Per Oral Endoscopic Myotomy for Esophageal Motility Disorders: A Review of Clinical, Cost-Effectiveness, and Guidelines
Authors: Kwakye Peprah, Lorna Adcock


Acknowledgments:

ISSN: 1922-8147 (online)

Disclaimer: The information in this document is intended to help Canadian health care decision-makers, health care professionals, health systems leaders, and policy-makers make well-informed decisions and thereby improve the quality of health care services. While patients and others may access this document, the document is made available for informational purposes only and no representations or warranties are made with respect to its fitness for any particular purpose. The information in this document should not be used as a substitute for professional medical advice or as a substitute for the application of clinical judgment in respect of the care of a particular patient or other professional judgment in any decision-making process. The Canadian Agency for Drugs and Technologies in Health (CADTH) does not endorse any information, drugs, therapies, treatments, products, processes, or services.

While care has been taken to ensure that the information prepared by CADTH in this document is accurate, complete, and up-to-date as at the applicable date the material was first published by CADTH, CADTH does not make any guarantees to that effect. CADTH does not guarantee and is not responsible for the quality, currency, propriety, accuracy, or reasonableness of any statements, information, or conclusions contained in any third-party materials used in preparing this document. The views and opinions of third parties published in this document do not necessarily state or reflect those of CADTH.

CADTH is not responsible for any errors, omissions, injury, loss, or damage arising from or relating to the use (or misuse) of any information, statements, or conclusions contained in or implied by the contents of this document or any of the source materials.

This document may contain links to third-party websites. CADTH does not have control over the content of such sites. Use of third-party sites is governed by the third-party website owners’ own terms and conditions set out for such sites. CADTH does not make any guarantee with respect to any information contained on such third-party sites and CADTH is not responsible for any injury, loss, or damage suffered as a result of using such third-party sites. CADTH has no responsibility for the collection, use, and disclosure of personal information by third-party sites.

Subject to the aforementioned limitations, the views expressed herein are those of CADTH and do not necessarily represent the views of Canada’s federal, provincial, or territorial governments or any third party supplier of information.

This document is prepared and intended for use in the context of the Canadian health care system. The use of this document outside of Canada is done so at the user’s own risk.

This disclaimer and any questions or matters of any nature arising from or relating to the content or use (or misuse) of this document will be governed by and interpreted in accordance with the laws of the Province of Ontario and the laws of Canada applicable therein, and all proceedings shall be subject to the exclusive jurisdiction of the courts of the Province of Ontario, Canada.

The copyright and other intellectual property rights in this document are owned by CADTH and its licensors. These rights are protected by the Canadian Copyright Act and other national and international laws and agreements. Users are permitted to make copies of this document for non-commercial purposes only, provided it is not modified when reproduced and appropriate credit is given to CADTH and its licensors.

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada’s health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.
Context and Policy Issues

Achalasia is an esophageal motility disorder characterized by the inability of the lower esophageal sphincter (LES) to relax. The continued contraction of the LES prevents the typical wave of esophageal contraction (peristalsis) which is required for swallowing. The resulting retention of food and saliva leads to the common presenting symptoms of achalasia such as dysphagia, regurgitation, chest pain and weight loss. The incidence and prevalence of achalasia in Canada are reported to be 1.63 per 100,000 per year and 10.82 per 100,000, respectively. It has been proposed that achalasia is a neurodegenerative disease caused by factors such as viral infections, inflammatory condition, and autoimmune processes targeting the esophageal ganglion cells. The primary goal of therapy is to lower the LES pressure and restore the ability to swallow. Treatments that have been used over the years include pharmacotherapy (calcium channel antagonists, nitrates), endoscopic pneumatic dilatation, surgical myotomies such as laparoscopic Heller myotomy (LHM) and injection of botulinum toxin.

Per Oral Endoscopic Myotomy (POEM) is a relatively new procedure, with the first human studies published in 2010. Although POEM is primarily used to treat achalasia, it has also been used to treat other esophageal motility disorders such as nutcracker esophagus, hypertensive peristalsis, and diffuse esophageal spasm.

This review aims to summarize evidence regarding the clinical and cost-effectiveness, as well as guidelines for the use of POEM in the treatment of achalasia and other esophageal motility disorders.

Research Question

1. What is the clinical effectiveness of per oral endoscopic myotomy for the treatment of achalasia?
2. What is the clinical effectiveness of per oral endoscopic myotomy for the treatment of other esophageal motility disorders?
3. What is the cost-effectiveness of per oral endoscopic myotomy for the treatment of esophageal motility disorders?
4. What are the evidence-based guidelines regarding the use of per oral endoscopic myotomy for the treatment of other esophageal motility disorders?

Key Findings

Findings from the included systematic reviews (SRs) indicate that POEM is an effective procedure to treat achalasia and spastic esophageal disorders (SED) as demonstrated by significant reductions in Eckardt score, LES pressure, timed barium esophagram (TBE) column height, and weight gain. The results from the SRs suggest that POEM is a safe procedure with the main adverse events (AEs) being mild gastroesophageal reflux (GER)-related conditions which were successfully controlled with proton pump inhibitors (PPI). No POEM-related deaths were reported. The findings from
two SRs indicate that POEM and LHM had similar effectiveness at resolving achalasia symptoms. However, results from one SR suggests that the proportion of patients who experience improvements in dysphagia may be higher with POEM than with LHM with fundoplication. Findings from three SRs indicate that post-operative pain score and analgesic dose were similar with POEM and LHM. On safety, results from two SR suggest that the incidence of GER-related AEs was higher with POEM than with LHM with fundoplication, whereas findings from two SRs indicate that the rate of postoperative symptomatic GER was similar between POEM and LHM. The results from four SRs indicate that procedural complications were similar between POEM and LHM. The included SRs in this review lacked long-term follow-up POEM data, and relevant randomized controlled trial (RCT). Therefore, a rigorous RCT comparing POEM with LHM over a long-term is needed to assess the comparative advantage of one procedure over the other.

The literature search did not identify cost-effectiveness studies or evidence-based clinical guidelines on the use of POEM for achalasia or other esophageal mobility disorders.

**Methods**

**Literature Search Methods**

A limited literature search was conducted on key resources including PubMed, The Cochrane Library, the University of York Centre for Reviews and Dissemination (CRD) databases and a focused Internet search. Methodological filters were applied to limit retrieval to health technology assessments, systematic reviews, meta-analyses, randomized controlled trials, non-randomized studies, economic studies, and guidelines. The search was limited to English language documents published between January 1, 2012, and December 7, 2017.

Rapid Response reports are organized so that the evidence for each research question is presented separately.

**Selection Criteria and Methods**

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed, and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text publications was based on the inclusion criteria presented in Table 1.

<table>
<thead>
<tr>
<th>Population</th>
<th>Q1: Patients with achalasia &lt;br&gt;Q2: Patients with other esophageal motility disorders (e.g., nutcracker esophagus, or hypertensive peristalsis, and diffuse esophageal spasm) &lt;br&gt;Q3-4: Patients with esophageal motility disorders (e.g., achalasia, nutcracker esophagus, or hypertensive peristalsis, and diffuse esophageal spasm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>Per Oral Endoscopic Myotomy (POEM)</td>
</tr>
<tr>
<td>Comparator</td>
<td>Q1-2: Endoscopic pneumatic dilation or laparoscopic surgical myotomy, no treatment &lt;br&gt;Q3: Endoscopic pneumatic dilation or laparoscopic surgical myotomy &lt;br&gt;Q4: No comparator</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Q1-2: Clinical effectiveness (benefit/harm), safety, timed barium swallow, long-term follow-up, hospital length of stay, readmission rate, reflux, peri-treatment complications (e.g., esophageal perforation), long-</td>
</tr>
</tbody>
</table>
Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, or were published before 2012. Due to a large number of studies identified by the literature search, study selection for this review was limited to systematic reviews (SRs) with or without meta-analysis. Where there was complete overlap of included primary studies with a selected SR, the SR with the larger number of primary studies or latest data was selected.

Critical Appraisal of Individual Studies

The included SRs\(^1,2,5-12\) were critically appraised using the AMSTAR tool.\(^13\) Summary scores were not calculated for the included SRs; instead, a review of the strengths and limitations of each included study was described narratively.

Summary of Evidence

Quantity of Research Available

A total of 536 citations were identified in the literature search. Following the screening of titles and abstracts, over 100 potentially relevant studies were identified. Given the time and resource constraints, it was decided to limit the study selection for this review to SRs, with or without meta-analysis. Therefore, 528 citations were excluded, and 11 potentially relevant SRs from the electronic search were retrieved for full-text review. The grey literature search did not identify additional potentially relevant SRs. Of the 11 potentially relevant articles, one SR\(^14\) was excluded because a larger SR\(^5\) which had already been select included all of its primary studies and reported similar results. Thus, ten SRs\(^1,2,5-12\) met the inclusion criteria and were included in this report. Appendix 1 describes the PRISMA flowchart of the study selection.

Additional references of potential interest are provided in Appendix 5.

Summary of Study Characteristics

Study Design

A total of ten systematic reviews (SRs)\(^1,2,5-12\) including eight with meta-analysis\(^5-12\) were identified which evaluated the clinical effectiveness of POEM or LHM for the treatment of achalasia or other esophageal motility disorders. Four SRs\(^1,10-12\) reported only non-comparative POEM outcomes while six SRs\(^2,5-9\) reported POEM versus LHM outcomes. POEM versus LHM outcomes were generated either from primary studies comparing the two procedures, or by separately pooling the outcomes of POEM only and LHM only studies and comparing the overall results. The primary studies which assessed POEM were published between 2012 and 2016 whereas studies which investigated LHM were published between 1997 and 2017. One RCT comparing two techniques for POEM was common to two SRs.\(^5,8\) The SR by Schlottmann et al.,\(^8\) had four additional RCTs on LHM.
The remaining eight SRs were of non-randomized studies only. Further information about the characteristics of included SRs is available in Appendix 2.

Country of Origin

Six SRs were conducted by authors located in single countries—one each from Canada¹ and the United Kingdom (UK),² and two each from the USA³-⁶ and Italy.⁵,¹⁰ Each of the remaining four SRs had authors from two or more countries including UK, USA, Italy, Germany, India, Japan, Malaysia, and Pakistan, as detailed in Appendix 2.

Patient Population

The total number of patients involved in the individual SRs ranged from 179¹¹ to 7,797.⁸ Primary studies of all the SRs were conducted in patients diagnosed with achalasia. One SR¹¹ included patients with spastic esophageal disorders (SEDs) such as type III achalasia, jackhammer esophagus, diffuse esophageal spasm (DES), and hypertensive non-relaxing LES. Patients’ demographic information and disease history were not uniformly reported. For instance, the overall sex distribution across the primary studies was reported by only two SRs,¹⁰,¹² and only Talukdar et al.⁵ reported the mean duration of disease. The most commonly reported demographic parameter was patients’ age. The SR by Crespin et al.¹ included one primary study in pediatric patients with a mean age of 13.8 years while the remaining eighteen were conducted in adult patients (> 18 years). Eight SRs²⁻⁵,⁹⁻¹² included only primary studies conducted in adults patients. The SR by Khan et al.¹¹ did not provide any patients’ demographic information. Eight SRs¹⁻²,⁵⁻⁹,¹¹ reported that their primary studies included treatment-naive patients as well as those who failed previous treatment or were experiencing a relapse. However, the reporting was inconsistent making it difficult to arrive at representative data for these patient subgroups.

Interventions and Comparators

POEM was the intervention of interest in all the included SRs.¹⁻²,⁵,⁶,⁸⁻¹² Six SRs¹⁻²,⁵,⁶,⁸⁻¹² compared POEM to LHM, including two SRs⁸,¹⁵ which excluded LHM studies that did not specify accompanying fundoplication. The SRs by Talukdar et al.⁵ and Patel et al.⁵ performed the comparison between POEM and LHM as sub-analyses. The remaining four SRs¹⁻⁵,¹⁰⁻¹² had POEM as the sole intervention without a comparator. Nine SRs²⁻⁵,¹² reported that the mean duration of follow-up after POEM in the majority of their included studies was less than 12 months, although there was a few outliers with longer periods of follow-up.

Outcomes

Efficacy measure

The most common efficacy outcome measure was Eckardt scores, reported in eight of the ten included SRs.¹⁻²,⁵⁻⁷,⁹⁻¹¹,¹² The Eckardt score is the sum of the scores for dysphagia, regurgitation, and chest pain on a scale from 0 to 3 (0 = absent, 1 = occasional, 2 = daily, 3 = each meal) and weight loss (0 = no weight loss, 1 for < 5 kg, 2 for 5–10 kg, 3 for >10 kg). The total score ranges from 0 to 12 points.¹ Eckardt scores of ≤3 post-treatment is widely accepted as clinical success.¹,²,⁵⁻⁶,⁹⁻¹² Improvement in dysphagia was the efficacy outcome in the SRs by Schlotthmann⁶ and Barbieri¹⁰ but they did not report how it was assessed. Other efficacy measures were LES pressure,¹⁻²,⁵,¹² postoperative pain score,⁴⁻⁹ analgesic dose,⁵,⁶,¹² length of hospital stay,⁵,⁶ procedure time,⁵,⁶ TBE column height,¹² and average weight gain.¹² Also, Barberie et al.¹⁰ assessed technical feasibility and clinical success rate. Technical feasibility was defined as the pooled prevalence of successfully-performed procedure, and clinical success rate was defined as the pooled rate of clinically-relevant
improvement of dysphagia. Khan et al.\textsuperscript{11} reported the cumulative clinical success in spastic esophageal disease (SED) and clinical success in subtypes of SEDs. Clinical success was defined as improvement in severity of dysphagia based on achalasia disease-specific health-related quality of life questionnaire.

Safety measures

Safety outcomes were postoperative GER symptoms and gastroesophageal reflux disease (GERD),\textsuperscript{1,2,5-12} reflux esophagitis and abnormal acid exposure,\textsuperscript{1,2,7,11,12} and operative (i.e. procedure-related) complications.\textsuperscript{1,2,5,6,9,11,12}

Summary of Critical Appraisal

All the SRs\textsuperscript{1,2,5-12} included in this review stated clear objectives, defined their inclusion and exclusion criteria and described the interventions of interest as well as their main study findings. Nine SRs\textsuperscript{1,2,5-12} conducted comprehensive literature searches for relevant studies using multiple electronic databases and supplementing with hand-searches of the bibliographies of all selected articles. However, Schlottmann et al.\textsuperscript{8} searched only one electronic database for potentially relevant studies. Thus, it is unclear whether their literature search was sufficiently comprehensive. In seven SRs,\textsuperscript{1,6,7,10-12} multiple reviewers independently screened citations and abstracts and reviewed full-text article for inclusion to minimize bias. However, three SRs\textsuperscript{2,8,9} did not describe their screening and study selection procedure adequately. The primary studies of all the SRs were non-randomized except one RCT which compared two different techniques for performing POEM. Thus, the primary studies of the included SRs had a high potential for selection and performance biases that could limit generalizability. Methodological quality appraisals performed by seven SRs\textsuperscript{2,5,6,9} rated the quality of their primary studies between fair and good. Two SRs\textsuperscript{8,12} did not evaluate the quality of their primary studies, while one SR\textsuperscript{1} did not report the findings of their assessment. All the SRs\textsuperscript{1,2,5-12} reported high heterogeneity across primary studies with eight reporting significant or indeterminate publication biases in some analysis. The analysis in one SR\textsuperscript{10} did not detect publication bias, and another SR\textsuperscript{8} did not assess for publication bias. The primary studies in the included SRs were performed in many countries across the globe including Canada. Thus, it is reasonable to expect broad generalizability of the findings. However, Akintoye et al\textsuperscript{12} reported that the heterogeneity in their SR was significantly explained by the distribution of female patients across the studies, with the continent of origin of the study contributing marginally. Therefore, further studies may be required to confirm the generalizability. None of the authors of the included SRs had competing interest with the potential to influence the reported outcomes.

Summary of Findings

1. What is the clinical effectiveness of per oral endoscopic myotomy for the treatment of achalasia?

Non-comparative SRs on POEM

Five SRs,\textsuperscript{1,2,10,12} evaluated the effectiveness of POEM for the treatment of achalasia without a comparator intervention. Further details are available in Appendix 4.
Success rate of Procedure

Two SRs\textsuperscript{1,10} found that POEM was successfully performed in 97% to 99% of patients with achalasia. In a meta-analysis of 16 studies, Barberie et al.\textsuperscript{10} found 97% (95% CI: 94, 98) successfully-performed procedures following a meta-analysis of 16 studies, Crespin et al.\textsuperscript{1} did not perform a meta-analysis but calculated the proportion of successfully completed POEM across 19 studies to be 99% using study-level data.

Improvement in dysphagia or Eckardt score

Two SRs\textsuperscript{1,2,10} reported that 94.5% and 98% of achalasia patients who underwent POEM achieved Eckardt scores ≤3. From a meta-analysis of 27 studies, Akintoye et al.\textsuperscript{12} found that 98% (95% CI: 97, 100) of patients achieved a post-POEM Eckardt scores ≤3, with the mean (standard deviation [SD]) decreasing from 6.9 (0.15) before POEM to 0.77 (0.10) within one month. The reduced Eckardt score was maintained at 1.0 (0.10) and 1.0 (0.08) within six and 12 months, respectively, after the treatment. From study-level data, Crespin et al.\textsuperscript{1} reported that an average of 94.5% of patients who successfully underwent POEM achieved Eckardt scores ≤3. Talukdar et al.\textsuperscript{5} and Patel et al.\textsuperscript{2} reported reductions in Eckardt score after POEM differently. Although both stated in the text that Eckardt scores ≤3 (or <4) is a measure of treatment success, they reported the magnitude of reductions from baseline without indicating the proportion of patients who achieved Eckardt scores ≤3. From a meta-analysis of 19 studies, Talukdar et al.\textsuperscript{5} found a significant decrease in Eckardt score with an overall effect size of -7.95 (p <0.0001). Patel et al.\textsuperscript{2} reported a significant improvement in achalasia symptoms demonstrated (in part) by a reduction in the mean (SD) Eckardt score of 6.8 (1.0) before POEM to 1.2 (0.6) after POEM in 66% of patients. A meta-analysis of 16 studies by Barbieri et al.\textsuperscript{10} found a clinical success rate of 93% (95% CIs: 90, 95%). They defined clinical success as clinically significant improvement in dysphagia, without specifying how that was determined.

Lower Esophageal sphincter (LES) pressure and integrated relaxation pressure (IRP)

Four SRs\textsuperscript{1,2,5,12} reported significant decreases in LES pressure in achalasia patients after they underwent POEM. Akintoye et al.\textsuperscript{12} reported a 57.6% decrease in LES pressures after POEM from a meta-analysis of 24 studies while Patel et al.\textsuperscript{2} found a 65% decrease from 22 studies. In a meta-analysis of 16 studies, Talukdar et al.\textsuperscript{5} showed that patients who underwent POEM had a significant improvement of LES pressure with an overall effect size of -7.28 (p<0.0001). Crespin et al.\textsuperscript{1} found that the mean (SD) LES pressure across ten studies significantly decreased from a pre-POEM range of 22.0 (10.0) to 71.2 (20.8) mmHg to a post-POEM range of 6.49 (1.42) to 29.8 (22.0) mmHg. Only Akintoye et al.\textsuperscript{12} evaluated IRP outcomes, reporting a 56.7% reduction within six months after POEM.

Timed barium esophagogram (TBE) column height

Two SRs\textsuperscript{2,12} reported decreases in TBE column height ranging from 70% to 79.7% in achalasia patients after they had undergone POEM. Akintoye et al.\textsuperscript{12} showed that the mean (SD) TBE column height at one minute decreased by 70% (i.e., from 14.0 (2.3) cm to 4.2 (0.77) cm) at one minute. Similarly, the TBE column height at five minutes reduced by 73% compared with pre-POEM (i.e., from 9.7 (1.9) cm to 2.6 (0.72) cm at five minutes). Patel et al.\textsuperscript{2} reported a 79.7% reduction in TBE column height compared to pre-POEM heights.

Weight gain

Only Akintoye et al.\textsuperscript{12} evaluated weight gain outcomes, reporting that the mean (SD) weight gain was 5.4 (0.73) kg after a mean follow-up of 7.4 months in 488 patients in six studies.
Clinical failure rate

Patel et al.\textsuperscript{2} defined treatment failure as an Eckardt score \( \geq 4 \) and showed that treatment failure of POEM occurred at rates of between 3.2\% and 8.0\% at various time points during a mean follow-up period of between six and 30 months.

Safety

Safety outcomes were group broadly as GER-related AEs and complications directly related to the POEM procedure. Akintoye et al.\textsuperscript{12} reported a mean symptomatic GER of 8.5\% (95 \%CI: 4.9, 13) within eight months follow-up period. Five SRs\textsuperscript{1,2,5,10,12} found that the rates of reflux esophagitis ranged from 10.9\% to 19.0\% after the POEM procedure. The reported abnormal esophageal acid exposure from two SRs\textsuperscript{1,12} ranged from 20\% to 53\%, and one SR\textsuperscript{1} reported a study-level abnormal acid exposure of 47\% (95 \%CI: 21, 74). The cases were mild and were mostly controlled symptomatically with PPIs. The most frequently reported procedural complications were subcutaneous emphysema (7.5\% to 31\%),\textsuperscript{1,2,5,12} pneumoperitoneum (6.8\% to 30.0\%),\textsuperscript{1,2,5} pleural effusion (10.2\%),\textsuperscript{1} pneumothorax (5.3\% to 11.0\%),\textsuperscript{1,2,5} mucosal perforation (4.8\% to 9.0\%),\textsuperscript{1,12} and pneumonia (7.9\% to 9.19\%).\textsuperscript{1,5} There were no reports of POEM-related deaths.

SRs comparing POEM versus LHM

Improvement in dysphagia or Eckardt scores

Schlottmann et al.\textsuperscript{8} found that an average of 93.2\% of patients who underwent POEM reported improvement in dysphagia compared with 87.7\% of patients who underwent LHM with fundoplication. However, Talukdar et al.\textsuperscript{5} and Marano et al.\textsuperscript{6} did not find a statistically significant difference in the reduction of postoperative Eckhart’s score between POEM and LHM. It should be noted that Talukdar et al.\textsuperscript{5} and Marano et al.\textsuperscript{6} did not specify the use of fundoplication, which is a surgical procedure to reinforce the LES to make it less likely that acid will back up in the esophagus.\textsuperscript{16}

Postoperative pain score and analgesic dose

Three SRs\textsuperscript{5,6,9} evaluated postoperative pain score and found no statistically significant difference between POEM and LHM in the reduction of post-operative pain score. Talukdar et al.\textsuperscript{5} and Marano et al.\textsuperscript{6} also assessed the dose (in morphine equivalent) of analgesic medication used by patients who had undergone POEM or LHM and found no statistically significant difference between the two groups.

Procedure time

Five SRs\textsuperscript{2,5,6,8,9} assessed the length of time it takes to perform POEM or LHM. While four SRs\textsuperscript{2,5,6,9} found no significant difference in operative time between the two procedures, Talukdar et al.\textsuperscript{5} reported that the time for the POEM procedure was significantly less than for LHM (\( Z = -2.220; p = 0.026 \)). It should be noted that this was a sub-analysis of five studies (out of a total of 74 primary studies) which compared POEM to LHM.\textsuperscript{5}

Length of hospital stay

Five SRs\textsuperscript{2,5,6,8,9} evaluated the length of hospital stay after POEM and LHM and reported inconsistent findings. Three SRs\textsuperscript{5,6,9} found no statistically significant difference in the duration of hospital stay between POEM and LHM after the operation. However, Schlottmann et al.\textsuperscript{8} reported a longer hospital stay for POEM compared with LHM with
fundoplication whereas Marano et al.\textsuperscript{6} found a shorter hospital for POEM than for LHM after the operation. However, the difference in the duration of hospital stay was marginally statistically significant in each of these studies\textsuperscript{6,8} (1.03 days $P = 0.04$,\textsuperscript{8} and 0.629 days; $P = 0.049$).\textsuperscript{6}

**Clinical failure**

Only Awaiz et al.\textsuperscript{9} assessed clinical failure outcomes, reporting a significantly higher short-term clinical treatment failure rate for LHM than POEM (OR = 9.82; 95% CI, 2.06, 46.80; $P < 0.01$). The criteria for treatment failure were not defined clearly.

**Safety and Complications**

Repici et al.\textsuperscript{7} Schlottmann et al.\textsuperscript{8} found that the overall post-operative pooled rate of GER symptoms was higher with POEM than with LHM with fundoplication. However, Talukdar et al.\textsuperscript{5} and Awaiz et al.\textsuperscript{9} found that the incidence of postoperative symptomatic GER was similar between POEM and LHM. Repici et al.\textsuperscript{7} and Schlottmann et al.\textsuperscript{8} also reported that the rates of abnormal acid exposure and esophagitis were significantly higher ($P < 0.0001$) in patients who underwent POEM than those who underwent LHM with fundoplication.

It is worth noting that Repici et al.\textsuperscript{7} Schlottmann et al.\textsuperscript{8} analyzed non-comparative studies for POEM separately from non-comparative studies on LHM with fundoplication and compared the outcomes of the two procedures, whereas Talukdar et al.\textsuperscript{5} and Awaiz et al.\textsuperscript{9} examined comparative studies between POEM and LHM with no criteria to exclude studies without fundoplication. Considering that fundoplication is intended to reduce acid reflux into the esophagus,\textsuperscript{9,16} its presence may explain the lower rates of GER symptoms and abnormal acid exposure associated with LHM with fundoplication compared with POEM.

Four SRs\textsuperscript{2,5,6,9} found that the risk of procedure-related complications was similar between POEM and LHM.

2. **What is the clinical effectiveness of per oral endoscopic myotomy for the treatment of other esophageal motility disorders?**

Only the SR by Khan et al.\textsuperscript{11} assessed the effectiveness of POEM in spastic esophageal disorders (SEDs). They found that overall, 87% (95% CI: 78, 93%) of SED patients who underwent POEM achieved clinical success. Clinical success was defined as Eckardt scores of $\leq 3$, or improvement in severity of dysphagia based on achalasia disease-specific health-related quality of life questionnaire, or both. For the individual subtypes of SED, the weighted pooled rate (WPR) of clinical success after POEM were 92% for type III achalasia, 88% for DES, and 72% for jackhammer achalasia.

3. **What is the cost-effectiveness of per oral endoscopic myotomy for the treatment of esophageal motility disorders?**

The literature search did not identify any cost-effectiveness studies on the use of POEM for the treatment of esophageal motility disorders.

4. **What are the evidence-based guidelines regarding the use of per oral endoscopic myotomy for the treatment of other esophageal motility disorders?**

The literature search did not identify any evidence-based guidelines on POEM for the treatment of achalasia or any other esophageal motility disorders.
Limitations

All the primary studies on POEM of the SRs included in this reviews were non-randomized except one RCT which compared two different techniques for POEM. Therefore, the SRs were based predominantly on studies with high potential for selection and performance biases. Also, the literature search did not identify any RCT comparing POEM with other interventions for achalasia or other esophageal motility disorders. Many of the primary studies across the included SRs were from the same authors or research groups possibly reporting outcomes from the same study population. Such substantial overlap of primary studies presented a risk of inflated beneficial effect. Many of the SRs did not report the duration of the disease or the treatment history before the patients underwent POEM. For SRs which indicated that their primary studies included both treatment-naive and treatment-exposed patients, analyses of outcomes were not stratified on treatment history. Thus, it is unknown whether the reported findings could differ in other patient populations based on previous treatment. All the SRs stated that the follow-up period in most of their primary studies on POEM was short (≤12 months). Therefore, it is unknown if the reported findings will be sustainable over the long-term. Further, because the primary studies on LHM had data over a longer term while studies on POEM did not; a meaningful comparison of the overall results and conclusions from POEM with LHM was not possible.

Setting for POEM

The included SRs did not indicate the settings where their primary studies were conducted. However, in their general discussion of the procedure, authors of three SRs reported that POEM should be performed only in specialized centers by skilled surgeons with expertise in endoscopy. Barbieri et al. further advised that POEM operators must be aware of AEs that are not typical of non-POEM endoscopy, and ensure that surgeons with extensive experience in esophageal surgery are promptly available to treat POEM AEs. Marano et al. noted that currently a major challenge to the uptake of POEM is that it is only performed at highly advanced endoscopy and surgery centers by expert interventional gastroenterologists and surgical endoscopists.

Conclusions and Implications for Decision or Policy Making

Findings from the included SRs indicate that POEM is an effective procedure to treat achalasia as demonstrated by the significant reductions in Eckardt score, LES pressure, TBE column height, and an increase in weight gain. The results from one SR suggest that between 3.2% and 8.0% of patients may experience failure within 30 months of follow-up after the initial success. The findings of one SR suggest that POEM was effective at treating SED. The POEM procedure was safe with the main AEs being mild GER-related cases which were successfully controlled with PPIs. The most frequently reported procedural complications were subcutaneous emphysema, pneumoperitoneum, pleural effusion, pneumothorax, mucosal perforation, and pneumonia. There were no reports of POEM-related deaths.

The findings from two SRs indicate that patients who underwent either POEM or LHM had similar improvements in Eckardt score. However, the results of one SR suggest that the proportion of patients who experience improvements in dysphagia may be higher with POEM than with LHM with fundoplication. The results from one SR indicate that the rate of treatment failure was higher with LHM than with POEM. Findings from three SRs suggest that postoperative pain score and analgesic dose were similar among patients who
underwent POEM compared to those who had LHM. However, results for procedure time and length of hospital stay were inconclusive as some SRs found no significant differences between POEM and LHM while others did. On safety, results from two SRs indicate that the incidence of GER-related AEs was higher with POEM than LHM with fundoplication, whereas findings from two SRs suggest no statistically significant difference in postoperative symptomatic GER between POEM and LHM. The results from four SR indicate that procedural complications were similar between POEM and LHM. Considering the lack of long-term data on POEM data and the fact that the SRs in this review did not include any relevant RCT, a rigorous study comparing POEM with LHM over a long-term is needed to assess the comparative advantage of one procedure over the other.

The literature search did not identify cost-effectiveness studies or evidence-based clinical guidelines on the use of POEM for achalasia or other esophageal mobility disorders.
References


Appendix 1: Selection of Included Studies

536 citations identified from electronic literature search and screened

525 citations excluded

11 potentially relevant articles retrieved for scrutiny (full text, if available)

No potentially relevant reports were retrieved from other sources (grey literature, hand search)

11 potentially relevant reports

1 report excluded: complete overlap of included studies with a selected systematic review

10 reports were included in the review
## Appendix 2: Characteristics of Included Publications

### Table A1: Characteristics of Included Systematic Reviews and Meta-Analyses

<table>
<thead>
<tr>
<th>Author, Publication Date, Country</th>
<th>Study Design</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparator(s)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awaiz, 2017(^7) Australia, Malaysia, Pakistan, The United Kingdom</td>
<td>A systematic review and meta-analysis of seven NRC studies published from 2013 to 2015. The countries of origin and settings of the primary studies were not specified.</td>
<td>483 adult patients (POEM=233, LHM=250), (&gt;18 years old) diagnosed with all grades and subtypes of achalasia.</td>
<td>POEM</td>
<td>LHM</td>
<td><strong>Efficacy</strong>&lt;br&gt;• Clinical treatment failure&lt;br&gt;• Operative time;&lt;br&gt;• Length of hospital stay;&lt;br&gt;• Postoperative pain score&lt;br&gt;<strong>Safety</strong>&lt;br&gt;• Complication rate;&lt;br&gt;• Postoperative GERD;&lt;br&gt;• Long-term GERD.</td>
</tr>
<tr>
<td>Crespin, 2017(^1) Canada</td>
<td>A systematic review of 19 NRC studies (ten retrospective and nine prospective studies) published from 2012 to 2015. The countries of origin and settings of the primary studies were not specified.</td>
<td>1228 patients with achalasia. The mean age for the majority studies that reported age for ranged from 36 years to 64 years. One study was conducted in a pediatric population with a mean age of 13.8 years. Some primary studies had treatment-naive and treatment-exposed patients or achalasia patients who had a sigmoid esophagus.</td>
<td>POEM</td>
<td>None</td>
<td><strong>Efficacy</strong>&lt;br&gt;Mean difference in postoperative&lt;br&gt;• Eckardt score,&lt;br&gt;• LES pressure&lt;br&gt;<strong>Safety</strong>&lt;br&gt;• Intra- and postoperative complications,&lt;br&gt;• Development of GERD as measured by postoperative 24-hour pH studies</td>
</tr>
<tr>
<td>Khan, 2017(^11) USA, Italy, Japan</td>
<td>A systematic review and meta-analysis of eight studies (three prospective and five retrospective) published from 2014 to 2016. The countries of</td>
<td>179 patients with SEDs (made up of 116 with type III achalasia; 37 with JH esophagus, 18 with DES, and eight had hypertensive non-relaxing</td>
<td>POEM</td>
<td>None</td>
<td><strong>Efficacy</strong>&lt;br&gt;Cumulative clinical success in SEDs and success in subtypes of SEDs. Clinical success was defined as&lt;br&gt;• Eckardt scores of ≤3&lt;br&gt;• Improvement in the severity of</td>
</tr>
<tr>
<td>Author, Publication Date, Country</td>
<td>Study Design</td>
<td>Population</td>
<td>Intervention</td>
<td>Comparator(s)</td>
<td>Outcomes</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------</td>
<td>------------</td>
<td>--------------</td>
<td>---------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| **Repici, 2017**<sup>a</sup> <br> Italy, USA, Germany | A systematic review and meta-analysis of a total of 45 non-randomized studies (17 for POEM and 28 for LHM). The POEM studies were published between 2012 and 2016, and LHM studies were published between 1997 and 2017. The included studies were performed in 18 countries including Australia, Brazil, Canada, China, Egypt, India, Japan, USA and ten European countries. | A total of 4423 patients (aged ≥18 years) with a diagnosis of achalasia or other SEDs (1542 had POEM, and 2581 had LHM). | POEM | LHM with fundoplication<sup>a</sup> | Safety  
Cumulative post-procedure AE rate,  
Rate of AEs in subtypes of SEDs  
Primary Incidence post-procedural GERD according to:  
- GER symptoms  
- pH monitoring finding (abnormal acid exposure)  
- Endoscopic monitoring findings (esophagitis)  
Secondary  
Factors influencing the risk of developing GERD after procedure |
| **Schlottmann, 2017**<sup>b</sup> <br> The USA | Systematic review and meta-analysis of six RCTs and 68 NRC studies (one RCT and 20 non-randomized for POEM and five RCTs and 48 non-randomized for LHM). The POEM studies were published between 2013 | A total of 7,797 patients with achalasia (1,958 for POEM and 5,834 for LHM). The mean patients’ age across studies was POEM 45.1 years for POEM and 46.0 years for LHM. | POEM | LHM with fundoplication | Efficacy  
- Improvement of dysphagia  
- GERD symptoms  
- Reflux esophagitis  
- Abnormal acid exposure  
- Length of hospital stay  
Safety  
- Post-operative complications |
<table>
<thead>
<tr>
<th>Author, Publication Date, Country</th>
<th>Study Design</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparator(s)</th>
<th>Outcomes</th>
</tr>
</thead>
</table>
| Akintoye, 2016¹² The USA | A systematic review and meta-analysis of 36 NRC studies from 12 countries with the majority conducted in China, Japan, and the USA. The specific number of studies per each country of origin was not provided. | A total of 2373 patients (mean age of 45 ± 0.96 years) who underwent a POEM procedure for achalasia (98%) and other esophageal motility disorders between 2008 and 2014. | POEM | None | Efficacy  
• Primary—proportion of patients with Eckardt score ≤3 post-POEM  
• Secondary—mean post-POEM  
  o Eckardt score,  
  o LES pressure,  
  o IRP,  
  o TBE column height,  
  o Average weight gain, and  
  o Duration of hospitalization  
Safety  
• Procedure-related AEs  
• Incidence of post-POEM GER |
| Marano, 2016 Italy | A systematic review and meta-analysis of seven retrospective case-control studies published from 2013 to 2015. Six of the studies were performed in the USA, and one was a multicenter study with data from Germany, Netherlands, Switzerland, and Canada). | A total of 486 patients with achalasia (196 in POEM group and 290 in LHM group). Mean (SD) age was 49.5 (11.8) years in POEM group and 47.8 (11.8) in the LHM group. | POEM | LHM | Efficacy  
Mean difference in postoperative  
• Eckardt score,  
• Pain score,  
• Analgesic requirement,  
• Length of hospital stay,  
• Procedure time,  
Safety  
Odds ratio of  
• Intra- and postoperative complications,  
• Post-procedure symptomatic GER |
| Patel, 2016² The United Kingdom | A systematic review² of 22 NRC studies, (19 case series and | A total 1,262 patients with achalasia (1,122 for POEM and | POEM | LHM² | Efficacy  
• Eckardt and dysphagia scores and |
<table>
<thead>
<tr>
<th>Author, Publication Date, Country</th>
<th>Study Design</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparator(s)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbieri, 2015&lt;sup&gt;™&lt;/sup&gt; Italy</td>
<td>Systematic review and meta-analysis of 16 NRC studies from Asia, Europe, and the USA published between 2010 and 2014</td>
<td>A total of 551 patients with achalasia, 48% female, the median of the mean ages 44 years (range: 32–64 years).</td>
<td>POEM</td>
<td>None</td>
<td>Efficacy: • Technical feasibility—defined as the pooled prevalence of successfully-performed procedure • The clinical success rate—defined as the combined rate of clinically-relevant improvement of dysphagia. Safety: The pooled rate of • Post-POEM esophagitis • AEs that required medical or surgical intervention</td>
</tr>
<tr>
<td>Talukdar, 2015&lt;sup&gt;™&lt;/sup&gt; India, Japan</td>
<td>Systematic review and meta-analyses of one RCT and 28 NRC studies. Countries of origin of the single-center studies were China, Germany, Hong Kong, Japan, Korea, Netherlands, and the USA. There was one multicenter study conducted at A total of 1,045 patients who underwent POEM for achalasia. The mean (SD) age and the duration of disease were 50.5 (14.1) years and 51.01 (104.6) months respectively.</td>
<td>POEM</td>
<td>LHM&lt;sup&gt;™&lt;/sup&gt;</td>
<td>Primary improvement in • Eckhart’s score and • LES pressure. Subgroup analysis outcomes from five studies comparing POEM and LHM were: • Eckhart score, • The duration of hospital stay, • Postoperative pain score, • Postoperative analgesic dose, • Procedure time, • AEs, and • Post-procedure...</td>
<td></td>
</tr>
<tr>
<td>Author, Publication Date, Country</td>
<td>Study Design</td>
<td>Population</td>
<td>Intervention</td>
<td>Comparator(s)</td>
<td>Outcomes</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------</td>
<td>------------</td>
<td>--------------</td>
<td>---------------</td>
<td>----------</td>
</tr>
<tr>
<td>sites in Germany, Netherlands, and Canada.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>symptomatic GER</td>
</tr>
</tbody>
</table>

AE = adverse event; BMI = body mass index; DES = diffuse esophageal spasm; GER = gastroesophageal reflux; GERD = gastroesophageal reflux disease; IRP = integrated relaxation pressure; JH = jackhammer or nutcracker esophagus; LES = lower esophageal sphincter; LHM = laparoscopic Heller myotomy; POEM = per oral-esophageal myotomy; SED = spastic esophageal disorders; TBE = timed barium esophagram

* POEM and LHM were not necessarily compared in the same studies. Rather, the authors performed two separate systematic searches—one for POEM and the other for LHM with fundoplication—and reported the number of studies identified for each. Studies with LHM without fundoplication were excluded.

* Data from three studies comparing POEM to LHM were included in a pooled analysis, which was a secondary objective. However, POEM data from these comparative studies were also extracted for main systematic review.

* The LHM comparator applied in only the three studies included in the pooled analysis as explained in * above. The remaining 19 studies did not have any comparators to the POEM intervention.

* POEM (n=90 patients) was compared with LHM (n=160) in five of the 29 included studies.
Appendix 3: Critical Appraisal of Included Publications

Table A2: Strengths and Limitations of Systematic Reviews and Meta-Analyses using the AMSTAR tool

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>- The objectives, inclusion and exclusion criteria and main study findings were provided.</td>
<td>- All included studies were non-randomized with inherent limitations such as selection and performance biases.</td>
</tr>
<tr>
<td>- A comprehensive literature search involving multiple electronic databases as well as hand-searching of the bibliography of all selected articles for relevant studies was performed.</td>
<td>- The processes to screen and select studies for inclusion and for data extraction were not adequately reported.</td>
</tr>
<tr>
<td>- Each primary study was assessed for quality using Modified Newcastle-Ottawa scale, and the quality of all the studies was rated as moderate to high.</td>
<td>- Some reported outcome measures (prior endoscopic treatment; previous medical treatment; and previous Heller myotomy) were inadequately explained in the studies, making it difficult to understand and interpret reported data associated with them. These outcomes were not reported in this review.</td>
</tr>
<tr>
<td>- The meta-analysis provided a more reliable estimate of outcomes (with 95% CI) than would be possible from primary studies.</td>
<td>- The primary studies included both treatment-naive treatment-exposed patients and it is unknown whether or not the reported outcomes were influenced by the differences in treatment history.</td>
</tr>
<tr>
<td>- The authors had no funding and conflicts of interest to disclose.</td>
<td>- A meaningful long-term comparison of POEM with LHM was not possible since follow-up data for POEM beyond one year were not available.</td>
</tr>
<tr>
<td>- The objective, intervention, outcome measure and main study findings were reported.</td>
<td>- The authors reported a high degree of heterogeneity (quantitative measure not provided) in the primary studies and indicated that the variables were too few to detect or refute publication bias. Taken together with the high potential for bias due to the designs of primary studies, the conclusions of the meta-analysis may not be definite.</td>
</tr>
<tr>
<td>- The inclusion and exclusion criteria were described and multiple databases, supplemented with hand-searching of the bibliographies of all the selected articles, were used to identify relevant studies.</td>
<td>- The included primary studies were all non-randomized with inherent limitations such as selection and performance biases.</td>
</tr>
<tr>
<td>- Two reviewers independently screened identified citations for potentially relevant reports, and selected studies for inclusion independently. A third reviewer resolved all disagreements.</td>
<td>- Effect estimates from pooling across the studies were not determined due to high clinical, methodological, and statistical heterogeneity. However, forest plots showing the study-level estimates were presented.</td>
</tr>
<tr>
<td>- Two reviewers performed data extraction independently.</td>
<td>- The risk of bias of the primary studies was not reported although the authors stated that they assessed such risk using NIH tools and guidelines for case series, case-control, and before-and-after studies.</td>
</tr>
<tr>
<td>- Data were examined for clinical, methodological, and statistical heterogeneity.</td>
<td>- Publication bias was not reported.</td>
</tr>
<tr>
<td>- The authors reported that they have no conflicts of interest or financial ties to disclose.</td>
<td>- Only 14 of the 19 primary studies were included in the analysis and five studies were omitted for not providing any measure of variability for the preoperative Eckardt score.</td>
</tr>
<tr>
<td>Crespin, 2017*</td>
<td>- Fourteen studies had achalasia patients with sigmoid esophagus (n=99), and 15 studies had patients (n=289) with previous treatments for achalasia before POEM. Without a stratified analysis comparison of direct</td>
</tr>
<tr>
<td>Strengths</td>
<td>Limitations</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>The objective, intervention, outcome measures and main study findings were described.</td>
<td>comparison of data could lead to inaccurate conclusions.</td>
</tr>
<tr>
<td>The inclusion and exclusion criteria were reported, and a comprehensive literature search for relevant studies was performed using multiple electronic databases, and manual searches of the bibliographies of all the selected articles.</td>
<td>Without a comparator intervention, the comparative efficacy and safety profile of POEM for achalasia is uncertain.</td>
</tr>
<tr>
<td>Two reviewers independently screened identified citations for potentially relevant reports, and selected studies for inclusion and extracted data independently. Any disagreement between reviewers was to be discussed with a third reviewer.</td>
<td>It is unknown if the reported findings will be sustainable over the long-term</td>
</tr>
<tr>
<td>The quality of the primary studies was independently assessed as by two reviewers using the NIH quality assessment tool for before-after studies with no control group. Any disagreement between reviewers was resolved by consensus.</td>
<td></td>
</tr>
<tr>
<td>Heterogeneity was assessed using $I^2$ statistics and Cochran Q test and found to be low across studies.</td>
<td></td>
</tr>
<tr>
<td>Egger’s test for publication bias was negative overall. However, a funnel plots assessment could not be done due to the small number of studies.</td>
<td></td>
</tr>
<tr>
<td>The authors reported that they have no conflicts of interest or financial relationships relevant to this article.</td>
<td></td>
</tr>
</tbody>
</table>

Khan, 2017

- All the included primary studies were non-randomized with inherent limitations such as selection and performance biases. Also, only two out of the eight studies were rated as good quality while the remaining six studies were of fair quality.
- Although the duration of follow-up in the primary studies ranged from three months to three years, the varying degrees of follow-up implies that the findings from the meta-analysis might not represent the long-term efficacy well.
- Without a comparator intervention, the comparative efficacy and safety profile of POEM for achalasia is uncertain.

Repici, 2017

- A comprehensive literature search for relevant studies was performed using multiple electronic databases supplemented by manual searches of references of included studies and review articles.
- The inclusion and exclusion criteria, as well as the reasons for excluding trials, were reported.
- Three review authors independently screened the titles and abstracts identified by the search against the inclusion criteria. Review author pairs then screened the full-text of potentially relevant reports to select studies for inclusion. Disagreements were resolved through discussion of all the authors.
- Two reviewers extracted data independently and in duplicate from each eligible study, resolving disagreements by discussion, and the arbitration with other reviewers to reconcile discrepancies.
- Two reviewers assessed the quality of included studies using a tool developed by Moga et al. 17, and discrepancies were resolved by discussion.
- Regarding after-POEM and/or LHM incidence of symptomatic GERD, definition of reflux symptoms was not standardized across the various studies.
- There was a high degree of heterogeneity in the estimates of abnormal acid exposure and esophagitis. Also, the Egger test of publication bias was positive for GER symptoms and abnormal acid exposure estimates after-LHM.
The authors report no conflicts of interest.

Schlottmann, 2017°

- The objectives of the study was provided, and the interventions, outcome measures, and study findings were well-reported,
- The inclusion and exclusion criteria presented, and two authors independently reviewed and selected studies for inclusion and abstracted the relevant data for analysis, resolving disagreements by consensus.
- The mean patients' age across studies and the proportion of patients who received prior treatment was not significantly different between POEM and LHM. However, the proportion of males was higher among LHM studies than among POEM studies. The clinical significance of the gender difference in the two procedures is unclear.
- Although there was a significant difference in the duration of follow-up between LHM and POEM studies, a sensitivity analysis involving studies with a comparable length of follow-up was performed to compare the main efficacy outcome (improvement of dysphagia) between LHM and POEM studies.
- The authors report no conflicts of interest.

- The literature search was based on only one electronic database so it is unknown if the search for literature was sufficiently comprehensive.
- Only one of the 74 included studies was an RCT. The remaining 73 primary studies were none-randomized and thus have a high potential for selection and performance biases.
- The methodological quality of the included studies was not assessed
- Publication bias was not determined.
- The POEM studies have significantly shorter follow-up than the LHM studies. Thus a comparison of long-term effects was not possible, and the difference may influence the overall results reported.
- It is unknown whether heterogeneity among studies was formally evaluated. However, the authors reported (without a quantitative score) that significant heterogeneity existed in the reporting of outcomes among studies which may have affected their results.
- The outcomes were reported as dichotomous variables—present or not present—which made no differentiation between clinically significant effects and minor changes with no clinical consequence.
- Although the observed post-procedure proportions of incidence of symptoms of GER and GERD based on EGD were not significantly different among POEM and LHM, the author's calculated odds ratios with significantly large differences in favor of LHM by adjusting for follow-up time as a significant confounder. However, the assumption for the confounding was not explained, and the factor (weight) by which the data was adjusted to arrive at the significantly different ORs was not provided.

Akintoye, 2016°

- The objectives, the intervention, outcome measures, and study findings were well-reported.
- The inclusion and exclusion criteria were provided.
- A comprehensive search involving multiple electronic databases was conducted for relevant literature without filtering by language and supplemented with hand-searching of the bibliography of all selected articles.
- The included primary studies were none-randomized and thus have a high potential for selection bias.
- Only one reviewer screened titles and abstracts from search results for relevant studies. Also, although two reviewers screened full-text articles for inclusion, there were no indications of how disagreements, if any, were resolved.
### Strengths

- Two of the reviewers independently screened full-text reports according to the specified inclusion criteria, and extracted data.
- Heterogeneity between studies was assessed via visual inspection of the forest plot, Cochran Q statistic, and the $I^2$ statistic.
- Evaluations using visual inspection of the funnel plot and Egger’s test were negative for publication bias. Also, meta-regression and sensitivity analysis showed that the results were not substantially influenced by the results of any single study, which suggest they are robust.
- The meta-analysis provided a more reliable estimate of outcomes (with 95% CI) than would be possible from primary studies.

### Limitations

- The quality of included studies was not assessed.
- Significant heterogeneity existed among studies which were pooled for outcome analysis. The heterogeneity was significantly explained by the distribution of female patients across the studies and marginally by the study design and the continent of origin of the study which may suggest limited generalizability of the findings.
- One of the five authors had a consultancy relationship with and had received research grants from some medical devices and healthcare products companies. However, it does not appear that this influenced the reported findings of the study.
- The duration of follow-up was not longer than one year (12 months). Therefore, it is unknown if the reported finding would be sustained over a long-term (>12 months).

### Marano, 2016

- The objectives, outcome measures, and study findings were well-reported, and the inclusion and exclusion criteria provided.
- The reviewers conducted a comprehensive search involving multiple electronic databases as well as hand-searching the bibliography of all selected articles for relevant studies.
- Two of the investigators independently reviewed all eligible studies for content and screened them according to the specified inclusion criteria, and extracted data. Disagreements were resolved by discussion between the two, and it was planned that a third author would be involved if no agreement could be reached.
- The risk of bias of each primary study was assessed the ACROBAT-NRSI with all the studies rated as having a moderate risk of bias.
- Although significant heterogeneity existed among studies which were pooled to determine individual efficacy outcomes, sensitivity analyses by sequentially omitting each study did not change the results; suggesting that the findings are robust. The test for heterogeneity among studies which were pooled for safety outcomes was not significant.
- The meta-analysis provided a more reliable estimate of outcomes (with 95% CI) than would be possible from primary studies.
- Publication bias was evaluated using Egger’s regression.
- The authors had no funding and conflicts of interest to disclose.

### Patel, 2016

- The included primary studies were none-randomized and thus have a high potential for selection bias.
- There was high heterogeneity among studies. However, it does not appear that this affected the finding since sensitivity analyses by sequentially omitting each study did not change the outcomes.
- Egger’s regression analysis suggested significant publication bias. However, a Funnel plot to test the publication bias could not be produced due to the limited number of studies included in each analysis (below 10).
- Most of the primary studies included both treatment-naïve patients and those who had undergone previous endoscopic or surgical interventions for achalasia. Without stratified analysis, it is unknown whether or not the prior treatment history influenced the reported outcomes.
- None of the primary studies report the results of follow-up longer than one year. With a mean follow-up duration of 6.4 ± 4.6 months, it is unknown if the reported finding would be sustained over a long-term (>12 months).
### Strengths
- A comprehensive systematic search for potentially relevant articles was performed in multiple electronic databases.
- Two reviewers independently screened the titles and abstracts of studies identified through the electronic search and obtained the full texts of potentially relevant articles. Further, they hand-searched the reference lists of all the selected studies for potentially relevant articles. Any disagreement was resolved by a third reviewer.
- Heterogeneity among studies was assessed using the Cochrane’s Q statistic.
- The quality of studies included in the systematic review was assessed using an assessment tool described by Downs et al.\(^{18}\)
- The authors stated that there were no funding sources for the study or any associated financial conflicts.

### Limitations
- in the majority of studies which made interpretation and summative analysis difficult and precluded any meaningful comparison between studies.
- There was limited data on long-term POEM outcomes.

<table>
<thead>
<tr>
<th>Barbieri, 2015 (^{19})</th>
</tr>
</thead>
<tbody>
<tr>
<td>- The objectives, outcome measures, and study findings were well reported.</td>
</tr>
<tr>
<td>- The inclusion and exclusion criteria specified, and the reasons for exclusion provided.</td>
</tr>
<tr>
<td>- Two reviewers screened titles and abstracts of papers, and two reviewers independently pre-screened the full-text study report for relevance and selection for inclusion. Any disagreements were resolved through discussion.</td>
</tr>
<tr>
<td>- One reviewer extracted data from relevant studies, which were checked by a second reviewer. Any disagreements were resolved through discussion with a third reviewer.</td>
</tr>
<tr>
<td>- The methodological quality of each primary study was evaluated using the STROBE checklist for observational studies and recorded.</td>
</tr>
<tr>
<td>- Egger’s regression did not detect any publication bias.</td>
</tr>
<tr>
<td>- By pooling finding from several primary studies, which on their own seemed underpowered, the meta-analysis provided estimate of POEM efficacy with a 95% CI that was more reliable.</td>
</tr>
<tr>
<td>- According to the authors, the research received no specific grant from any funding agency in public, commercial or not-for-profit sectors; and there were no conflicts of interest.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Talukdar, 2015 (^{15})</th>
</tr>
</thead>
<tbody>
<tr>
<td>- The included primary studies were non-randomized, and the outcomes assessors were not blinded to the procedure performed. Thus there was the potential for assessment bias, at least in theory.</td>
</tr>
<tr>
<td>- The primary studies applied different criteria to measure the clinical degree of dysphagia or the adverse events rates. However, given that the outcomes from the various studies were in the same direction, it is unlikely that this could change the reported results.</td>
</tr>
<tr>
<td>- The median post-POEM follow-up was six months (range: 3–12 months). Thus the sustainability of the reported finding in the long-term (&gt;12 months) is unclear.</td>
</tr>
</tbody>
</table>

---

**Note:**
- The information is extracted from a systematic review and meta-analysis focusing on the effectiveness of per oral endoscopic myotomy (POEM) for esophageal motility disorders.
- The review assesses the strengths and limitations of the included studies, highlighting the importance of comprehensive search strategies and the need for high-quality study design.
- The limitations include the lack of randomization in many studies, potential performance and selection biases, and limited long-term outcomes data.

---

**PER ORAL ENDOSCOPIC MYOTOMY FOR ESOPHAGEAL MOTILITY DISORDERS**

---

**References:**
- Barbieri, 2015
- Talukdar, 2015

**Footnotes:**
- \(^{18}\) Downs et al. (1998)
- \(^{19}\) Egger's regression
Strengths

- The quality of the included studies was assessed using a study rigor table that was previously developed and validated.
- Between-study heterogeneity was assessed by the $I^2$ statistic, and the Q measure was used to evaluate the significance of heterogeneity.
- Publication bias was initially evaluated and quantified by the Egger’s test.
- Two of the three authors stated they had no conflicts of interest or financial ties to disclose while the third was an advisor to an optics and reprography products company. However, it is unlikely that his advisory relationship with the firm influenced the study and its reporting.

Limitations

- There were significant publication bias and heterogeneity among the studies that reported data for the primary outcomes—change in Eckhart’s score and LES pressures.
- Most of the studies included patients who were both treatment naïve and underwent previous endoscopic or surgical interventions for achalasia. However, the analysis did not stratify the differences in treatment exposure. Thus it is unclear how the efficacy of POEM in treatment naïve patients compared to that in those who failed to respond to previous interventions or had a relapse.
- Majority of the studies did not provide results of long-term follow-up.

ACROBAT-NRSI = A Cochrane risk of bias assessment tool for nonrandomized studies of interventions; CI = confidence interval; $I^2$ = I-squared statistic; NIH = National Institutes of Health; POEM = per oral esophageal myotomy; STROBE = Strengthening the Reporting of Observational Studies in Epidemiology; TBE = timed barium esophagram
## Table A3: Summary of Findings of Included Studies

<table>
<thead>
<tr>
<th>Main Study Findings</th>
<th>Author’s Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Efficacy outcomes</strong></td>
<td><strong>“POEM compares favorably to LHM for achalasia treatment in short-term perioperative outcomes. However, there was a significantly higher clinical treatment failure rate for LHM on short-term postoperative follow-up. Presently long-term postoperative follow-up data for POEM beyond one year are unavailable and eagerly awaited.“</strong>&lt;sup&gt;9&lt;/sup&gt; page 123</td>
</tr>
<tr>
<td>Clinical treatment failure (three studies)</td>
<td>There was a significantly higher short-term clinical treatment failure rate for LHM (OR = 9.82; 95% CI, 2.06, 46.80; <em>P</em> &lt; 0.01).</td>
</tr>
<tr>
<td>Operative time in minutes (six studies)</td>
<td>LHM had longer operative time than POEM but the difference was not statistically significant (WMD = 26.28; 95% CI: 11.20, 63.70; <em>P</em> = 0.17).</td>
</tr>
<tr>
<td>Length (in days) of hospital stay (six studies)</td>
<td>Length of hospital stay after procedure was not significantly different between POEM and LHM (WMD = 0.30; 95% CI: 0.24, 0.85; <em>P</em> = 0.28).</td>
</tr>
<tr>
<td>Postoperative pain score</td>
<td>Pain score after procedure was not significantly different between POEM and LHM (WMD = 0.26; 95% CI: 1.58, 1.06; <em>P</em> = 0.70).</td>
</tr>
<tr>
<td><strong>Safety</strong></td>
<td>“POEM is a safe and effective alternative for the treatment of achalasia. However, only short-term follow-up data compared with LHM are available. RCTs and long-term follow-up studies are needed to establish the efficacy and safety of POEM in the management of patients with achalasia.“&lt;sup&gt;1&lt;/sup&gt; page 2188</td>
</tr>
<tr>
<td>Complication rate (six studies)</td>
<td>The overall complication rate was comparable between POEM and LHM (OR = 1.25; 95% CI: 0.56, 2.77; <em>P</em> = 0.59).</td>
</tr>
<tr>
<td>Postoperative GERD (four studies)</td>
<td>GERD rate after procedure was not significantly different between POEM and LHM (OR = 1.27; 95% CI: 0.70, 2.30; <em>P</em> = 0.44).</td>
</tr>
<tr>
<td>Long-term GERD</td>
<td>The authors reported that long-term GERD symptoms based on two studies were not significantly different between POEM and LHM (WMD = 1.06; 95% CI: 0.27, 4.1; <em>P</em> = 0.08). However, they did not define &quot;long-term&quot;. This is important since they had stated in the text that &quot;long-term postoperative follow-up data for POEM beyond one year are unavailable in all the comparative trials and therefore meaningful comparison with LHM is impossible.“&lt;sup&gt;9&lt;/sup&gt; page 130</td>
</tr>
</tbody>
</table>

| Cressin, 2017<sup>1</sup> | |
**Main Study Findings**

<table>
<thead>
<tr>
<th>POEM data (n=760 patients), the study-level mean (SD) post-POEM Eckardt score ranged 0.4 (0.7) to 2.7 (1.9) and were lower than the Pre-POEM scores which ranged from 6.3 (1.0) to 8.4 (3.2).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LES pressure</strong></td>
</tr>
<tr>
<td>Overall, there was a significant improvement in post-POEM LES pressure compared with pre-procedure values.</td>
</tr>
<tr>
<td>From 10 studies with the corresponding pre- and post-POEM data (n=483 patients), the study-level mean (SD) post-POEM LES pressure score ranged 6.49 (1.42) to 29.8 (22.0) and were lower than the Pre-POEM scores which ranged from 22.0 (10.0) to 71.2 (20.8).</td>
</tr>
</tbody>
</table>

A high level of heterogeneity in mean change from pre-operation was observed across the studies ($I^2 = 97.10\%$).

**Safety**

- **Complications**
  - Nineteen studies (n=1299 patients) reported complication. The most frequently reported complications were subcutaneous emphysema 19% (n = 131), pneumoperitoneum 17% (n = 221), pleural effusion 10.2% (n = 132), mucosal perforation 9% (n = 118), pneumonia 7.9% (n = 103), pneumothorax 5.3% (n = 69), and pneumomediastinum 4.5% (n = 58). |
  - All complications related to the procedure were managed conservatively using endoscopic clipping, suturing, or hemostasis interventions. There was no mortality reported. None of the POEM procedures had to be converted to surgery. |

- **Development of GERD**
  - Only four studies evaluating 147 patients assessed GERD after POEM, using 24-h pH monitoring. The range of abnormal esophageal acid exposure varied from 20 to 53%; all patients were symptomatically controlled with proton pump inhibitors. |

**Efficacy**

- **Clinical success in all SEDs (eight studies)**
  - The overall WPR for clinical success of POEM in all SEDs was 87% (95% CI: 78, 93), Cochran Q test $P = 0.12$, $I^2 = 37\%$. |

- **Clinical success in individual SEDs**
  - From eight studies, the WPR for clinical success of POEM in type III achalasia was 92% (95% CI: 84, 96), Cochran Q test $P = 0.73$, $I^2 = 0\%$. |
  - From five studies, the WPR for clinical success in JH esophagus was 72% (95% CI: 55, 83), Cochran Q test $P = 0.91$, $I^2 = 0\%$. |

**Author’s Conclusion**

- "On the basis of this meta-analysis, POEM seems to be a logical, efficacious, and safe treatment modality for patients with SEDs. Jackhammer esophagus was associated with decreased clinical response rate in comparison with spastic achalasia and diffuse esophageal spasm. Before we can surely say that POEM is ready for prime time in managing patients with spastic esophageal disorders, larger prospective studies are required."
Main Study Findings

- From four studies, the WPR for clinical success of POEM in DES was 88% (95% CI: 61, 97), Cochran Q test P = 0.80, I² = 0%.

Safety

- The overall WPR for post-POEM AEs of in all SEDs of 14% (95% CI: 9, 20), Cochran Q test P = 0.57, I² = 0%.
- From eight studies, the WPR for post-POEM AEs in patients with type III achalasia was 11% (95% CI: 6, 19), Cochran Q test P = 0.77, I² = 0%.
- From five studies, the WPR for post-POEM AEs in patients with jackhammer esophagus was 16% (95% CI: 7, 32), Cochran Q test P = 0.94, I² = 0%.
- From four studies, the WPR for post-POEM AEs in patients with DES was 14% (95% CI: 4, 39%), Cochran Q test P = 0.89, I² = 0%.
- Most (74%) AEs did not require further intervention and were managed conservatively.
- Only 3.3% of patients had adverse events that required subsequent intervention and/or prolonged hospitalization.

Repici, 2017

Overall GER symptoms

- POEM
  In a meta-analysis of 17 studies (1275 patients), the pooled rate of overall GER symptoms post-POEM was 19.0% (95% CI: 15.7, 22.8), Between-study heterogeneity was moderate (I² = 43.3%; P = 0.024).
- LHM with fundoplication.
  In a meta-analysis of 20 studies (1136 patients), the pooled rate of overall GER symptoms post-LHM was 8.8% (95% CI: 5.3, 14.1), Between-study heterogeneity was high (I² = 80%).

Abnormal acid exposure

- POEM
  A meta-analysis of five studies (289 patients) found that the pooled rate estimate of abnormal acid exposure post-POEM was 39.0% (95% CI: 24.5, 55.8). (I² = 85.0%; P < 0.0001).
- LHM with fundoplication
  A meta-analysis of 14 studies (1022 patients) found that the pooled rate estimate of abnormal acid exposure post-LHM was 16.8% (95% CI: 10.2, 26.4). Heterogeneity was high (I² = 86.0%)

Esophagitis

- POEM
  In a meta-analysis of 12 studies (1056 patients), the pooled rate estimate of esophagitis was 29.4% (95% CI: 18.5, 43.3) post-PEM. Heterogeneity across was high (I² = 93%).
- LHM with fundoplication.

Author’s Conclusion

- “The incidence of reflux disease appears to be significantly more frequent after POEM than after LHM with fundoplication. Monitoring pH and ensuring appropriate treatment after POEM should be considered in order to prevent long-term reflux-related adverse events.”
Main Study Findings

In a meta-analysis of five studies (752 patients), the pooled rate estimate of esophagitis was 7.6% (95% CI: 4.1, 13.7) post-LHM. The between study heterogeneity was high ($I^2 = 62\%$; $P = 0.03$).

Schlottmann, 2017

**Efficacy**
- Across all the studies, an average of 93.2% of patients who underwent POEM reported improvement in dysphagia compared with 87.7% of patients who underwent LHM.
- Predicted probabilities for improvement in dysphagia after POEM and LHM were 93.5% for and 91.0% ($P = 0.01$), respectively, at 12 months; and 92.7% and 90.0% ($P = 0.01$), respectively, at 24 months.
- On average, length of hospital stay was 1.03 days longer after POEM ($P = 0.04$).

**Safety**
- 18.5% patients who underwent POEM reported GERD symptoms compared with 17.5% of those who underwent LHM. It was estimated that patients undergoing POEM were more likely to develop GERD symptoms (OR = 1.69, 95% CI: 1.33; 2.14; $P <0.0001$).
- Using EGD, GERD was observed in 22.4% of patients treated with POEM compared with 11.5% of those who underwent LHM. Thus patients undergoing POEM were estimated to be more likely to develop erosive esophagitis than those receiving LHM (OR = 9.31, 95% CI: 4.71, 18.85; $P < 0.0001$).
- Abnormal 24 hour pH monitoring GERD findings were observed in 47.5% of POEM patients compared with 11.1% in LHM patients. It was estimated that abnormal acid exposure was more like in patients undergoing POEM that those receiving LHM (OR = 4.30, 95% CI: 2.96, 6.27; $P < 0.0001$).

Author’s Conclusion

- “Short-term results show that POEM is more effective than LHM in relieving dysphagia, but it is associated with a very high incidence of pathologic reflux.”
- “The available data suggest that POEM is very effective in the relief of symptoms in patients with achalasia. However, this technique is associated with a very high incidence of pathologic reflux. The clinical sequelae of the increase in pathologic reflux are currently incompletely understood, but it is plausible that longer-term outcomes associated with POEM may demonstrate GERD complications such as stricture and/or Barrett’s esophagus. Longer follow-up studies and randomized trials comparing POEM to LHM are necessary to establish the role of this new technique in the treatment algorithm of achalasia.”

Akintoye, 2016

**Efficacy**
- A meta-analysis of 27 studies involving 2065 patients found that clinical success, defined as Eckardt score ≤3, was achieved in 98% (95% CI: 97, 100; $I^2 = 67.9\%$) of patients after the procedure.
- The mean (SD) weight gain after a mean follow-up of 7.4 months was 5.4 (0.73) kg (488 patients from six studies).
- The mean (SD) Eckardt score decreased from a pre-operative score of 6.9 (0.15) (from 2,223 patients in 31 studies) to 0.77 (0.10) within one month post-POEM (648 patients in 11 studies), 1.0 (0.10) within six months (1,437 patients in 17 studies), and 1.0 (0.08)

“...In conclusion, POEM is a safe and effective technique according to the large and broad body of current medical literature. It compares favorably to surgical management. The procedure warrants consideration as first-line therapy when an expert operator is available.”

PER ORAL ENDOSCOPIC MYOTOMY FOR ESOPHAGEAL MOTILITY DISORDERS
Main Study Findings | Author’s Conclusion
--- | ---
within 12 months (1,249 patients in 11 studies). |  
- The six-month postoperative analysis showed that the mean (SD) LES pressure decreased from 33 (1.7) mmHg pre-POEM (1,976 patients in 24 studies) to 14 (1.2) mmHg post-POEM (1,569 patients in 18 studies).
- The six-month postoperative analysis showed that the mean (SD) IRP decreased from 30 (1.4) mmHg pre-POEM (470 patients in 11 studies) to 13 (1.6) mmHg post-POEM (456 patients in 10 studies). Overall, compared with preoperative values, there were significant decreases (P<0.05) in the average post-POEM LES pressure and the IRP.
- The one minute mean (SD) TBE column height decreased from 14 (2.3) cm pre-POEM to 4.2 (0.77) cm post-POEM (186 patients in three studies).
- The five minutes mean (SD) TBE column height decreased from 9.7 (1.9) cm and 2.6 (0.72) cm post-POEM (232 patients in four studies). Overall, compared with preoperative values, there were significant decreases (P< 0.05) in the postoperative Eckardt scores, LES pressure, IRP, and the TBE column heights in each study.

Safety
AEs after a mean follow-up of eight months the post-POEM were as follows:
- The rates of symptomatic GER (from 29 studies 2,142 patients) was 8.5% (95% CI: 4.9, 13; I² =87%),
- Esophagitis on EGD (20 studies 1,762 patients) was 13% (95% CI: 5.0; 23; I² = 96%),
- Abnormal acid exposure (five studies 336 patients) was 47% (95 % CI: 21, 74; I² = 96%)

The most common perioperative complications were
- Subcutaneous emphysema—7.5% (95 % CI: 3.5, 12; I² = 85%) (24 studies 1,286 patients)
- Pneumoperitoneum—6.8% (95 % CI: 1.9, 14; I² = 94) (22 studies 1,760 patients)
- Mucosal injury—4.8% (95 % CI: 2.0, 8.5; I² = 79%) (20 studies 1,682 patients)

Marano, 2016

Efficacy
- **Eckardt Score Reduction (five studies)**
  No significant difference was found between POEM and LHM (MD = -0.659, 95% CI: -1.70, 0.38; P = 0.217). Heterogeneity among the pooled studies was significant (I² = 94.3%, P <0.001).
- **Operative Time (meta-analysis of four studies)**
  The mean time for POEM was not significantly different from that of LHM (MD = -0.354, 95% CI: -1.12, 0.41; P = 0.36). Heterogeneity among the pooled studies was significant (I² = 82.1%, P <0.001).
- **Length of Hospital Stay (four studies)**
  POEM was associated with a comparable decrease in Eckardt score as LHM, and there no significant difference between the two interventions as regards the time for the procedure, postoperative pain score, and analgesic requirement. Length of hospital stay was significantly less in POEM compared to LHM.
- Pooled outcomes on post-procedure symptomatic GER showed a significant reduction in favor of LHM compared to POEM group. Nevertheless, future studies are needed to investigate the long-term effects of POEM on acid reflux.
- “This meta-analysis points to the urgent need for high-quality clinical trials with long-term evaluation
## Main Study Findings

<table>
<thead>
<tr>
<th>Study Findings</th>
<th>Author’s Conclusion</th>
</tr>
</thead>
</table>
| The length of hospital stay after intervention was shorter with POEM than LHM (MD = -0.629, 95% CI: -1.256, -0.002; P = 0.049). Heterogeneity among the pooled studies was significant (I² = 71.8%, P <0.014). | comparing POEM with other standard procedures, including surgical myotomy, in a randomized manner. 
Patel, 2016*6 page 7 |
| **Postoperative Pain Score and Analgesic Dose (two studies)**                  |                                                                                     |
| The reduction in pain score after procedure was not statistically significantly different between POEM and LHM (MD = -1.86, 95% CI: -5.17, 1.44; P = 0.268). The reduction in dose of analgesic medication (in morphine equivalent) was also not statistically significant between the two groups (MD = -0.74, 95% CI: -2.65, 1.16; P = 0.445) |                                                                                     |
| Although significant heterogeneity existed among studies which were pooled to determine individual efficacy outcomes, sensitivity analyses by sequentially omitting each study did not change the outcomes. |                                                                                     |
| **Safety**                                                                    |                                                                                     |
| **Symptomatic GER (five studies)**                                             |                                                                                     |
| The POEM procedure was significantly more likely to be associated with symptomatic GER than LHM (OR = 1.81, 95% CI: 1.11, 2.95; P = 0.017). The test for heterogeneity was not significant (I² = 0%, P = 0.851). |                                                                                     |
| **Complications (four studies)**                                               |                                                                                     |
| Overall, the rate of complications did not differ between POEM and LHM (OR = 1.11, 95% CI: 0.5, 2.44; P = 0.78). The test for heterogeneity was not significant among the pooled studies (I² = 0%, P = 0.482). |                                                                                     |
| Similarly, there was no statistically significant difference between POEM and LHM in either the major complication rate (OR = 1.33, 95% CI: 0.24, 7.35; P = 0.74) or the minor complication rate (OR = 1.82, 95% CI: 0.48, 6.9; P = 0.375). |                                                                                     |
| Major complications were bleeding and perioperative perforation, whereas minor complications were urinary retention, splenic capsular injury, atrial fibrillation, vagal nerve injury, and subcutaneous emphysema. |                                                                                     |
| **Efficacy**                                                                  |                                                                                     |
| Symptoms of achalasia improved as demonstrated by a decrease in mean (SD) Eckardt score from the pre-POEM value of 6.8 (1.0) to post-POEM score of 1.2 (0.6). |                                                                                     |
| The post-POEM LES pressure was 65.0% lower than the pre-POEM value.            |                                                                                     |
| The TBE column height at 5 minutes decreased by 79.7% post-POEM compared the pre-POEM height. |                                                                                     |
| Five studies evaluated treatment failure of POEM (Eckardt score ≥4); with two reporting 4.3% and 5.4% failure at six months, and one study each reporting failure rate of 8% at 15 months, 5.5% at 24 months, |                                                                                     |
| “In conclusion, POEM is a safe and effective treatment for achalasia, showing significant improvements in objective metrics and achalasia-related symptoms. Randomized comparative studies of LHM and POEM are required to determine the most effective treatment modality, with additional investigation of more novel metrics such as EndoFLIP and objective measures of reflux (e.g., impedance studies) pre- and post-treatment.”** page 818 |                                                                                     |
### Main Study Findings

#### Safety
- Endoscopy showed esophagitis in 19% of patients post-POEM compared with none observed pre-POEM.
- Complications were mostly minor, with the most commonly reported included being subcutaneous emphysema (31.6%), pneumo-peritoneum (30.6%), and pneumothorax (11.0%) and.
- The most common major complications were postoperative bleeding (1.1%) and mediastinal leak (0.3%). One person (0.09%) died from cachexia due to lack of oral. The death was considered a consequence of long-term treatment failure rather than a direct result of the POEM procedure.

### POEM versus LHM (Three studies, 213 patients)
- The operative time for POEM and LHM were (WMD = −15.13 minutes; 95% CI: −50.09, 19.83; P = 0.4). There was evidence of significant statistical heterogeneity (Cochran Q = 581.38; P < 0.0001).
- The length of hospital stay was marginally but not significantly shorter for POEM (WMD = −0.95 days; 95% CI: −1.92, 0.01; P = 0.05). There was evidence of statistical heterogeneity (Cochran Q = 10.93; P = 0.004).
- The pooled ORs of AEs was 1.2 (95% CI: 0.5, 2.86; P = 0.69). There was no evidence of statistical heterogeneity (Cochran Q = 2.22; P = 0.33).
- The pooled ORs for perforation was 0.81 (95% CI: 0.28, 2.3; P = 0.69). There was no evidence of statistical heterogeneity (Cochran Q = 1.24; P = 0.54).

### Efficacy
#### Successful procedures
- The pooled prevalence of successfully-performed procedures was 97% (95% CI: 94, 98). The analysis did not detect significant heterogeneity in the pooled primary studies (χ² = 13.59, p = 0.557, I² = 0%; 95% CI: 0, 47.3).

#### Clinical success
- The pooled rate of clinical success prevalence was 93% (95% CI: 90, 95), with no statistically significant heterogeneity among pooled studies (χ² = 13.86, p = 0.460, I² = 0%; 95% CI: 0, 53).

#### Dysphagia scores
- The mean of the dysphagia scores was reduced from 7 pre-POEM to 1 post-POEM (p<0.001; Student test).

### Safety
- Overall, 47 out of 354 patients had post-POEM esophagitis resulting in a pooled rate of 13% (95% CI: 10, 17).

### Author’s Conclusion

> “In conclusion, we showed that POEM is a highly effective and safe procedure, when performed in a dedicated setting, and it also appeared to be reproducible and generalizable in different continents; however, when considering the observational type of studies included, we believe further data from randomized studies, as well as on the long-term outcome, are needed.”

---

**Barbieri, 2015**

---

**PER ORAL ENDOSCOPIC MYOTOMY FOR ESOPHAGEAL MOTILITY DISORDERS**
### Main Study Findings

- 79 out of 551 experienced AEs that required medical or surgical intervention giving a rate of 14% (95% CI: 11.17). One patient out of 551 patients required post-POEM surgery resulting in a pooled rate of 0.2% (95% CI: 0.05).
- There was no POEM-related death.

**Efficacy**

- A meta-analysis of 19 studies (662 patients) found a significant reduction in Eckhart's score in the patients post-POEM. The overall effect size (Z) was -7.95 (p <0.0001), with a standardized difference in means of -0.938 (95% CI: -1.169, -0.706). Heterogeneity among the studies was significant (Q = 83.06; I² = 78.33%; p<0.0001), and Egger's test for publication bias was positive (intercept = -3.19; 95% CI: -3.56, -2.82; p<0.0001).
- A meta-analysis of 16 studies (568 patients) showed a significant improvement in the resting LES pressure. The overall Z was -7.28 (p<0.0001), with a standardized difference in means of -0.869 (95% CI: -1.102, -0.635). The heterogeneity between studies was significant (Q = 61.44; I² = 75.68%; p<0.0001) and Egger's test for publication bias was positive (intercept = -3.06; 95% CI: -3.58, -2.53; p<0.0001).

**Safety**

- Twenty-six of the 29 included studies reported complications of the POEM. The most commonly reported were subcutaneous emphysema (20.35%), pleural effusion (16.25%), pneumoperitoneum (15.09%), and pneumonia (9.19%), pneumothorax (8.13%), mediastinal emphysema (4.55%); and esophageal and gastric perforation (2.41%). Most of the complications were minor and self-limited and could be managed conservatively.
- Overall, post-POEM GER/reflux esophagitis was reported in 10.9% of patients with symptomatic GER could be satisfactorily managed with proton pump inhibitor and antacids.
- There was no mortality, and none of the POEM procedure had to be converted to surgery.

**POEM versus LHM (Five studies, 250 patients)**

- A sub-analysis of five studies comparing POEM to LHM showed that the time for the procedure was significantly less for POEM than for LHM (Z = -2.220; p = 0.026).
- There was no statistically significant difference between POEM and LHM in the reduction of post-operative Eckhart's score, pain score, the analgesic dose in morphine equivalent, and the length of hospital stay.
- Also, the risk of AEs and procedural complications, and the development of symptomatic GER did not differ.

### Author’s Conclusion

“POEM is effective for achalasia and has similar outcomes as LHM. Multicenter randomized trials need to be conducted to further compare the efficacy and safety of POEM between treatment naïve achalasia patients and those who failed treatment.”

---

**Talukdar, 2015**

---

**POEM is effective for achalasia and has similar outcomes as LHM. Multicenter randomized trials need to be conducted to further compare the efficacy and safety of POEM between treatment naïve achalasia patients and those who failed treatment.”

---

**Page 3030**
### Main Study Findings

<table>
<thead>
<tr>
<th>Author’s Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>between POEM and LHM.</td>
</tr>
</tbody>
</table>

AE = adverse event; CI = confidence interval; DES = diffuse esophageal spasm; EGD = esophagogastroduodenoscopy; MD = mean difference; GER = gastroesophageal reflux; GERD = gastroesophageal reflux disease; JH = jackhammer; LES = lower esophageal sphincter; LHM = laparoscopic Heller myotomy; OR = odds ratio; POEM = per oral esophageal myotomy; RCT = randomized controlled trial; SD = standard deviation; SED = spastic esophageal disorders; TBE = timed barium esophagram; WMD = weighted mean difference; WPR = weighted pooled rate;

*Clinical success was defined as Eckardt scores of ≤3 and/or improvement in severity of dysphagia based on achalasia disease-specific health-related quality of life questionnaire.*
Appendix 5: Additional References of Potential Interest


