

Published in final edited form as:

*Clin Gastroenterol Hepatol.* 2013 October ; 11(10): . doi:10.1016/j.cgh.2013.03.039.

## Efficacy and Durability of Radiofrequency Ablation for Barrett's Esophagus: Systematic Review and Meta-analysis

ERIC S. ORMAN, NAN LI, and NICHOLAS J. SHAHEEN

Center for Esophageal Diseases and Swallowing, Division of Gastroenterology and Hepatology, Department of Medicine, University of North Carolina, Chapel Hill, North Carolina

### Abstract

**BACKGROUND & AIMS**—In patients with Barrett's esophagus (BE), radiofrequency ablation (RFA) safely and effectively eradicates dysplasia and intestinal metaplasia. We aimed to determine the efficacy and durability of RFA for patients with dysplastic and nondysplastic BE.

**METHODS**—We performed a systematic review and meta-analysis of studies identified in PubMed and EMBASE that reported the proportion of patients treated with RFA who had complete eradication of dysplasia (CE-D) and intestinal metaplasia (CE-IM), and the proportion of patients with recurrent IM after successful treatment. Pooled estimates of CE-D, CE-IM, IM recurrence, and adverse events were calculated.

**RESULTS**—We identified 18 studies of 3802 patients reporting efficacy and 6 studies of 540 patients reporting durability. Ten were prospective cohort studies, 9 were retrospective cohort studies, and 1 was a randomized trial. CE-IM was achieved in 78% of patients (95% confidence interval [CI], 70%–86%) and CE-D was achieved in 91% (95% CI, 87%–95%). After eradication, IM recurred in 13% (95% CI, 9%–18%). Progression to cancer occurred in 0.2% of patients during treatment and in 0.7% of those after CE-IM. Esophageal stricture was the most common adverse event and was reported in 5% of patients (95% CI, 3%–7%). Confidence in most summary estimates was limited by a high degree of heterogeneity, which did not appear to be caused by single outlier studies.

**CONCLUSIONS**—Treatment of BE with RFA results in CE-D and CE-IM in a high proportion of patients, with few recurrences of IM after treatment and a low rate of adverse events. Despite the large amount of study heterogeneity, these data provide additional information for patients and providers to make informed treatment decisions.

### Keywords

Esophageal Cancer; Gastroesophageal Reflux; Prevention; Endoscopy

---

Barrett's esophagus (BE) is a precancerous condition characterized by the replacement of the normal stratified squamous epithelium of the distal esophagus by intestinal metaplasia (IM), affecting 1%–2% of the general population.<sup>1–3</sup> Multiple endoscopic ablative

---

© 2013 by the AGA Institute

Reprint requests: Address requests for reprints to: Nicholas J. Shaheen, MD, MPH, Center for Esophageal Diseases and Swallowing, University of North Carolina School of Medicine, CB #7080, Chapel Hill, North Carolina 27599-7080. nicholas\_shaheen@med.unc.edu; fax: (919) 843-2508.

#### Conflicts of interest

This author discloses the following: Dr Shaheen has received research grants from BARRX Medical, Oncoscope, CSA Medical, Takeda, and AstraZeneca; and is a consultant for CSA Medical, Astra-Zeneca, Takeda, Oncoscope, and NeoGenomics. The remaining authors disclose no conflicts.

techniques have been developed for BE, with the goal of eradicating IM and preventing neoplastic progression to esophageal adenocarcinoma (EAC).<sup>4-7</sup> Of these techniques, radiofrequency ablation (RFA) is used commonly and was shown in a randomized controlled trial (RCT) to have low complication rates, substantial rates of complete eradication of dysplasia (CE-D) and IM (CE-IM), and a decrease in progression to cancer.<sup>8</sup>

Although RFA is safe and effective in eradicating IM, the absolute magnitude of the benefit has not been well described. Different studies have varied considerably in their reports of absolute rates of CE-D and CE-IM and in estimates of durability of the neosquamous epithelium that appears after RFA.<sup>9-11</sup> These studies are often from tertiary care centers, subject to local expertise with limited generalizability; have small sample sizes that limit the precision of effect estimates; and have varied in the inclusion of different histologic grades of BE. The latter issue is important because the risk-benefit ratio of ablation changes with the risk of malignancy, which is tied closely to histologic grade. In addition, pretreatment histology may predict treatment efficacy, although such a relationship has been seen inconsistently.<sup>12,13</sup> Inclusion of a limited histology distribution in an individual study may not allow for adequate power to make such comparisons. Estimates of adverse events also may be unreliable because of insufficient power in individual studies.

We therefore performed a systematic review and meta-analysis of studies of RFA for dysplastic and nondysplastic BE (NDBE). We aimed to determine the proportion of patients achieving CE-D and CE-IM after treatment with RFA as well as the proportion with recurrence of IM after successful treatment. We also sought to evaluate the association of pretreatment histology with CE-D and CE-IM and the incidence of adverse events.

## Methods

### Search Strategy and Eligibility Criteria

We followed the Meta-analysis Of Observational Studies in Epidemiology guidelines for the conduct and reporting of systematic reviews of observational studies.<sup>14</sup> Two authors (E.S.O. and N.L.) independently searched the PubMed and EMBASE databases for relevant articles on August 24, 2012. PubMed was queried with the search terms “((Barrett) OR Barrett’s) AND (radiofrequency OR radio-frequency OR ablation)” and EMBASE was queried with “Barrett AND (‘radiofrequency’/exp OR radiofrequency OR radio+frequency OR ablation).” Abstracts identified in EMBASE that were published in 2011 and 2012 were included. Two additional manuscripts from the authors’ institution accepted for publication and in press at the time of the search but not yet indexed in either database also were included.<sup>15,16</sup>

The identified records then were reviewed independently according to strict eligibility criteria. We excluded letters to the editor; editorials; non-English language studies; nonclinical or nonhuman studies; review articles; case reports; studies that were not observational or RCTs; studies without biopsy-proven BE; studies that did not use focal RFA; studies of patients who had received other ablative therapies (with the exception of endoscopic mucosal resection [EMR] in combination with RFA); studies not reporting either efficacy (CE-IM or CE-D) or durability (histologic recurrence after CE-IM or CE-D); efficacy studies with mean or median follow-up periods less than 12 months from the first RFA session; durability studies with mean or median follow-up periods less than 12 months from the first post-RFA endoscopy showing eradication; studies with fewer than 20 subjects receiving RFA; duplicate reports of study samples; abstracts published before 2011; and studies not reporting a follow-up duration. Abstracts published before 2011 were excluded because those that have not been converted into manuscript form in the 2 years after submission may have significant flaws in methodology or interpretation that have limited

their acceptability for publication. After exclusion of ineligible abstracts, a full-text review of remaining studies was performed independently using the same eligibility criteria, with discrepancies resolved by consensus. We attempted to contact study authors for clarification when multiple reports potentially described the same patient population. For multiple reports describing a common patient sample, we included the most recent report unless only the older report described our a priori outcome variables. Duplicate reports were included, however, if 1 reported efficacy outcomes and the second reported durability. After exclusion of full-text records, the reference sections of included articles were searched manually for additional records.

### Data Collection

Data were abstracted from each study and organized into formalized evidence tables independently by the authors. We recorded study characteristics including year of publication, country, study design, setting, inclusion and exclusion criteria, RFA and surveillance protocol, sample size, proton pump inhibitor use, use of circumferential RFA and EMR, number of RFA sessions and surveillance endoscopies, and follow-up duration. For the RFA protocols, we recorded the timing and type of RFA (circumferential vs focal). For the surveillance protocols, we recorded intervals of surveillance endoscopy, inclusion of cardia biopsies, and follow-up evaluation start time (ie, the time at which treatment was considered complete, and surveillance was begun). For studies that reported both efficacy and durability, sample size was considered separately for each outcome. Recorded patient characteristics included age, sex, BE length, and pretreatment dysplasia assessment. The quality of each study was assessed using the previously validated Downs and Black<sup>17</sup> instrument, which can assess both randomized and nonrandomized studies. This tool assesses the quality of reporting, external validity, bias, confounding, and power using a checklist of 27 items, and scores reports from 0 to 32, with higher scores representing greater methodologic quality. After abstraction, the authors reviewed the evidence tables and discrepancies again were resolved by consensus.

### Outcomes

The primary efficacy outcomes were CE-IM, defined as the absence of IM on endoscopy and histology after RFA; and CE-D, defined as the absence of dysplasia on histology (in subjects with baseline dysplastic BE). Residual IM at the gastro-esophageal junction was considered to be failure of CE-IM in the efficacy analysis and recurrence in the durability analysis, but IM in the cardia was not. The presence of IM in these locations was determined according to the descriptions in the manuscripts. Efficacy outcomes were tabulated according to baseline histology where available. Progression to EAC during treatment was recorded as well. The primary durability outcome was recurrence of IM defined histologically after CE-IM. The presence of dysplasia or EAC at the time of recurrence also was recorded. Adverse events (most commonly strictures, pain, and bleeding) were recorded as secondary outcomes. These were ascertained based on the individual study definitions of adverse events.

### Statistical Analysis

We calculated the proportion of patients who met the primary efficacy and durability outcomes in each study. The denominator for the IM recurrence durability outcome comprised only those patients who had achieved CE-IM after RFA. The proportion of patients who achieved CE-IM and CE-D was calculated for each pretreatment histologic group, and unadjusted risk ratios (RRs) were calculated to compare the outcomes between these groups, with low-grade dysplasia (LGD) as the comparator. LGD was chosen as the comparison group because it was present in more studies than NDBE and because it is the lowest grade of dysplasia that can be compared for the CE-D end point.

Statistical analysis was performed using Stata version 12.1 (StataCorp LP, College Station, TX) and Open Meta-Analyst (Tufts Medical Center, Boston, MA).<sup>18</sup> To determine the pooled proportion of patients achieving CE-IM and CE-D and the proportion of those who had IM recurrence after CE-IM, we performed meta-analyses using a random-effects model.<sup>19</sup> We used the Freeman–Tukey double-arcsine transformation for variance stabilization for meta-analysis of proportions.<sup>20</sup> Random-effects models also were used to estimate the pooled proportion of patients experiencing adverse events and the RR of CE-IM and CE-D for patients with pretreatment NDBE (for CE-IM only), high-grade dysplasia (HGD), and intramucosal carcinoma (IMC) compared with LGD. Heterogeneity was assessed using the  $I^2$  statistic, which estimates the percentage of variation across studies caused by study heterogeneity.<sup>21</sup>  $I^2$  values of 25%, 50%, and 75% represent low, moderate, and high levels of heterogeneity, respectively. To identify potential sources of heterogeneity, analyses were repeated for each subset of the studies obtained by excluding potential outliers. In addition, analyses were stratified according to whether they were published as an abstract or peer-reviewed article. Additional stratified analyses were performed by study quality (score above vs below the median score), design (prospective vs retrospective), and sample size (above vs below the median sample size). Non-normally distributed continuous variables were compared using the Wilcoxon rank-sum test. Ninety-five percent confidence intervals (CIs) were calculated for all point estimates, and a  $P$  value less than .05 was considered statistically significant.

## Results

### Search Results

Of the 1191 unique records identified, 1069 were excluded based on title and abstract review alone (Figure 1). An additional 102 records were excluded after full-text review because of non-English language ( $n = 2$ ), no RFA ( $n = 3$ ), no focal RFA ( $n = 4$ ), the use of other ablative therapies ( $n = 16$ ), no report of either efficacy or durability ( $n = 15$ ), mean follow-up period less than 12 months ( $n = 4$ ), fewer than 20 subjects ( $n = 3$ ), duplicate reports of study samples ( $n = 23$ ), abstracts published before 2011 ( $n = 30$ ), and no report of follow-up duration ( $n = 2$ ). We therefore included 18 efficacy studies of 3802 patients (13 full-text articles and 5 abstracts),<sup>9,10,12,13,15,22–34</sup> and 6 durability studies of 540 patients (5 full-text articles and 1 abstract).<sup>10,11,16,24,25,32</sup>

All included studies were published between 2008 and 2012 and were performed in either North America or Europe (Table 1). One study was an RCT of RFA vs stepwise radical endoscopic resection.<sup>25</sup> Another was a cohort extension of the sham-controlled Ablation of Intestinal Metaplasia (AIM) RCT,<sup>8</sup> in which subjects initially randomized to the sham group were offered a cross-over to RFA after 1 year with continued follow-up evaluation for both efficacy and durability in the overall cohort.<sup>24</sup> The remaining studies were either prospective or retrospective cohort studies.

### Study Quality

Study quality ranged from 7 to 24 on the Downs and Black<sup>17</sup> instrument, with a median of 13.5 (Table 1) of a total possible score of 32. Published articles had a trend toward higher quality than abstracts (median, 15 vs 11;  $P = .07$ ). Studies consistently scored well on description of study aims, description of main findings, clarity in reporting of unplanned retrospective analyses, appropriate use of statistical tests, and use of accurate main outcome measures, and consistently scored poorly on generalizability of study sample, generalizability of treatment, blinding of subjects and assessors, and randomization. In aggregate, the studies performed the best on the reporting sub-scale and the worst on the external validity subscale.

## Efficacy Study Characteristics

There was substantial variability among the efficacy studies in terms of the allowed pretreatment histology, BE length restrictions, and RFA protocols (Table 1). Of the 18 studies, 11 included patients with pretreatment NDBE, 14 included LGD, 17 included HGD, and 10 included IMC. Five studies included all histologic groups, and 1 included only 1 group (NDBE). Five studies specified BE length restrictions, and the specified maximum ranged from 5 to 12 cm. Two studies required a minimum length of 2 or 3 cm. The protocol for 9 of the 18 efficacy studies specified the use of circumferential ablation for the initial treatment session, and the protocol was not reported in 5 studies (1 article and 4 abstracts). One study required an absence of IM on 2 consecutive endoscopies to qualify as CE-IM<sup>32</sup>; the remainder required only 1 endoscopy. Sample sizes ranged from 22 to 2135, although only 3 of the 18 studies had more than 137 patients. The largest was a report from the US RFA Registry, which collected data from patients treated at 148 community-based and academic institutions across the United States.<sup>34</sup> The sample size from this study was greater than all of the other studies combined. Follow-up evaluation ranged from 12 to 31 months, with a median of 20.5 months. Two studies did not specify an exact follow-up duration but were noted to be longer than 12 months.

Patient and treatment characteristics in the efficacy studies are shown in Table 2. A total of 41% of patients with reported pretreatment histology had NDBE, 23% had LGD, 31% had HGD, and 6% had IMC. Mean age was in the 60s for most studies (range, 56–71 y). Men comprised 66% to 91% of patients in the studies (mean, 80%). Mean BE length ranged from 3 to 8 cm. Patients were treated with 1 to 3.4 RFA sessions on average, with 2 to 3 sessions in most studies. The vast majority of studied patients received circumferential ablation. Only 4 of 11 studies reported fewer than 100% use of circumferential RFA, and none reported less than 50%. In contrast, there was wide variability in the use of EMR, ranging from 0% to 96%.

## Durability Study Characteristics

Details of the durability studies are shown in Table 1. The abstract included patients with any pretreatment histology, whereas the published articles were more restrictive. All but 1 study included patients with dysplasia. Similar to the efficacy studies, BE length restrictions were variable. With the exception of the AIM II cohort, which specified a single surveillance endoscopy 5 years after starting treatment for NDBE,<sup>11</sup> surveillance protocols were similar across studies in terms of timing. Most specified surveillance endoscopy at 6 and 12 months after treatment, then annually, with several deviations according to the pretreatment grade of dysplasia. However, the protocols differed in their specification of cardia biopsies, and only 3 of 6 started follow-up evaluation at the first post-treatment endoscopy with biopsies showing eradication. Two started follow-up evaluation at the last treatment session, and 1 started after eradication was confirmed on a second post-treatment endoscopy with biopsies. The median sample size was 80 (range, 20–218), and the median follow-up period was 16.5 months (range, 13–51 mo).

Pretreatment histology distributions varied substantially, although HGD was well represented in most of the durability studies (Table 2). Twenty-seven percent of patients with reported pretreatment histology had NDBE, 13% had LGD, 52% had HGD, and 8% had IMC. Similar to the efficacy studies, the mean age was in the 60s for most studies, most subjects were male, and the mean BE length was between 3 and 8 cm. Only 4 of the studies reported the number of surveillance sessions, which was between 1 and 3.

## Efficacy Outcomes

The pooled percentage of patients achieving CE-IM was 78% in the random-effects model (95% CI, 70%–86%) and was slightly higher among published articles compared with abstracts (Figure 2A). Three of the 18 efficacy studies reported point estimates of 50% or less, whereas 6 studies reported CE-IM of more than 90%. Studies with quality scores above the median had a greater percentage with CE-IM (84%; 95% CI, 77%–90%) than the lower-quality studies (68%; 95% CI, 51%–83%). Similarly, the prospective studies had a greater percentage with CE-IM (84%; 95% CI, 71%–93%) than the retrospective studies (71%; 95% CI, 58%–82%). CE-IM did not differ by study sample size.

CE-D was achieved in 91% of patients (95% CI, 87%–95%) in the random-effects model and did not differ substantially between published articles and abstracts (Figure 2B). CE-D also was unchanged after stratification by study quality, design, and sample size.

Response to therapy varied according to pretreatment histology (Table 3). Patients with pretreatment HGD were less likely than patients with LGD to achieve CE-IM (RR, 0.92; 95% CI, 0.87–0.98) and CE-D (RR, 0.94; 95% CI, 0.91–0.97). CE-D also was less likely to occur for patients with IMC compared with patients with LGD, although this difference was not statistically significant (RR, 0.94; 95% CI, 0.87–1.01). CE-IM was more likely to occur in patients with NDBE, although this difference also was not statistically significant (RR, 1.12; 95% CI, 0.90–1.40).

A total of 9 patients in 4 studies<sup>9,15,24,27</sup> progressed to EAC during treatment, for an overall risk of 0.1% per year.

## Durability Outcomes

After CE-IM, IM recurred in 13% (95% CI, 9%–18%) (Figure 3) and was not substantially different when the abstract was excluded (11%; 95% CI, 8%–15%). Higher-quality studies had lower recurrences (11%; 95% CI, 7%–16%) than lower-quality studies (17%; 95% CI, 13%–22%). IM recurrence was similar in larger and smaller studies. At recurrence, dysplasia was noted in 5 patients from 1 study<sup>16</sup> and EAC was noted in 4 patients from 2 studies,<sup>16,24</sup> for an overall risk of 0.9% and 0.7%, respectively, over 1.5 years.

## Adverse Events

Strictures were the most common reported adverse event (pooled estimate, 5%; 95% CI, 3%–7%), followed by pain (3%; 95% CI, 1%–6%) and bleeding (1%; 95% CI, 1%–2%). Pain was more common in abstracts (8%; 95% CI, 0%–31%) and prospective studies (6%; 95% CI, 1%–13%) compared with published articles (3%; 95% CI, 1%–5%) and retrospective studies (2%; 95% CI, 1%–4%), mostly as a result of 1 outlier study that reported pain in 44% of patients (Martinek et al).<sup>33</sup> Other adverse events were similar in stratified analyses.

## Heterogeneity

There was substantial heterogeneity in the primary efficacy and durability outcomes, with  $I^2$  values of 96% for CE-IM, 78% for CE-D, and 46% for IM recurrence (Figures 2 and 3). Stratification by publication status (article vs abstract) partially attenuated the heterogeneity in the efficacy analyses; however, in general, heterogeneity persisted. Stratification by publication status could not be performed for IM recurrence because there was only 1 abstract examining this outcome. Stratification of CE-IM by quality, design, and sample size did not reduce heterogeneity significantly (minimum  $I^2 = 85%$  for high-quality studies). For CE-D, heterogeneity for small studies was low ( $I^2 = 0%$ ) but was otherwise substantial, with the next lowest value of 62% for prospective studies. Heterogeneity was low for IM

recurrence in low- ( $I^2 = 0\%$ ) and high-quality ( $I^2 = 26\%$ ) studies considered separately and in small studies ( $I^2 = 0\%$ ). To search for specific sources of heterogeneity, summary effect estimates were regenerated while excluding 1 study at a time from the analyses. In no case did the exclusion of a single study result in a shift of greater than 3% in the summary estimate, suggesting that no single outlier was the predominant cause of the heterogeneity noted in the results.

There was also heterogeneity in the model comparing CE-IM between patients with pretreatment NDBE and LGD (Table 3). Excluding the study by Lyday et al<sup>13</sup> resulted in a statistically significant difference favoring NDBE (RR 1.21; 95% CI, 1.13–1.30;  $I^2 = 0\%$ ).

There was significant heterogeneity for the estimate of strictures ( $I^2 = 59\%$ ) and pain ( $I^2 = 85\%$ ). Stratification of studies estimating strictures resulted in improved heterogeneity for prospective studies ( $I^2 = 35\%$ ) and abstracts ( $I^2 = 0\%$ ). Stratification did not resolve the heterogeneity in the pooled estimate of pain. Excluding the Martinek et al<sup>33</sup> study reduced the pain estimate from 3% to 2% and did not markedly improve heterogeneity ( $I^2 = 74\%$ ).

## Discussion

RFA has emerged as a preferred method for the endoscopic ablation of BE.<sup>1</sup> It safely and effectively eradicates dysplasia and IM, and for those with HGD it reduces progression to cancer.<sup>8</sup> In this systematic review and meta-analysis, we estimated the proportion of patients treated with RFA who achieved eradication of dysplasia and IM and the proportion of patients with recurrence of IM after successful treatment. We found that RFA resulted in CE-D in 91% of patients and CE-IM in 78% of patients, and that 13% of patients had recurrent IM after successful eradication.

To make appropriate informed decisions about the use of RFA, patients and providers need to be well versed in the risks and benefits of the procedure. However, making predictions about treatment outcome can be difficult when based on single-center studies with small sample sizes and large standard errors. This systematic review provides further guidance for these decisions by combining the results of several studies. Although heterogeneity was substantial, the pooled estimates likely are representative of the true efficacy and effectiveness of RFA as estimated for baseline degree of dysplasia. CE-D and CE-IM proportions reported here are similar to the AIM dysplasia RCT of RFA compared with sham ablation for dysplastic BE.<sup>8</sup> That trial was conducted rigorously, with close follow-up evaluation of subjects, and therefore may be the best estimate of efficacy in ideal circumstances. In contrast, the US RFA Registry contains data collected from 148 academic and nonacademic institutions across the country, without standardized treatment protocols.<sup>34</sup> Our pooled CE-D and CE-IM proportions also were similar to this registry, reflecting a good estimate of effectiveness in a real-world setting. In addition, exclusion of individual studies did not change the pooled point estimates substantially, so the estimates appear robust to the effects of individual studies despite the large amount of heterogeneity.

We found that efficacy differed based on the baseline degree of dysplasia, with a trend toward reduced efficacy for those with more advanced histology (Table 3). This relationship was seen in some studies,<sup>9,12,24,34</sup> but not others,<sup>13,15</sup> and the reason for such a relationship is not entirely clear. RFA using standard energy settings results in a uniform depth of tissue destruction.<sup>35</sup> Therefore, in the absence of nodularity or invasive cancer, RFA might be expected to result in equivalent eradication of IM regardless of the degree of dysplasia. One possible explanation for this finding is that histology is a surrogate marker of BE length, with greater length associated with more advanced histology.<sup>36</sup> CE-IM rates are lower in patients with longer BE segments because of the greater surface area that must be treated

and the greater likelihood of missing a portion of BE.<sup>8,13,15</sup> It was not possible to adjust our findings for BE length because studies vary in the BE length cut-off value used for comparison. However, a meta-analysis of individual patient-level data potentially could resolve this issue. Another potential explanation is that HGD is more likely than LGD or NDBE to harbor malignancy,<sup>37</sup> which is also more likely to penetrate into the submucosa, beyond the depth of ablation.<sup>38</sup> This explanation is less likely because such failures of CE-IM would be expected to translate into increased progression to EAC, which nevertheless was low in this review. Finally, BE of more advanced histology may be qualitatively different than NDBE with respect to its susceptibility to destruction by thermal methods. Whether the same mutations that cause the phenotypic aberration of these cells also might confer resistance to destruction by RFA is unclear.

EAC developed in 9 of 3802 patients during 20.5 months of treatment (0.1% per year). Of the 7 patients with pretreatment histology reported, 1 patient had LGD, 5 patients had HGD, and 1 patient had IMC. Assuming a worst-case scenario allotting the remaining 2 cases to each group, the worst rate of progression to EAC would be 0.09% for NDBE, 0.2% for LGD, 0.4% for HGD, and 0.9% for IMC. In comparison, studies have shown a 0.12% to 0.6% annual risk of EAC for patients with NDBE<sup>3,36</sup> and annual risks of 1.7% and 6.6% for patients with LGD and HGD, respectively.<sup>39</sup> Although the rates of progression reported in postablation patients in this review appear lower than those reported in natural history studies, direct comparisons between the 2 groups cannot be made owing to potential differences in patient populations, as well as the varied follow-up periods and surveillance protocols.

Durability of RFA appeared reasonable with a low, but not inconsequential, rate of IM recurrence. This finding suggests that the majority of patients who achieve CE-IM will continue to maintain squamous epithelium after 1.5 years of follow-up evaluation. However, it also should serve as a warning that patients need to be maintained in a surveillance program after completion of therapy because a sizeable minority (9%–18%) did develop recurrence, and EAC occurred in 0.7%.

The safety profile of RFA showed strictures in 5% of patients. This figure compares favorably with other endoscopic interventions for BE. Lewis et al<sup>40</sup> reported symptomatic stricture formation in 25% of patients undergoing EMR monotherapy at a single center. In a multicenter study of photodynamic therapy, 36% developed esophageal strictures.<sup>41</sup> Strictures arose in only 3% of patients treated with endoscopic spray cryotherapy in a multicenter cohort study,<sup>6</sup> although there are less data for this modality compared with RFA. In a head-to-head comparison, strictures developed in 88% of patients treated with stepwise radical endoscopic resection compared with just 14% of those treated with RFA.<sup>25</sup>

Interestingly, pain was reported in only 3% of patients. However, there was significant heterogeneity in the estimate of pain ( $I^2 = 85%$ ), with 2 outlier estimates of 15% and 44%.<sup>22,33</sup> This heterogeneity may be caused by differences in ascertainment of pain because modest discomfort is common after RFA. In the AIM Dysplasia Trial, the median chest pain score after RFA was 23 on a visual analogue scale from 0 to 100.<sup>8</sup> However, only 2 patients were classified as having an adverse event related to chest pain or discomfort. In clinical practice, practitioners often prescribe oral narcotics and viscous lidocaine after RFA sessions.

There were several limitations of this systematic review. The quality of the included studies was suboptimal, with a median score of 13.5 of a total possible score of 32.<sup>17</sup> Only 2 studies had scores greater than 20. External validity was particularly poor, with 15 of 20 studies scoring zero on this subscale. This deficiency is not surprising because most of the studies



were performed in referral centers with expertise in RFA. However, because summary estimates were comparable with the US RFA Registry<sup>34</sup> and with a multicenter community practice registry,<sup>13</sup> these results may be generalizable to the broader population. In addition, stratification by quality actually showed improved CE-IM and IM recurrence for the higher-quality studies. This review also was limited by substantial heterogeneity.  $I^2$  values for the primary efficacy and durability outcomes were moderate to high and are likely owing to differences in patient samples, study settings and protocols, and provider factors. This heterogeneity calls into question the reliability of the estimates, which will need to be confirmed as experience with RFA continues to accumulate in the future. However, the fact that estimates did not vary greatly with the inclusion and exclusion of the US RFA Registry data (which comprised more than half of the overall patients in this review), or with the removal of any other single study, speaks to the reliability of our findings despite the heterogeneity. Furthermore, although our analysis represents a detailed documentation of the data available to date, it can be no stronger than the underlying data, and the precision of our effect estimates may change as new data become available. Finally, the retrospective nature of several studies may have resulted in an underestimate of adverse events.

In summary, in this systematic review and meta-analysis of efficacy and durability studies of RFA for BE, we found that a high proportion of treated patients achieve CE-D and CE-IM and that most patients maintain the neosquamous epithelium after successful treatment without IM recurrence. We also observed a low incidence of adverse events with RFA. Despite the poor quality and limited external validity of the studies, these data may be helpful for patients and providers in weighing the absolute risks and benefits of this therapy. Further analysis and follow-up evaluation of the US RFA Registry in particular will be helpful in enhancing generalizability and defining RFA effectiveness in the broader population of patients with BE.

## Acknowledgments

### Funding

Supported in part by the National Institutes of Health (T32 DK07634 and UL1TR000083).

## Abbreviations used in this paper

<b>AIM</b>	Ablation of Intestinal Metaplasia
<b>BE</b>	Barrett's esophagus
<b>CE-D</b>	complete eradication of dysplasia
<b>CE-IM</b>	complete eradication of intestinal metaplasia
<b>CI</b>	confidence interval
<b>EAC</b>	esophageal adenocarcinoma
<b>EMR</b>	endoscopic mucosal resection
<b>HGD</b>	high-grade dysplasia
<b>IM</b>	intestinal metaplasia
<b>IMC</b>	intramucosal carcinoma
<b>LGD</b>	low-grade dysplasia
<b>NDBE</b>	nondysplastic Barrett's esophagus

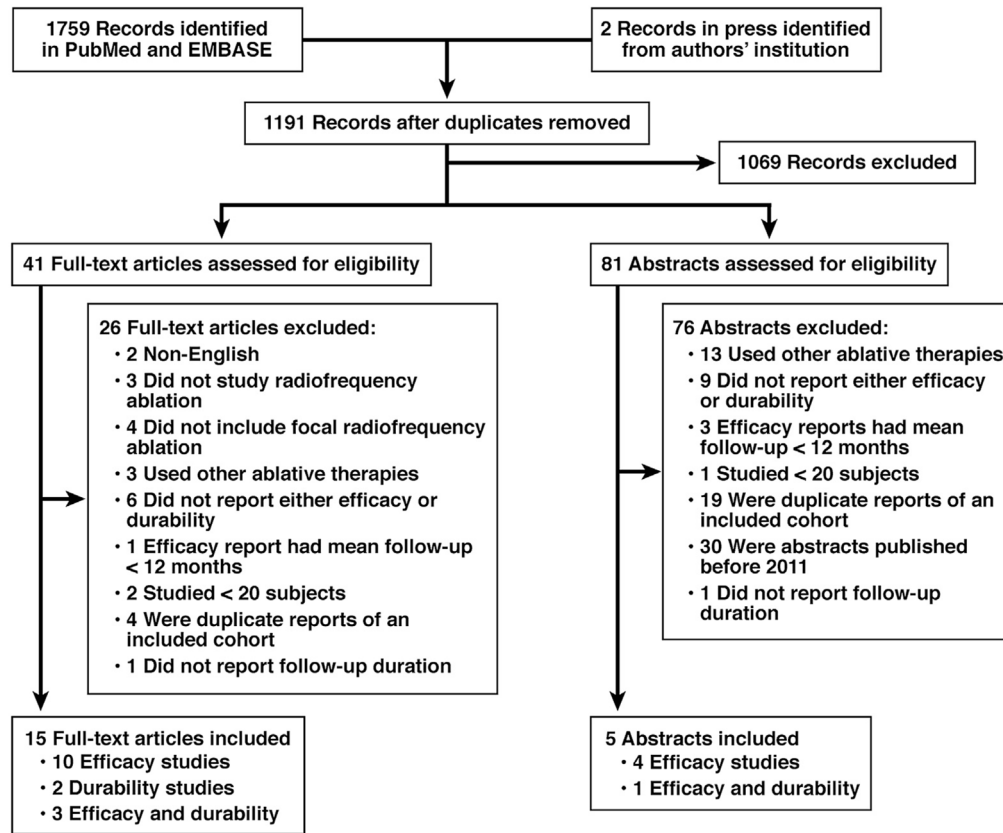
<b>RFA</b>	radiofrequency ablation
<b>RCT</b>	randomized controlled trial
<b>RR</b>	risk ratio

## References

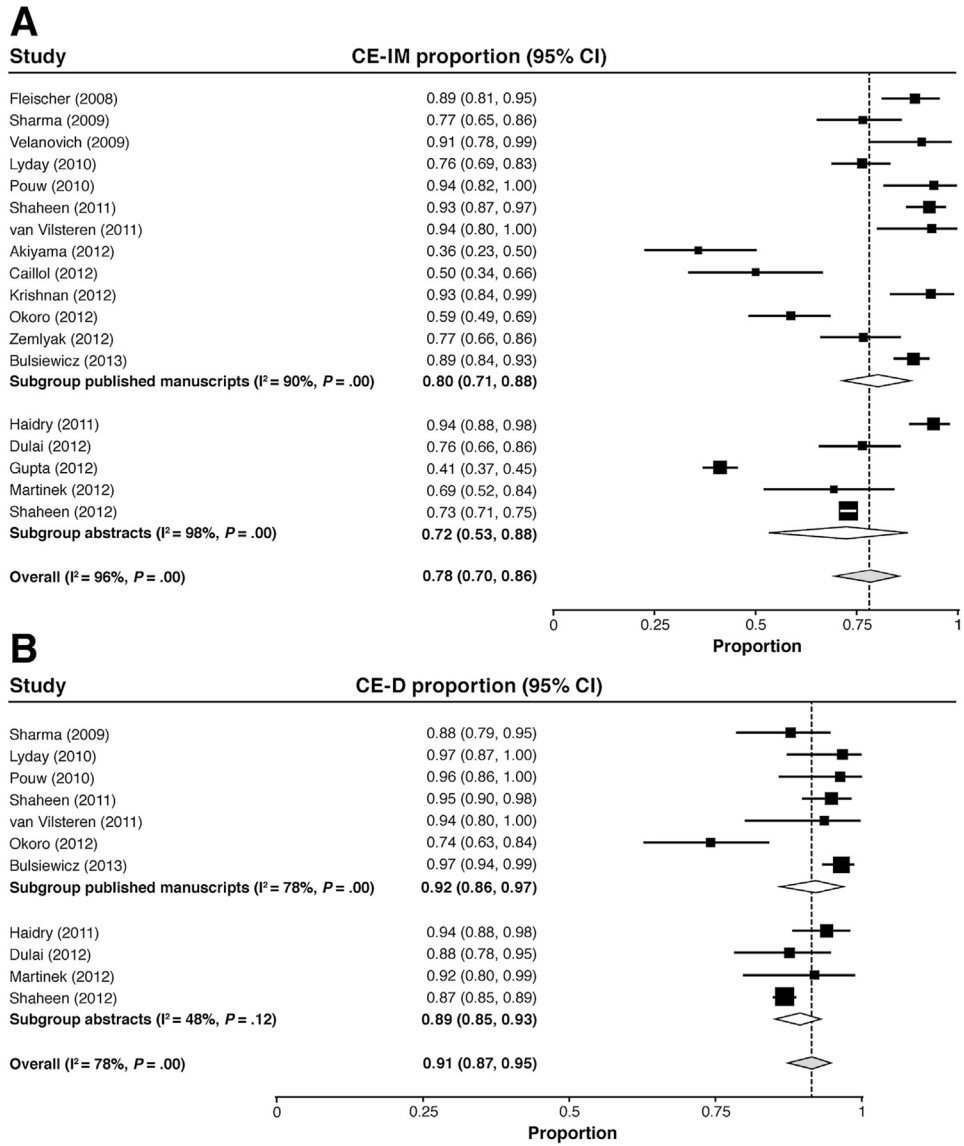
1. Spechler SJ, Sharma P, Souza RF, et al. American Gastroenterological Association medical position statement on the management of Barrett's esophagus. *Gastroenterology*. 2011; 140:1084–1091. [PubMed: 21376940]
2. Ronkainen J, Aro P, Storskrubb T, et al. Prevalence of Barrett's esophagus in the general population: an endoscopic study. *Gastroenterology*. 2005; 129:1825–1831. [PubMed: 16344051]
3. Hvid-Jensen F, Pedersen L, Drewes AM, et al. Incidence of adenocarcinoma among patients with Barrett's esophagus. *N Engl J Med*. 2011; 365:1375–1383. [PubMed: 21995385]
4. Sharma VK, Wang KK, Overholt BF, et al. Balloon-based, circumferential, endoscopic radiofrequency ablation of Barrett's esophagus: 1-year follow-up of 100 patients. *Gastrointest Endosc*. 2007; 65:185–195. [PubMed: 17258973]
5. Overholt BF, Panjehpour M, Haydek JM. Photodynamic therapy for Barrett's esophagus: follow-up in 100 patients. *Gastrointest Endosc*. 1999; 49:1–7. [PubMed: 9869715]
6. Shaheen NJ, Greenwald BD, Peery AF, et al. Safety and efficacy of endoscopic spray cryotherapy for Barrett's esophagus with high-grade dysplasia. *Gastrointest Endosc*. 2010; 71:680–685. [PubMed: 20363409]
7. Basu KK, Pick B, Bale R, et al. Efficacy and one year follow up of argon plasma coagulation therapy for ablation of Barrett's oesophagus: factors determining persistence and recurrence of Barrett's epithelium. *Gut*. 2002; 51:776–780. [PubMed: 12427775]
8. Shaheen NJ, Sharma P, Overholt BF, et al. Radiofrequency ablation in Barrett's esophagus with dysplasia. *N Engl J Med*. 2009; 360:2277–2288. [PubMed: 19474425]
9. Okoro NI, Tomizawa Y, Dunagan KT, et al. Safety of prior endoscopic mucosal resection in patients receiving radiofrequency ablation of Barrett's esophagus. *Clin Gastroenterol Hepatol*. 2012; 10:150–154. [PubMed: 22056303]
10. Pouw RE, Wirths K, Eisendrath P, et al. Efficacy of radiofrequency ablation combined with endoscopic resection for Barrett's esophagus with early neoplasia. *Clin Gastroenterol Hepatol*. 2010; 8:23–29. [PubMed: 19602454]
11. Fleischer DE, Overholt BF, Sharma VK, et al. Endoscopic radiofrequency ablation for Barrett's esophagus: 5-year outcomes from a prospective multicenter trial. *Endoscopy*. 2010; 42:781–789. [PubMed: 20857372]
12. Sharma VK, Jae Kim H, Das A, et al. Circumferential and focal ablation of Barrett's esophagus containing dysplasia. *Am J Gastroenterol*. 2009; 104:310–317. [PubMed: 19174783]
13. Lyday WD, Corbett FS, Kuperman DA, et al. Radiofrequency ablation of Barrett's esophagus: outcomes of 429 patients from a multi-center community practice registry. *Endoscopy*. 2010; 42:272–278. [PubMed: 20146164]
14. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. 2000; 283:2008–2012. [PubMed: 10789670]
15. Bulsiewicz WJ, Kim HP, Dellon ES, et al. Safety and efficacy of endoscopic mucosal therapy with radiofrequency ablation for patients with neoplastic Barrett's esophagus. *Clin Gastroenterol Hepatol*. 2013; 11:636–642. [PubMed: 23103824]
16. Orman ES, Kim HP, Bulsiewicz WJ, et al. Intestinal metaplasia recurs infrequently in patients successfully treated for Barrett's esophagus with radiofrequency ablation. *Am J Gastroenterol*. 2013; 108:187–195. [PubMed: 23247578]
17. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health*. 1998; 52:377–384. [PubMed: 9764259]

18. Wallace BC, Schmid CH, Lau J, et al. Meta-Analyst: software for meta-analysis of binary, continuous and diagnostic data. *BMC Med Res Methodol.* 2009; 9:80. [PubMed: 19961608]
19. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986; 7:177–188. [PubMed: 3802833]
20. Freeman MF, Tukey JF. Transformations related to the angular and the square root. *Ann Math Stat.* 1950; 21:607–611.
21. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002; 21:1539–1558. [PubMed: 12111919]
22. Fleischer DE, Overholt BF, Sharma VK, et al. Endoscopic ablation of Barrett's esophagus: a multicenter study with 2. 5-year follow-up. *Gastrointest Endosc.* 2008; 68:867–876. [PubMed: 18561930]
23. Velanovich V. Endoscopic endoluminal radiofrequency ablation of Barrett's esophagus: initial results and lessons learned. *Surg Endosc.* 2009; 23:2175–2180. [PubMed: 19263139]
24. Shaheen NJ, Overholt BF, Sampliner RE, et al. Durability of radio-frequency ablation in Barrett's esophagus with dysplasia. *Gastroenterology.* 2011; 141:460–468. [PubMed: 21679712]
25. van Vilsteren FG, Pouw RE, Seewald S, et al. Stepwise radical endoscopic resection versus radiofrequency ablation for Barrett's oesophagus with high-grade dysplasia or early cancer: a multicentre randomised trial. *Gut.* 2011; 60:765–773. [PubMed: 21209124]
26. Akiyama J, Marcus SN, Triadafilopoulos G. Effective intra-esophageal acid control is associated with improved radiofrequency ablation outcomes in Barrett's esophagus. *Dig Dis Sci.* 2012; 57:2625–2632. [PubMed: 22878916]
27. Caillol F, Bories E, Pesenti C, et al. Radiofrequency ablation associated to mucosal resection in the oesophagus: experience in a single centre. *Clin Res Hepatol Gastroenterol.* 2012; 36:371–377. [PubMed: 22361442]
28. Krishnan K, Pandolfino JE, Kahrilas PJ, et al. Increased risk for persistent intestinal metaplasia in patients with Barrett's esophagus and uncontrolled reflux exposure before radiofrequency ablation. *Gastroenterology.* 2012; 143:576–581. [PubMed: 22609385]
29. Zemlyak AY, Pacicco T, Mahmud EM, et al. Radiofrequency ablation offers a reliable surgical modality for the treatment of Barrett's esophagus with a minimal learning curve. *Am Surg.* 2012; 78:774–778. [PubMed: 22748537]
30. Haidry RJ, Dunn JM, Thorpe S, et al. Radio frequency ablation is more effective in shorter segments of Barrett's oesophagus for eradication of high grade dysplasia/intramucosal cancer - results from the UK RFA HALO registry. *Gastroenterology.* 2011; 140:S215.
31. Dulai PS, Fadden LL, Gordon SR, et al. Size really doesn't matter - a single center experience with radiofrequency ablation of ultra long-segment Barrett's esophagus. *Gastrointest Endosc.* 2012; 75:AB451.
32. Gupta M, Lutzke LS, Wang KK, et al. Safety and outcomes of radiofrequency ablation for Barrett's esophagus in older subjects: results from a betnet consortium. *Gastrointest Endosc.* 2012; 75:AB159.
33. Martinek J, Stefanova M, Suchanek S, et al. Endoscopic radiofrequency ablation combined with endoscopic resection for early neoplasia in Barrett's esophagus longer than 7 CM. *Gastrointest Endosc.* 2012; 75:AB459–AB460.
34. Shaheen NJ, Bulsiewicz WJ, Rothstein RI, et al. Eradication rates of Barrett's esophagus using radiofrequency ablation (RFA): results from the U.S. RFA registry. *Gastrointest Endosc.* 2012; 75:AB460.
35. Ganz RA, Utley DS, Stern RA, et al. Complete ablation of esophageal epithelium with a balloon-based bipolar electrode: a phased evaluation in the porcine and in the human esophagus. *Gastrointest Endosc.* 2004; 60:1002–1010. [PubMed: 15605025]
36. Sikkema M, Looman CW, Steyerberg EW, et al. Predictors for neoplastic progression in patients with Barrett's esophagus: a prospective cohort study. *Am J Gastroenterol.* 2011; 106:1231–1238. [PubMed: 21577245]
37. Sharma P, Falk GW, Weston AP, et al. Dysplasia and cancer in a large multicenter cohort of patients with Barrett's esophagus. *Clin Gastroenterol Hepatol.* 2006; 4:566–572. [PubMed: 16630761]

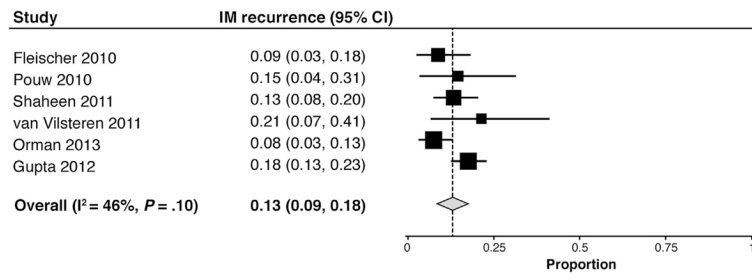
38. Peters FP, Brakenhoff KP, Curvers WL, et al. Histologic evaluation of resection specimens obtained at 293 endoscopic resections in Barrett's esophagus. *Gastrointest Endosc.* 2008; 67:604–609. [PubMed: 18155214]
39. Wani S, Puli SR, Shaheen NJ, et al. Esophageal adenocarcinoma in Barrett's esophagus after endoscopic ablative therapy: a meta-analysis and systematic review. *Am J Gastroenterol.* 2009; 104:502–513. [PubMed: 19174812]
40. Lewis JJ, Rubenstein JH, Singal AG, et al. Factors associated with esophageal stricture formation after endoscopic mucosal resection for neoplastic Barrett's esophagus. *Gastrointest Endosc.* 2011; 74:753–760. [PubMed: 21820109]
41. Overholt BF, Lightdale CJ, Wang KK, et al. Photodynamic therapy with porfimer sodium for ablation of high-grade dysplasia in Barrett's esophagus: international, partially blinded, randomized phase III trial. *Gastrointest Endosc.* 2005; 62:488–498. [PubMed: 16185958]



**Figure 1.**  
Study flow diagram.



**Figure 2.** Forest plots of the proportion of patients achieving (A) CE-IM and (B) CE-D after treatment of BE with RFA, stratified according to publication as a peer-reviewed article vs abstract.



**Figure 3.** Forest plot of the proportion of patients with recurrent IM after RFA-induced CE-IM.

Table 1

## Characteristics of Efficacy and Durability Studies

Study	Country	Design	Setting	Histology included	BE length criteria, cm	Sample size	RFA protocol	Follow-up duration, mo	Quality score
<i>Efficacy studies</i>									
<i>Published articles</i>									
Fleischer et al, <sup>22</sup> 2008	United States	Prospective cohort	Multicenter	NDBE	2–6	70	Circumferential at 0 and 4 mo, then up to 3 focal at least 2 mo apart	30	15
Sharma, <sup>12</sup> 2009	United States	Prospective cohort	Single-center	LGD, HGD	Any	63	Circumferential at time 0, then circumferential or focal every 3 mo	21	14
Velanovich, <sup>23</sup> 2009	United States	Retrospective cohort	Single-center	NDBE, LGD, HGD	Any	27	Circumferential at time 0, then circumferential or focal every 3 mo until visual eradication, with biopsy 12 mo later	12	7
Lyday, <sup>13</sup> 2010	United States	Retrospective cohort	Multicenter	NDBE, LGD, HGD	Any	137	Circumferential at time 0, then focal every 2–4 mo, alternating with biopsy sessions at some centers	20	13
Pouw, <sup>10</sup> 2010	The Netherlands, Germany, Belgium	Prospective cohort	Multicenter	HGD, IMC	<12	24	Circumferential at time 0, then circumferential or focal every 2 mo, maximum 2 circumferential and 3 focal ablations	22	13
Shaheen, <sup>24</sup> 2011	United States	Prospective cohort	Multicenter	LGD, HGD	8	119	Circumferential at time 0, then focal at 2, 4, 9, and 15 mo	24	23
van Vilsteren, <sup>25</sup> 2011	The Netherlands, Germany	RCT	Multicenter	HGD, IMC	5	22	Circumferential at time 0, then focal every 2–3 mo up to 4 total sessions	22	24
Akiyama, <sup>26</sup> 2012	United States	Retrospective cohort	Single-center	NDBE, LGD, HGD, IMC	Any	45	Circumferential or focal every 3 mo	>12	10
Caillol, <sup>27</sup> 2012	France	Retrospective cohort	Single-center	NDBE, LGD, HGD	Any	34	NR	12	12
Krishnan, <sup>28</sup> 2012	United States	Prospective cohort	Single-center	NDBE, LGD, HGD	Any	37	Circumferential at time 0, then focal every 2 mo up to 1 year	12	19
Okoro, <sup>9</sup> 2012	United States	Retrospective cohort	Single-center	NDBE, LGD, HGD, IMC	Any	90	Circumferential or focal at time 0, then focal every 10–12 wk, maximum 5 sessions	20.5 for RFA after EMR; 32 for RFA alone	18
Zemlyak, <sup>29</sup> 2012	United States	Retrospective cohort	Single-center	NDBE, LGD, HGD	Any	70	Circumferential or focal every 2 months	16	12
Bulsiewicz, <sup>15</sup> 2013	United States	Retrospective cohort	Single-center	LGD, HGD, IMC	Any	210	Circumferential at time 0, then circumferential or focal at least 2 mo apart	>12	18
<i>Abstracts</i>									
Haidry, <sup>30</sup> 2011	United Kingdom	Prospective cohort	Multicenter	HGD, IMC	Any	91	Every 3 mo, maximum 4 sessions	12	11
Dulai, <sup>31</sup> 2012	United States	Retrospective cohort	Single-center	NDBE, LGD, HGD, IMC	>3	69	NR	37 for BE >8 cm; 27 for BE <8 cm	13
Gupta, <sup>32</sup> 2012	United States	Prospective cohort	Multicenter	NDBE, LGD, HGD, IMC	Any	529	NR	21	11
Martinek, <sup>33</sup> 2012	Czech Republic	Prospective cohort	Single-center	LGD, HGD, IMC	Any	30	NR	12	14



Study	Country	Design	Setting	Histology included	BE length criteria, cm	Sample size	RFA protocol	Follow-up duration, mo	Quality score
Shaheen, <sup>34</sup> 2012	United States	Prospective cohort	Multicenter	NDBE, LGD, HGD, IMC	Any	2135	NR	17	9
<b>Durability studies</b>									
<b>Published articles</b>									
Fleischer, <sup>11</sup> 2010	United States	Prospective cohort	Multicenter	NDBE	2-6	50	60	51	15
Pouw, <sup>10</sup> 2010	The Netherlands, Germany, Belgium	Prospective cohort	Multicenter	HGD, IMC	<12	23	2, 6, 12, annual	22	13
Shaheen, <sup>24</sup> 2011	United States	Prospective cohort	Multicenter	LGD, HGD	8	110	LGD: 6, 12, annual; HGD: 3, 6, 9, 12, annual	13	23
van Vilsteren, <sup>25</sup> 2011	The Netherlands, Germany	RCT	Multicenter	HGD, IMC	5	20	6, 12, annual	15	24
Orman, <sup>16</sup> 2013	United States	Retrospective cohort	Single-center	LGD, HGD, IMC	Any	119	LGD: 6, 12, annual; HGD before 2009: 3, 6, 9, 12, 18, 24, annual; HGD after 2009: 6, 12, annual	13	16
<b>Abstracts</b>									
Gupta, <sup>32</sup> 2012	United States	Prospective cohort	Multicenter	NDBE, LGD, HGD, IMC	Any	218	NR	18	11

**Table 2**

**Patient and Treatment Characteristics in Efficacy and Durability Studies**

Study	Histology, %NDBE/LGD/HGD/IMC	Age, mean or median	Sex, % male	BE length, mean or median	RFA sessions, mean or median	Circumferential RFA, % of patients	EMR, % of patients
<b>Efficacy studies</b>							
Published articles							
Fleischer, <sup>22</sup> 2008	100/0/0/0	56	74	3.2	3.4	100	0
Sharma, <sup>12</sup> 2009	0/62/38/0	71	90	5	2	100	NR
Velanovich, <sup>23</sup> 2009	NR/NR/18/0	62	80	3	1	100	NR
Lyday, <sup>13</sup> 2010	80/10/7/0	60	66	3	2.1	NR	1.6
Pouw, <sup>10</sup> 2010,	0/0/33/67	65	83	8	2	100	96
Shaheen, <sup>24</sup> 2011	0/49/51/0	66	86	4.9	NR	100	8.4
van Vilsteren, <sup>25</sup> 2011	0/0/32/68	69	86	4	3	100	82
Akiyama, <sup>26</sup> 2012	58/11/20/11	62	73	4.6	1.7	80	18
Caillol, <sup>27</sup> 2012	9/50/41/0	60	82	6.8	1.3	NR	47
Krishnan, <sup>28</sup> 2012	30/46/24/0	64	NR	5	2	100	NR
Okoro, <sup>9</sup> 2012	30/18/48/4	63	91	5.7	2.2	79	49
Zemlyak, <sup>29</sup> 2012	26/63/11/0	61	80	NR	2.8	NR	NR
Bulsiewicz, <sup>15</sup> 2013	0/21/64/15	64	78	4	2.8	60	33
<b>Abstracts</b>							
Haidry, <sup>30</sup> 2011	NR	70	75	6	2.6	NR	42
Dulai, <sup>31</sup> 2012	13/17/49/20	67	80	7.2	2.8	NR	NR
Gupta, <sup>32</sup> 2012	22/NR/33/NR	65	85	4.3	2.3	54	50
Martinek, <sup>33</sup> 2012	0/27/30/43	61	NR	5.8	2.2	NR	80
Shaheen, <sup>34</sup> 2012	45/22/22/4	NR	NR	NR	2.8	NR	NR

**Durability studies**

Study	Histology, % NDBE/LGD/HGD/IMC	Age, mean or median	Sex, % male	BE length, mean or median	Surveillance sessions, mean or median
<b>Durability studies</b>					
Published articles					
Fleischer, <sup>11</sup> 2010	100/0/0/0	54	74	3.1	1
Pouw, <sup>10</sup> 2010	0/0/33/67	65	83	8	3

Study	Histology, % NDBE/LGD/HGD/IMC	Age, mean or median	Sex, % male	BE length, mean or median	Surveillance sessions, mean or median
Shaheen, <sup>24</sup> 2011	0/49/51/0	66	86	4.9	NR
van Vilsteren, <sup>25</sup> 2011	0/0/32/68	69	86	4	3
Orman, <sup>16</sup> 2013	0/21/63/15	64	80	4	2
Abstracts					
Gupta, <sup>32</sup> 2012	22/NR/33/NR	65	85	4.3	NR

**Table 3**

Eradication According to Pretreatment Histology

	Total studies	Eradication, n/N	Studies in model	RR	95% CI	I <sup>2</sup> , %
CE-IM	4	941/1161	3	1.12	0.90–1.40	61
	6	458/633	–	–	–	–
LGD	6	458/633	–	–	–	–
HGD	7	490/721	6	0.92	0.87–0.98	0
IMC	3	89/121	3	0.99	0.82–1.18	43
CE-D	6	581/633	–	–	–	–
	6	607/714	6	0.94	0.91–0.97	8
LGD	6	607/714	6	0.94	0.91–0.97	8
HGD	3	104/121	3	0.94	0.87–1.01	0
IMC	3	104/121	3	0.94	0.87–1.01	0