS</td>
<td>0.003</td>
</tr>
<tr>
<td>FP vs TN1</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>FP vs TN2</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>TN1 vs TN2</td>
<td>NS</td>
<td>NS</td>
<td>0.04*</td>
</tr>
</tbody>
</table>

p<0.05 was considered statistically significant.
* significance was lost after Bonferroni correction for multiple testing (p<0.01)

Table 3a. Cyclin A expression in dysplastic (HGIN/EC) vs. non-dysplastic (NDBE) samples. Dysplastic samples showed significantly higher Cyclin A expression (p=0.003).

<table>
<thead>
<tr>
<th>Cyclin A</th>
<th>positive</th>
<th>negative</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGIN/EC</td>
<td>7</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>NDBE</td>
<td>3</td>
<td>39</td>
<td>42</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>48</td>
<td>58</td>
</tr>
</tbody>
</table>
Table 3b. Cyclin A expression in samples from AFI positive lesions (true positive; HGIN/EC and false positive; ND8E) vs. AFI negative (ND8E) lesions. No significant difference in cyclin A expression was observed between both groups (p=0.73).

<table>
<thead>
<tr>
<th>Cyclin A</th>
<th>positive</th>
<th>negative</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFI + (ND8E/HGIN/EC)</td>
<td>7</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>AFI – (ND8E)</td>
<td>3</td>
<td>20</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>48</td>
<td>58</td>
</tr>
</tbody>
</table>

Table 3c. Cyclin A expression in samples from AFI positive lesions without HGIN/EC (false positive) vs. AFI negative lesions (TN1 + TN2). No significant difference was observed between both groups (p=0.24).

<table>
<thead>
<tr>
<th>Cyclin A</th>
<th>positive</th>
<th>negative</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFI + (ND8E)</td>
<td>0</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>AFI – (ND8E)</td>
<td>3</td>
<td>20</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>39</td>
<td>42</td>
</tr>
</tbody>
</table>

Table 3d. p53 expression in dysplastic (HGIN/EC) vs. non-dysplastic (ND8E) samples. Dysplastic samples showed significantly higher p53 expression (p<0.001).

<table>
<thead>
<tr>
<th>P53</th>
<th>positive</th>
<th>negative</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGIN/EC</td>
<td>13</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>ND8E</td>
<td>4</td>
<td>38</td>
<td>42</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>39</td>
<td>56</td>
</tr>
</tbody>
</table>

Table 3e. p53 expression in samples from AFI positive lesions (true positive; HGIN/EC and false positive; ND8E) vs. AFI negative lesions. P53 was significantly higher in samples with a AFI positive status (p=0.003).

<table>
<thead>
<tr>
<th>P53</th>
<th>positive</th>
<th>negative</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFI + (ND8E/HGIN/EC)</td>
<td>15</td>
<td>18</td>
<td>33</td>
</tr>
<tr>
<td>AFI – (ND8E)</td>
<td>2</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>39</td>
<td>56</td>
</tr>
</tbody>
</table>

correlated with dysplasia in BE. When we looked at the correlation with AFI status, we found that only AFI true-positive areas had higher methylation levels, but not the AFI false-positive areas, suggesting that dysplasia rather than AFI positivity correlates with p16 methylation. Interestingly, we found higher p16 methylation levels in non-dysplastic BE biopsies from patients with dysplasia elsewhere in the esophagus compared to non-dysplastic samples from patient without any previous evidence of dysplasia. This finding suggests that BE has some degree of field defect for p16 methylation and this is in agreement with the previous evidence that p16 methylation can be used as biomarker of progression. However, significance was lost.
after correcting for multiple testing and we did not see any difference in p16 methylation in the subgroup analysis between FP and TN2, which is likely due to the small sample size.

With regard to the other two methylation markers, we detected positive HPP1 methylation in about 90% of samples, whereas about 50% of the biopsies tested showed positive RUNX3 methylation. However, we found no statistical significant correlation either for HPP1 or for RUNX3 with dysplasia, nor with AFI positivity. This is in contrast with previous evidence that these methylation markers correlate with dysplasia. It is unlikely that this difference depends on technical aspects, since we worked in identical experimental conditions and we applied the same cut-offs previously published and validated. One of the possible explanations for this discrepancy is that the majority of non-dysplastic samples in our retrospective cohort was derived from patients who had dysplasia elsewhere and therefore a molecular field defect may have played a role. When we looked at raw methylation values, only RUNX3 showed a trend towards higher methylation levels in dysplastic samples, which was not significant. Therefore, the absence of correlation of RUNX3 methylation with dysplasia could depend on our small sample size. Overall, these results would suggest that methylation at these two genes, particularly at HPP1, are early events in BE pathogenesis and may not play a significant role in the malignant progression.

With regards to the two immunohistochemical markers analysed, we found that both Cyclin A and p53 expression was significantly higher in dysplastic compared to non-dysplastic samples (p=0.003 and p<0.001, respectively), which is in concordance with previously reported observations. While no difference was seen between AFI positive and AFI negative groups for cyclin A, p53 did show increased expression in the AFI positive group, which is likely due to the enrichment of dysplasia in the the AFI positive group. In fact, after excluding HGIN/EC samples, the correlation between p53 positivity and AFI status was lost.

For the clinical applicability of biomarkers, it is crucial that they are validated by independent laboratories outside the patient cohorts where they have been first identified. The second aim of this study was to externally validate previously published markers for neoplastic progression in BE. The results of our retrospective study indicate that p16 methylation, cyclin A surface

Figure 3A and 3B. Cyclin A (A) and p53 (B) stained slide, with positive staining epithelial surface cells and High Grade Intraepithelial Neoplasia.
expression and p53 staining are promising biomarkers for the presence of dysplasia in BE. In previous studies, p53 positivity increased the OR for therapeutic intervention to 3.8 (95% CI 1.5-10), and cyclin A positivity or p16 methylation corresponded to an OR for progression of 7.5 (95% CI 1.8-30.7) and 1.74 (95% CI 1.33-2.20) respectively\(^{12,15}\). The results of the present study should be taken into account when planning further prospective validation in future studies.

With regards to the correlation between biomarkers and AFI, none of the biomarkers tested showed positive correlation with the AFI status. A few possible causes for the lack of correlation can be identified.

First of all, autofluorescence imaging is to date one of the few endoscopic wide-field “red-flag techniques” that is able to highlight areas of possible dysplasia in BE. AFI, however, suffers from a high false-positive rate and, as was shown by Curvers et al, dysplastic lesions are still missed with AFI\(^{13}\). It is suggested that with the current AFI algorithm, secondary tissue reactions in the mucosa have major influence on the AFI appearance. Architectural and structural changes, deriving from dysplasia as well as from inflammation can therefore cause AFI positivity on endoscopic imaging. AFI may thus not be the ideal technique to completely correctly identify dysplastic areas and AFI false-positive areas may very well reflect patches of active inflammation. Indeed, Kara et al. describe in a prospective randomized cross-over study on one of the first autofluorescence endoscopy systems a significantly increased rate and severity of acute inflammation in false-positive areas\(^{31}\). Although the system used in the study by Kara et al has a different configuration and operating algorithm than the ETMI system, it targets the same fluorophores in the same wavelentgh-range.

Secondly, although biopsies were carefully targeted to AFI positive areas, sampling error can always occur, leading to misclassification of AFI true-positive areas as false-positive. This should have led to higher biomarker positivity in AFI false-positive compared to AFI true-negative samples, which was not the case in our study. Since all endoscopies were performed by one experienced endoscopist, with a track record for AFI imaging in BE, the real-time assessment of the AFI image was assumed to be very reliable and consistent, as was the performance of the biopsy protocol.

Third, our study has built upon a retrospective cohort of patients whose size may not be large enough to detect subtle differences in biomarkers positivity.

Fourth, and most likely, the biomarkers investigated correspond to subtle molecular changes and may not have the impact at a cellular level to influence the AFI appearance of the tissue. Other biomarkers related to BE malignant progression, such as gross DNA abnormalities (aneuploidy and tetraploidy) can affect nuclear size and density and may potentially affect the AFI feature of BE epithelium.

Therefore, a larger prospective study investigating a more comprehensive panel of biomarkers for neoplasia is needed to address the question of whether AFI can improve detection rate of biomarkers and finally improve patient risk stratification. This is warranted in light of recent evidence that malignant progression of BE may be lower than previously thought\(^{7-9}\). Attention will have to shift to concentrating surveillance and possible treatment on high risk individuals, identified in a risk-stratification model where biomarkers for progression will play an important role.
REFERENCES


