Does Ablative Therapy for Barrett Esophagus Affect the Depth of Subsequent Esophageal Biopsy as Compared With Controls?

Bergein F. Overholt, MD,* Patrick J. Dean, MD,† Joseph A. Galanko, PhD,‡ and Charles J. Lightdale, MD§

Background: Photodynamic therapy (PDT) and radiofrequency ablation (RFA) are associated with high rates of complete eradication of Barrett esophagus (BE). However, if ablation were to induce fibrosis in the regenerated squamous epithelium, then postablation biopsies may not penetrate deeply enough to detect subsquamous intestinal metaplasia (SSIM) and, therefore, complete response rates could be over-estimated.

Goals: To assess the depth of esophageal biopsies from the squamous epithelium of ablation-naive controls and from the neo-squamous epithelium of post-PDT and post-RFA patients to determine if prior ablation results in a reduced proportion of biopsies containing lamina propria (LP) as compared with controls.

Study: Review of archived esophageal specimens from a prospective multicenter cohort study (post-RFA) and 2 retrospective consecutive case series (ablation-naive controls, post-PDT).

Setting: Eight US centers and 1 US gastrointestinal pathology laboratory.

Patients: Ablation-naive controls with GERD, dyspepsia, and/or BE. Post-PDT and post-RFA BE patients with biopsies more than 6 months after achieving complete eradication of BE.

Interventions: Review of endoscopic biopsies from ablation-naive controls, post-PDT patients, and post-RFA patients.

Main Outcome Measurements: Each GI pathology lab processed all tissue and slides. One expert GI pathologist, blinded to cohort, graded the depth of each esophageal specimen as: partial epithelium, full epithelium, LP, muscularis mucosae, or submuco-sa. Each specimen was also evaluated for SSIM.

Results: There were 82 patients [ablation-naive (12), post-PDT (10), post-RFA (60)] with 899 biopsy specimens. The proportion of specimens containing “LP or deeper” was similar between groups: ablation-naive (88%), post-PDT (88%), post-RFA (91%) (\(P > 0.05\)). No SSIM was detected in any group.

Conclusions: There is no difference in esophageal biopsy depth between ablation-naive squamous epithelium and post-PDT/post-RFA neo-squamous epithelium, thus refuting the concern of ablation-induced mucosal resistance to procurement of adequate biopsy specimens. Most biopsies (89% to 91%) from both ablation cohorts were deep enough to detect SSIM, in that they included “LP or deeper.”

Key Words: Barrett esophagus, biopsy depth, photodynamic therapy, radiofrequency ablation, subsquamous intestinal metaplasia (J Clin Gastroenterol 2010;00:000–000)

Barrett esophagus (BE) is a precancerous condition characterized by the conversion of the esophageal lining from a stratified squamous epithelium to a columnar epithelium containing goblet cells (also known as intestinal metaplasia or IM).1,2 This process occurs in the setting of chronic acid exposure and mucosal injury associated with GERD.3 The clinical significance of BE lies in its role as the precursor to esophageal adenocarcinoma, a cancer displaying a rapid increase in incidence and a poor 5-year relative survival rate.4,5

On the basis of the hypotheses that complete eradication of BE could impart a benefit to a patient population with BE in terms of cancer prevention, reduction in need for surveillance, improved cost-effectiveness, and improvement in patient quality of life, among other variables, a number of endoscopic interventions have been evaluated for this purpose, including photodynamic therapy (PDT), argon plasma coagulation (APC), laser, cryotherapy, multipolar electrocoagulation (MPEC), and radiofrequency ablation (RFA). Effective BE ablation presumes complete eradication of the abnormal epithelium, inclusive of its stem cells that are believed to accumulate oncogenetic abnormalities that lead to the phenotypic expression of dysplasia and cancer in the epithelial cells.6-9 Ineffective (incomplete) ablation leaves IM behind and increases the risk for IM to become buried beneath the neo-squamous epithelium.10 This latter phenomenon is also known as subsquamous intestinal metaplasia (SSIM). Safe BE ablation presumes avoiding injury to the subepithelial structures, specifically.
the submucosa and muscularis propria, thereby avoiding complications of stricture formation, bleeding, and perforation.

After the eradication of the BE epithelium, wound healing ensues followed by restoration, in most cases, of a thin nascent squamous epithelium. This neo-squamous epithelium thickens over time to a normal stratified squamous epithelium. Several theories exist as to the source of the neo-squamous epithelium, including encroachment of adjacent squamous epithelium, extension of cells from the submucosal gland duct lining with conversion to squamous epithelium, and circulating pluripotent stem cells which deposit in the wound and transform to squamous stem cells.\(^{10-14}\)

The acute primary endpoint of ablative therapy for BE is to achieve complete endoscopic and histologic eradication of all IM within the esophageal body. Some have hypothesized that if ablative therapy induces fibrosis in the new squamous epithelium and underlying structures, then biopsies may not penetrate deeply enough to detect residual SSIM, should it be present.\(^{7,15}\) If this were the case, a “complete response” could be incorrectly assigned to a patient after ablation. Biddlestone et al\(^{14}\) have reported that SSIM, when it occurs, resides in the deep portion of the epithelium or in the lamina propria (LP), so a postablation biopsy should contain at least LP to be considered “adequate” for ruling out the presence of SSIM.

To address this question of biopsy depth after ablative therapy, we evaluated the depth of esophageal biopsies from squamous epithelium of ablation-naive controls and from neo-squamous epithelium of post-PDT and post-RFA patients to determine if prior ablation reduced the ability to sample the LP with biopsy as compared with control.

**MATERIALS AND METHODS**

**Patients**

Ablation-naive control biopsies were obtained from consecutive patients at 1 center (B.F.O.) who had undergone an endoscopy with biopsy of the esophagus for GERD, dyspepsia, and/or BE and had at least 2 biopsy specimens within the set that contained exclusively squamous epithelium. No patient in this cohort had prior endoscopic esophageal therapy of any type (ablation, resection, injection, banding, radiation, etc.).

Post-PDT biopsies were obtained from consecutive patients from 1 center (B.F.O.) who had undergone PDT with or without endoscopic mucosal resection (EMR) followed by focal rescue ablation with neodymium: yttrium-aluminum-garnet (Nd:YAG) laser or electrocautery for BE containing high-grade dysplasia (BE-HGD). All had shown a complete histologic response for IM (CR-IM) defined as all postablation biopsies negative for IM, and all biopsies were obtained at least 6 months after the last treatment session.

Post-RFA biopsies were obtained from patients as part of a prospective 8-center US clinical trial (AIM-II Trial). At baseline, patients had up to 6 cm of nondysplastic BE and had not undergone earlier ablative therapy or EMR. Treatment included circumferential RFA followed by focal RFA. No therapy other than RFA was applied. All had shown CR-IM and all biopsies were obtained at least 6 months after the last treatment session.

Biopsies obtained from the post-RFA cohort had been collected in an IRB-approved, prospective cohort trial under patient informed consent, whereas biopsies of ablation-naive and post-PDT patients had been collected as part of standard clinical practice and were reviewed as a retrospective consecutive case series. All ablation-naive and post-PDT patients signed a separate, institution-specific patient informed consent form before undergoing any endoscopic procedure.

**Esophageal Biopsy Procurement and Handling**

Biopsies from ablation-naive controls and post-PDT patients were obtained using large capacity biopsy forceps (Radial Jaw 3 forceps, Boston Scientific, Natick, MA) by 1 physician (B.F.O.), whereas biopsies from post-RFA patients were obtained using biopsy forceps deemed large capacity, jumbo, or large cup (Radial Jaw 3 forceps, Boston Scientific, Natick, MA; Large Cup with Needle biopsy forceps, Olympus America, Center Valley, PA; Jumbo GI biopsy forceps, Conmed Corporation, Utica, NY; Captura biopsy forceps, Cook Endoscopy, Bloomingtion, IN) by physicians from 8 centers using a standardized mapping protocol. Post-PDT and post-RFA patients had 4-quadrant biopsies obtained from every 1 to 2 cm of the earlier ablation zone (neo-squamous epithelium). An identical biopsy specimen kit was used for all patients from the 3 cohorts (GI Pathology, Memphis, TN), which included specimen containers with zinc formalin fixative, and a container for overnight shipment to the central pathology laboratory.

**Central Pathology Processing**

Fixed specimens from each container were embedded in paraffin, sectioned in 4-micron slices, affixed to glass slides, and stained with hematoxylin and eosin.

**Histopathologic Evaluation**

One board certified pathologist specializing in gastrointestinal pathology (P.J.D.) interpreted all slides, blinded to patient cohort. A standardized case report form was used for reading each slide set corresponding to 1 patient. For each individual specimen on each slide, the deepest histologic layer present was recorded as: partial epithelium (P-EPI), full epithelium containing complete basement membrane (F-EPI), LP, muscularis mucosae (MM), or submucosa (SM) (see Fig. 1 for representative photomicrographs).

**Study Outcomes**

All outcomes are histology based. Excluded from the analysis were specimens obtained from outside the esophagus (ie, cardia), ablation-naive control specimens that included IM, and any specimens obtained proximal to the earlier ablation zone in post-PDT and post-RFA patients.

The primary outcome is the proportion of esophageal squamous and neo-squamous specimens from each cohort with a depth of LP, MM, or SM (collectively, “LP or deeper”). Secondary outcomes include: (1) the proportion of biopsy specimens from each cohort that include each depth category (P-EPI, F-EPI, LP, MM, SM); (2) the proportion of biopsies for each patient from each cohort that contain “LP or deeper”; and (3) the proportion of biopsy specimens and patients from the post-PDT and post-RFA cohorts contain SSIM.

**Statistical Analysis**

Means, standard errors, and percentages were computed for each cohort for variables of interest. For summary statistics, patient cohorts were compared on categoric
variables through Fisher Exact Test and on continuous variables through Analysis of Variance for means and Kruskal-Wallis test for medians. To compare the proportion of biopsy specimens in each cohort showing each depth category, a logistic regression model was fit with the correlation of the repeated measures for each patient modeled by a compound symmetry correlation matrix. Overall tests were first carried out to determine if there were any differences between the 3 groups. If an overall test was statistically significant, then corresponding pairwise tests were carried out to determine which groups were different. SAS statistical software (SAS Version 9.1, Cary, NC) was used. P values less than 0.05 were considered statistically significant.

RESULTS

Baseline characteristics for patients from each cohort are shown in Table 1. There were 82 total patients: ablation-naive (12), post-PDT (10), and post-RFA (60). Post-PDT and post-RFA patients were more likely to be male compared with ablation-naive controls. Post-PDT patients were older than ablation-naive controls and post-RFA patients. Ablation-naive controls and post-PDT patients were more likely to be of white versus nonwhite race compared with post-RFA patients.

Patients from the post-RFA cohort had earlier undergone a mean of 3.4 ablation sessions without prior EMR and without subsequent stricture formation or dilation. Patients from the post-PDT cohort had undergone baseline EMR in half the cases, followed by a mean of 1.3 PDT and 2.7 focal laser or electrocuttery rescue ablation sessions. PDT patients developed strictures in 50% of cases, with a mean of 4 dilations per patient.

There were 899 esophageal biopsy specimens that met the eligibility criteria for depth analysis (Table 2). The post-RFA group contributed the most specimens (734, 82%) owing to the large size of the cohort (n = 60, 73% of all patients). There were fewer squamous biopsies obtained on average per patient in the ablation-naive cohort (4.2), as compared with the number of neosquamous biopsies obtained from post-PDT (11.5) and post-RFA (12.2).

For the primary outcome, the majority of squamous and neosquamous biopsies contained “LP or deeper” regardless of cohort: ablation-naive (88%), post-PDT (88%), and post-RFA (91%). There was no statistically significant difference between the groups for this outcome (Table 2).

For the secondary depth outcomes, the proportion of specimens containing each of the 5 depth categories was similar between cohorts. Taking this a step further and comparing depth of biopsies on a per patient basis, the
median proportion of biopsies that were “LP or deeper” was high; ablation-naive (95%), post-PDT (94%), and post-RFA (100%) and there was no difference between groups. This indicates that for each patient, rather than exclusively for a cohort’s aggregated biopsies, one can expect that the majority of specimens will be “LP or deeper.”

### TABLE 1. Baseline Characteristics of Study Patients*

<table>
<thead>
<tr>
<th>Ablation-naive Controls</th>
<th>Post-PDT Patients</th>
<th>Post-RFA Patients</th>
<th>Overall</th>
<th>Control Versus Post-PDT</th>
<th>Control Versus Post-RFA</th>
<th>Post-PDT Versus Post-RFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>12</td>
<td>10</td>
<td>60</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Male sex—no. (%)</td>
<td>3 (25)</td>
<td>10 (100)</td>
<td>44 (73)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001 NS</td>
</tr>
<tr>
<td>Age—year</td>
<td>53.3 ± 3.9</td>
<td>71.4 ± 2.4</td>
<td>58.1 ± 1.4</td>
<td>&lt;0.001</td>
<td>&lt;0.05</td>
<td>NS</td>
</tr>
<tr>
<td>Weight—kg</td>
<td>79.7 ± 5.0</td>
<td>84.4 ± 3.0</td>
<td>80.8 ± 2.3</td>
<td>0.79</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ethnicity, no. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>12 (100)</td>
<td>10 (100)</td>
<td>42 (70)</td>
<td>&lt;0.05 NS</td>
<td>NS</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>16 (27)</td>
<td>na</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Black</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (3)</td>
<td>na</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*bPlus-minus values are means ± standard error, “na” indicates that a statistical test was not indicated for comparison between groups, “—” in the pair-wise comparison columns indicates that the group-wise comparison was either not significant or not indicated so a pair-wise comparison was not carried out.

### TABLE 2. Histologic Assessment of Esophageal Biopsy Specimens According to Patient Cohort*

<table>
<thead>
<tr>
<th>Ablation-naive Controls</th>
<th>Post-PDT Patients</th>
<th>Post-RFA Patients</th>
<th>Overall</th>
<th>Control Versus Post-PDT</th>
<th>Control Versus Post-RFA</th>
<th>Post-PDT Versus Post-RFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>12</td>
<td>10</td>
<td>60</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Specimens, no.</td>
<td>50</td>
<td>115</td>
<td>734</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Specimens per patient</td>
<td>4.2 ± 0.7</td>
<td>11.5 ± 1.7</td>
<td>12.2 ± 0.6</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001 NS</td>
</tr>
<tr>
<td>Specimens “LP or Deeper,” no. (%)</td>
<td>44 (88)</td>
<td>101 (88)</td>
<td>668 (91)</td>
<td>0.39</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Specimens “LP or Deeper,” per patient</td>
<td>3.7 ± 0.7</td>
<td>10.1 ± 1.6</td>
<td>11.1 ± 0.6</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001 NS</td>
</tr>
</tbody>
</table>

*bPlus-minus values are means ± standard error, “na” indicates that a statistical test was not indicated for comparison between groups, “—” in the pair-wise comparison columns indicates that the group-wise comparison was either not significant or not indicated so a pair-wise comparison was not carried out.

**F-EPI indicates full epithelium containing complete basement membrane; LP, lamina propria; MM, muscularis mucosae; P-EPI, partial epithelium; SM, submucosa.**
No specimen from any cohort was found to have SSIM. We did not include a fibrosis outcome in our prospectively designed case report forms for this histology assessment; therefore, we were unable to assess fibrosis or degree of fibrosis in these study specimens.

DISCUSSION

We evaluated the deepest histologic layer present in a large number of esophageal biopsies obtained from the squamous epithelium of ablation-naive controls and from the neosquamous epithelium of post-PDT and post-RFA patients. Our aim was to determine whether ablative therapy altered the likelihood that a biopsy specimen would contain “LP or deeper” structures, as compared with a control. We found that the proportion of biopsies that contained “LP or deeper” was not different between the ablation-naive controls (88%) and post-PDT and post-RFA patients (88% and 91%, respectively), therefore, it seems that ablation does not alter the esophageal epithelium to prevent sampling of the deeper layers of the mucosa.

Although ablative therapy has been shown in many studies to achieve a high rate of complete eradication of BE, one concern is that residual BE tissue can become buried after therapy. This finding is concerning in that SSIM cannot be endoscopically visualized or targeted for biopsy during surveillance, and therefore, may progress to cancer in an occult manner.16,17 Recent studies have revealed that SSIM is actually present in a large number of ablation-naive BE patients. Sharma et al18 found SSIM in 38.5% of squamous island biopsy specimens from ablation-naive BE patients, suggesting that this may result from frequent surveillance biopsies while on PPI therapy. Bronner et al19 compared 2 groups of BE-HGD patients: PDT plus omeprazole versus omeprazole only; all underwent the same PPI and biopsy surveillance regimen. In 33,568 biopsy specimens obtained during follow-up, the proportion of patients with SSIM was similar between groups (30% in PDT vs. 33% in PPI only). Shaheen et al20 reported the results of a randomized sham-controlled trial comparing RFA versus sham for patients with BE containing LGD or HGD. At baseline, 25% of patients had SSIM, whereas at 1 year after randomization to RFA or sham, SSIM was present in 5% of RFA patients and 40% of sham patients. Chennat et al21 evaluated initial mucosectomy specimens in ablation-naive individuals with HGD or IMC undergoing complete Barrett esophagus EMR and reported SSIM in 28% of patients. These data strongly suggest that future trials of endoscopic interventions for BE should report not only the posttreatment SSIM prevalence, but the baseline prevalence as well, given the high proportion of patients with SSIM as part of their natural history.

Although there is a risk of occult cancer progression that has been suggested in relation to SSIM, Hornick et al22 reported that SSIM exhibits fewer oncogenetic abnormalities than IM present at the surface of the epithelium. They hypothesized that SSIM is isolated from luminal acid, bile, and enzymes by the overlying squamous epithelium and therefore, has less inflammation and likelihood for genetic alteration. Nonetheless, the objective of ablative therapy is complete eradication of all IM within the esophageal body, including any SSIM present at baseline.

What tissue layer must a postablation esophageal biopsy contain to confirm, with a high degree of certainty, the absence of SSIM? Biddlestone et al23 have reported that, when present, SSIM resides in the deep portion of the epithelium and the LP, so biopsies must contain at least those layers to rule out the presence of SSIM. On the basis of these data, we designed our study’s primary endpoint to compare the proportion of biopsies between cohorts that contained “LP or deeper.”

Others have recently assessed biopsy depth after ablative therapy, albeit in different patient populations. Shaheen et al24 evaluated histologic depth in squamous and columnar tissue pre-RFA and post-RFA. They found “LP or deeper” in 78% of squamous and 98% of columnar biopsy specimens, regardless of whether the tissue had been ablated or not, and therefore, concluded that tissue type, rather than history of ablation predicted how deep a biopsy penetrated. In comparison, Pouw et al25 reported “LP or deeper” in 37% of post-RFA neo-squamous biopsies and 36% of ablation-naive squamous biopsies.

Our study’s outcomes, and that of Shaheen and Pouw, found no difference between ablation-naive and postablation biopsy depth. All 3 studies concluded that ablative therapy did not impair the ability to obtain “LP or deeper” in subsequent biopsies compared with their controls. However, Pouw reports “LP or deeper” in only about 40% of neosquamous specimens, whereas Shaheen and our study report “LP or deeper” in over 3 quarters of neosquamous specimens. The reason for this difference is unclear. Although the interstudy absolute proportion of subepithelial biopsies postablation vary somewhat, each study confirmed internally that biopsy depth of untreated squamous epithelium and earlier treated neosquamous epithelium were similar.

Strengths of our study include the use of a central gastrointestinal pathology laboratory with standardized specimen processing, interpretation by a single expert gastrointestinal pathologist blinded to cohort, and that the largest group of specimens (post-RFA) was collected in a controlled manner in a prospective cohort study.

Limitations of our study include the retrospective review of pathology specimens and the inclusion of some specimens that were collected as part of retrospective consecutive case series (ablation-naive control and post-PDT). Furthermore, the number of specimens evaluated in each cohort was not equally distributed because we relied on the large prospectively collected specimens from the AIM-II RFA trial as the main ablation group, and then collected ablation-naive controls and post-PDT specimens as comparators. In future studies, perhaps a more optimal design would comprise a prospective collection of biopsies from patients who have undergone successful ablation, obtaining samples from the proximal untreated esophagus to compare with the treated distal esophagus, allowing patients to act as their own controls. Additional limitations include operator-dependent biopsy procurement techniques, lack of uniformity of specimen orientation, and variation in biopsy forceps type. The reported time interval between the last ablation and the procurement of biopsies assessed for depth (at least 6 mo), and the use of EMR in some PDT patients, may have confounded our results if the resultant neosquamous epithelium in some patients had not had enough time to mature and established a normal thickness.

In conclusion, we found no significant difference in esophageal biopsy depth when comparing specimens from ablation-naive squamous epithelium with post-PDT/post-RFA neo-squamous epithelium, thereby refuting the hypothesis that ablation imparts a change to the mucosa that
limits the ability to sample structures deep to the epithelium. The majority of biopsies obtained from the postablation cohorts included a depth of “LP or deeper,” indicating a depth adequate to detect SSIM, should it be present. Our findings indicate that published reports have not overestimated the efficacy of ablation interventions by failing to identify occult SSIM owing to inadequate biopsy depth.

ACKNOWLEDGMENTS

The authors would like to acknowledge other participants in the AIM Trial Investigator Group for their contributions in the conduct of the AIM-H clinical trial and the collection of biopsy specimens that were analyzed for depth in this study: David E. Fleischer, Virender K. Sharma, Alvaro Reymunde, Michael B. Kimney, Ram Chuttani, Kenneth J. Chang, Nilda Santiago, Douglas K. Pleskow, and Kenneth K. Wang.

REFERENCES