White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/166987/

Version: Accepted Version

Article:
Mundre, P, Black, CJ, Mohammed, N et al. (2021) Efficacy of surgical or endoscopic treatment of idiopathic achalasia: a systematic review and network meta-analysis. The Lancet Gastroenterology & Hepatology, 6 (1). pp. 30-38. ISSN: 2468-1253

https://doi.org/10.1016/s2468-1253(20)30296-x

© 2020 Elsevier Ltd. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Accepted for publication 24th August 2020

TITLE PAGE

Title: Efficacy of Surgical or Endoscopic Treatment of Idiopathic Achalasia: Systematic Review and Network Meta-analysis.

Short title: Treatment Interventions for Idiopathic Achalasia: Network Meta-analysis.

Authors: Pradeep Mundre MBBS[1*], Christopher J. Black MBBS (Hons)[2,3*], Noor Mohammed MD[2], Professor Alexander C. Ford MD[2,3].

*Joint first author

[1]Bradford Teaching Hospitals, Bradford, UK

[2]Leeds Gastroenterology Institute, St. James's University Hospital, Leeds, UK.

[3]Leeds Institute of Medical Research at St. James's, University of Leeds, Leeds, UK.

Abbreviations:

ACG	American College of Gastroenterology
Botox	botulinum toxin
CI	confidence interval
LHM	laparoscopic Heller's myotomy
LOS	lower oesophageal sphincter
PD	pneumatic dilation
POEM	per-oral endoscopic myotomy
RCT	randomised controlled trial
RR	relative risk

Correspondence: Professor Alex Ford

Leeds Gastroenterology Institute

Room 125

4th Floor

Bexley Wing

St. James's University Hospital

Beckett Street

Leeds

United Kingdom

LS9 7TF

Email: alexf12399@yahoo.com

Telephone: +441132684963

ORCID ID: 0000-0001-6371-4359

Twitter: @alex_ford12399

Key words: achalasia; dysphagia; myotomy; dilation

Word count: 4054

SUMMARY

Background: Treatment of achalasia has evolved substantially over the last 20 years. Therapeutic options offered to patients vary, depending on access to both resources and expertise, and include pneumatic dilation (PD), laparoscopic Heller's myotomy (LHM), or per-oral endoscopic myotomy (POEM). Although there are head-to-head trials of these interventions, many of these are small and underpowered, so relative efficacy is unknown. We performed a systematic review and network meta-analysis to try to resolve this uncertainty.

Methods: We searched the Cochrane register of controlled trials, MEDLINE, EMBASE, and EMBASE Classic through 11th June 2020 for randomised controlled trials (RCTs) assessing efficacy of POEM, LHM, or PD, compared with each other in adults with idiopathic achalasia. Trials reported a dichotomous assessment of treatment failure or success after completion of therapy. We pooled data using a random effects model, and assessed heterogeneity between studies using the I^2 statistic. Risk of bias was examined for all studies. Efficacy was reported as a pooled relative risk (RR) of treatment failure, with a 95% confidence interval (CI), for each comparison tested, and ranked by therapy according to P-score.

Findings: We identified nine eligible RCTs, containing 911 participants. None were at low risk of bias. In total, 372 participants were randomised to LHM, 317 to PD, and 222 to POEM. Of the three strategies, POEM was ranked first (RR of failure of treatment = 0.33; 95% CI 0.15 to 0.71, P-score 0.89), followed by LHM (RR = 0.45; 95% CI 0.26 to 0.78, P-score 0.61). There was moderate heterogeneity between studies (I^2 = 61.5%). Both were superior to PD on direct and indirect comparison, but neither was significantly more effective than each other. There were no significant differences in perforation rates, need for re-intervention or surgery, or serious adverse events, but PD was less likely to lead to adverse events than POEM.

Interpretation: POEM and LHM should be the preferred treatments for idiopathic achalasia. PD performed worst in terms of treatment success, and therefore its role in management of patients with achalasia is less certain.

Funding: None.

RESEARCH IN CONTEXT

Evidence before this study

The incidence of achalasia is 2 to 3 per 100,000 people per year. Treatment options demonstrating long-term efficacy include per-oral endoscopic myotomy (POEM), laparoscopic Heller's myotomy (LHM), or pneumatic dilation (PD). Although previous randomised controlled trials (RCTs) have compared efficacy of all these interventions head-to-head, these trials are relatively small and probably underpowered, so the optimal treatment is unclear**.** Although trial-based meta-analyses have been conducted, they are limited to comparing efficacy of two treatments, and cannot encapsulate the range of treatment options completely, thus providing a rationale for this network meta-analysis. We searched the Cochrane register of controlled trials (CENTRAL), MEDLINE, EMBASE, EMBASE Classic, clinicaltrials.gov, and the International Clinical Trials Registry Platform to 11th June 2020 for RCTs assessing efficacy of POEM, LHM, or PD, compared with each other in adults with idiopathic achalasia. Risk of bias was examined for all studies. Efficacy was reported as a pooled relative risk (RR) of treatment failure, with a 95% confidence interval (CI), for each comparison tested, and ranked by therapy according to P-score.

Added value of this study

This network meta-analysis included only RCTs comparing efficacy of all these interventions for achalasia. The network allowed us to make direct and indirect comparisons between over 900 participants in nine RCTs. We now have a better understanding of the relative efficacy of POEM, LHM and PD for the treatment of the condition.

Implications of all the available evidence

Both POEM and LHM were superior to PD in this network meta-analysis of RCTs. Although LHM had comparable efficacy to POEM, POEM was ranked first. There were no significant differences between the three treatments in terms of likelihood of perforation, need for re-intervention or surgery, or serious adverse events. POEM is a reasonable first treatment for idiopathic achalasia, where there are facilities and expertise, although risk of gastro-oesophageal reflux disease (GORD) is higher. LHM with fundoplication is least likely to lead to GORD. Developing reliable factors to predict GORD after myotomy will help in directing patients for LHM with fundoplication. There should be a greater focus on training in POEM. However, PD is still a valid treatment option and should be considered, taking comorbidity, cost, and patient preference into account.

INTRODUCTION

Achalasia is a primary motility disorder of the oesophagus characterised by abnormal oesophageal peristalsis and incomplete relaxation of a hypertensive lower oesophageal sphincter (LOS). [1] The annual incidence is approximately 2 to 3 per 100,000 people, and prevalence is 10 per 100,000 people. [2-5] Patients have an increased incidence of oesophageal cancer, aspiration pneumonia, lower respiratory tract infections, and higher mortality. [3] Presentation is variable, but more than 90% of patients report dysphagia to both solids and liquids, and more than three-quarters experience regurgitation of undigested food. [6] Other commonly reported symptoms include nocturnal cough, aspiration, chest pain, heartburn, and weight loss. [7] The condition leads to substantial morbidity due to the above symptoms.

The diagnosis is made in a patient presenting with typical symptoms and one or more objective findings on oesophageal manometry, Barium swallow, or upper gastrointestinal endoscopy. [8] Diagnostic features on manometry are incomplete relaxation of the LOS, as reflected by increased integrative relaxation pressure and absence of normal peristalsis. [8] Once the diagnosis is made, possible interventions include pharmacotherapy with drugs such as nitrates or calcium channel antagonists, surgery in the form of laparoscopic Heller's myotomy (LHM), or endoscopic interventions, including botulinum toxin injection into the LOS, pneumatic dilation (PD), or per-oral endoscopic myotomy (POEM).

Pharmacotherapy is neither very effective nor long-lasting, [9] and compliance is often affected by side effects; current European guidelines do not recommend its use. [8] Although injection of botulinum toxin (Botox) is widely used, it is only effective in two-thirds of patients and benefits are temporary; [8] most patients relapse within 1 year and repeat treatments are ineffective. [10,11] Nevertheless, it is useful in patients who are unsuitable for more durable treatment options.

The mainstays of treatment, therefore, include LHM, POEM, or PD. In LHM, laparoscopic dissection of the anterior muscle fibres of the lower oesophagus and cardia is undertaken, usually combined with a fundoplication to prevent gastro-oesophageal reflux. This is a well-established treatment, but it is technically challenging to perform longer myotomies with LHM,[12] as it involves mobilising thoracic contents.[13,14] During POEM a submucosal tunnel is created from the mid-oesophagus to the gastric cardia and myotomy performed using electrocautery. The procedure was first described in animal models, [15] with the first human case performed in Japan. [16] Despite increased acceptance of POEM, influenced by outcomes from eastern centres, the required skills are relatively complex. In PD, dilation of the LOS is performed with a specially designed balloon, up to 40mm in diameter. However, this is an uncontrolled method of dilating the LOS, and perforation rates have been reported to be almost 2%.[17]

Choice of treatment is usually determined by availability, local expertise, and patient choice. Although there are several head-to-head randomised controlled trials (RCTs) comparing efficacy of LHM, POEM, and PD, the results are conflicting. [18-20] In addition, some of these trials are small, and likely underpowered to detect significant differences in efficacy. Trial-based meta-analyses have been conducted, [21,22] but they are limited to comparing efficacy of two treatments and cannot encapsulate the range of options completely. One such meta-analysis was abandoned by the authors, due to perceived variability in both PD techniques used and the definition of outcome measures. [23]

As a result, there is no clear evidence base to optimise treatment selection, and these interventions are in equipoise. We conducted a network meta-analysis of LHM, POEM, and PD in achalasia to estimate the relative efficacy of these interventions. This approach allows indirect, as well as direct, comparisons to be made across different RCTs, increasing the number of participants' data available for analysis. In addition, it provides a credible ranking system of the likely efficacy of different interventions. Knowledge of the most effective therapy overall may help inform future management guidelines and clinical decision-making.

METHODS

Search Strategy and Selection Criteria

We searched the Cochrane register of controlled trials (CENTRAL), MEDLINE (1946 to 11th June 2020), EMBASE (1946 to 11th June 2020), EMBASE Classic (1946 to 11th June 2020), clinicaltrials.gov, and the International Clinical Trials Registry Platform. We searched for key words in the title and abstract and under medical subject headings. We also performed a recursive search of bibliographies of all included studies, published guidelines on achalasia, and studies included in any previously published trial-based systematic review and meta-analysis. Studies published in abstract form were eligible for inclusion.

Only RCTs that compared the efficacy of any of POEM, LHM, or PD with each other in adult participants (≥18 years) with idiopathic achalasia were included (Table 1). We did not include sham-controlled trials of any of these interventions, due to potential differences in the sham procedure, depending on the active intervention used, meaning that these could not be treated as a single comparator, as would be the case in a network meta-analysis of placebo-controlled drug trials.[24-27]We also did not include trials that compared the above interventions against Botox therapy, as this would potentially introduce a selection bias for patients, due to the higher likelihood of the inclusion of comorbid patients, who would not be considered fit enough for either POEM or LHM, in such trials. In fact, the recent American College of Gastroenterology (ACG) guideline for management of achalasia recommends that, due to its short-lived benefits, Botox be reserved for those who cannot undergo the above definitive therapies.[28] Trials had to recruit patients with a diagnosis of achalasia based on clinical grounds, along with typical findings on at least one of manometry, radiology, or upper gastrointestinal endoscopy. The minimum duration of follow-up was 1 year. Trials had to report a dichotomous measure of treatment success or failure, according to incomplete or poor symptom control, need for retreatment, or symptom relapse on measurable outcome scores, (i.e.

Eckardt score >3 or another achalasia-specific symptom score)[29,30] during follow-up. Ethical approval for this evidence synthesis was not required.

Two investigators (PM and NM) conducted the literature search independently from each other. The search strategy is provided in the Supplementary Materials (pages 1 to 4). There were no language restrictions. Two investigators (PM and NM) evaluated all abstracts identified by the search for eligibility, again independently from each other. We obtained all potentially relevant papers, and evaluated them in more detail, to assess eligibility independently, according to the pre-defined criteria. We translated foreign language papers, where required. Any disagreements in eligibility were resolved by discussion between the investigators (PM and NM). Where disagreements arose, we asked a third person (ACF) to arbitrate.

Data Analysis

The primary outcome was efficacy, in terms of a dichotomous measure of treatment success or failure, after a minimum of 12 months of follow-up. Treatment failure was defined as incomplete or poor symptom control, need for retreatment, or symptom relapse on measurable outcome scores, (i.e. Eckardt score >3 or another achalasia-specific symptom score) during follow-up. Details of the criteria used in individual trials are provided in Table 2. Secondary outcomes included occurrence of perforation, adverse events, serious adverse events (including death), need for re-intervention, need for surgery as a result of complications, development of gastro-oesophageal reflux (either according to symptoms or confirmed on ambulatory pH monitoring), or erosive oesophagitis (as seen at upper gastrointestinal endoscopy).

Two investigators (PM and ACF) extracted all data independently onto a Microsoft Excel spreadsheet (XP professional edition; Microsoft Corp, Redmond, WA, USA) as dichotomous outcomes (treatment success or failure). For all included studies, we also extracted the following data for each trial, where available: country of origin, number of centres, duration of follow-up, and

primary outcome measure used to define treatment success or failure. Data were extracted as intention-to-treat analyses, with all dropouts assumed to be treatment failures (*i.e.* symptomatic at final point of follow-up), wherever trial reporting allowed this. However, due to the interventions involved, and the duration of follow-up, we also performed a per protocol analysis to assess the robustness of our findings, with only patients receiving the intervention to which they were allocated, and successfully followed up, considered in the analysis.

We performed risk of bias assessment at the study level. Two investigators (PM and ACF) assessed this independently using the Cochrane risk of bias tool RoB 2. [31] Disagreements were resolved by discussion. We recorded the methods used to generate the randomisation schedule and conceal treatment allocation. Due to the nature of the interventions studied, blinding could not be implemented for either participants or personnel. However, this was possible for the outcome's assessors, so we extracted this in the assessment, where reported. We also assessed for evidence of incomplete outcomes data or selective reporting of outcomes.

A network meta-analysis was performed using the frequentist model, with the statistical package "netmeta" (version 0.9-0, https://cran.r-project.org/web/packages/netmeta/index.html) in R (version 3.4.2). We reported the network meta-analysis according to the PRISMA extension statement for network meta-analyses. [32] Network meta-analysis results usually give a more precise estimate, compared with results from standard, pairwise analyses, [33] and can rank treatments to inform clinical decisions. [34]

We examined the symmetry and geometry of the evidence by producing a network plot with node and connection size corresponding to the number of study subjects and number of studies, respectively. We aimed to produce comparison-adjusted funnel plots to explore publication bias or other small study effects, where sufficient trials (≥10) existed. [35] We produced a pooled relative risk (RR) with a 95% confidence interval (CI) to summarise the efficacy of each of the interventions tested, using a random effects model as a conservative estimate. We used a RR of treatment failure at

the final point of follow-up; where the RR is less than 1 and the 95% CI does not cross 1, there is a significant benefit of one intervention over another.

We assessed global statistical heterogeneity across all comparisons using the I^2 measure from the "netmeta" statistical package. The I^2 value ranges between 0% and 100%. Values of 25% to 49%, 50% to 74%, and ≥75% are considered low, moderate, and high levels of heterogeneity, respectively. [36] We ranked the interventions according to their P-score, which is a value between 0 and 1. P-scores are based solely on the point estimates and standard errors of the network estimates, and measure the mean extent of certainty that one intervention is better than another, averaged over all competing interventions. [37] Higher scores indicate a greater probability of the intervention being ranked as best, [37] but the magnitude of the P-score should be considered, as well as the rank. As the mean value of the P-score is always 0.5, individual treatments that cluster around this value are likely to be of similar effectiveness. However, when interpreting the results, it is also important to take the RR and corresponding 95% CI for each comparison into account, rather than relying on rankings alone. [38] In our primary analysis, we pooled data for the risk of being symptomatic at the final point of follow-up in each study for all included RCTs using an intention-to-treat analysis.

Role of the Funding Source

There was no funding source for this study. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

RESULTS

The search strategy generated 1044 citations. After review of titles and abstracts, we retrieved 35 articles for further assessment (Figure 1). Some articles were duplicates and reported outcomes from the same cohort of patients at different points of follow-up. In this situation, the article reporting the primary end point was used, but we examined all other publications, to ensure there were no missing data in the primary publication. One of the included studies was a subsequent analysis of an earlier trial, but reported more of the data of interest.[39,40] In total, 26 articles were excluded, leaving nine eligible RCTs, [18-20,39,41-45] containing 911 patients. Seven of these trials were fully published, and two were in abstract form only. [41,42]

Detailed characteristics of individual RCTs, including the comparisons made, are provided in Table 2. Technical aspects of each intervention in each trial are reported in Supplementary Table 1 (page 5). Risk of bias for all included studies is provided in Supplementary Table 2 (page 6) and Supplementary Figure 1 (page 14). All nine RCTs were high risk of bias due to the impossibility of blinding. Five stated the method of randomisation, four the method of concealment of allocation, and four reported an intention-to-treat analysis. None of the trials had evidence of selective reporting of outcomes.

Efficacy

Intention-to-treat Analysis

All nine RCTs provided dichotomous data for likelihood of failure of therapy at between 1 and 3 years. [18-20,39,41-45] In total, 372 participants were randomised to LHM, 317 to PD, and 222 to POEM. The network plot is provided in Figure 2. When data were pooled, there was moderate heterogeneity between studies ($I^2 = 61.5\%$). There were too few studies to assess for publication bias, or other small study effects. Of the three strategies, POEM was ranked first (RR of failure of

treatment = 0.33; 95% CI 0.15 to 0.71, P-score 0.89) (Figure 3), followed by LHM (RR = 0.45; 95% CI 0.26 to 0.78, P-score 0.61). Both were more effective than PD on direct and indirect comparison (Table 3), but neither was significantly more effective than each other.

Per Protocol Analysis

All nine RCTs provided dichotomous data for likelihood of failure of therapy at the last point of follow-up according to a per protocol analysis. [18-20,39,41-45] In this analysis, there were data available for 797 participants, of whom 328 participants were randomised to LHM, 264 to PD, and 205 to POEM. When data were pooled, there was moderate heterogeneity (I^2 = 56.5%). Once again, POEM was ranked first (RR of failure of treatment = 0.29; 95% CI 0.10 to 0.80, P-score 0.89) (Figure 4), and LHM second (RR = 0.42; 95% CI 0.20 to 0.90, P-score 0.60). Both were more effective than PD on indirect comparison (Table 4), and POEM was more effective than PD on direct comparison, but neither was significantly more effective than each other.

Rates of Perforation, Need for Re-intervention or Surgery, Adverse Events, or Gastro-oesophageal Reflux.

Reporting of these endpoints varied among the nine RCTs. Individual trials contributing data to each analysis, number of patients, and summary effects from the network meta-analysis for each are provided in Table 5 and in the Supplementary Materials (pages 7 to 13). There were insufficient trials reporting deaths to perform an analysis. There were no significant differences on either indirect or direct comparison for any of the other secondary endpoints of interest, with the exception that PD was significantly less likely to lead to adverse events than POEM on both indirect and direct comparison. LHM was the intervention least likely to lead to gastro-oesophageal reflux, and PD the least likely to lead to erosive oesophagitis, whereas POEM was least likely to lead to perforation, need for re-intervention, need for surgery, or serious adverse events.

DISCUSSION

This systematic review and network meta-analysis of RCTs of surgical or endoscopic interventions for achalasia has demonstrated that POEM is the best ranked treatment, in terms of efficacy, and is likely to be superior to PD, based on very low-quality evidence. It was also the least likely intervention to lead to perforation, need for re-intervention, need for surgery, or serious adverse events, although these endpoints did not reach statistical significance. However, it was ranked last for development of oesophagitis. LHM was ranked second, with comparable efficacy to POEM, and was also likely to be more effective than PD. It was the intervention least likely to lead to development of gastro-oesophageal reflux. PD was ranked last for efficacy, although it was significantly less likely to lead to adverse events than POEM. However, most of these adverse events were minor. The results of this network meta-analysis therefore suggest that the mainstay of therapy for achalasia should be either POEM or LHM. PD performed worst in terms of treatment success and therefore it has a limited role in the treatment of achalasia, but should still be considered, taking comorbidity, cost, and patient preference into account. Although POEM was ranked first, developing reliable factors to predict GORD after myotomy may help select the right patient group for this treatment. Until then, LHM with fundoplication may still be preferable for some patients. The choice of therapy should, therefore, be guided by shared decision-making, where patients are provided with the risks and benefits of both modalities.

To our knowledge, this is the first network meta-analysis including only RCTs comparing all of these interventions for achalasia. The network allowed us to make direct and indirect comparisons between over 900 participants in these nine RCTs. The trials themselves took place in a wide variety of settings, and countries meaning the results are likely to be generalisable to many patients with achalasia. We used an intention-to-treat analysis, with all trial dropouts assumed to be treatment failures. We extracted data during longer term follow-up, between a minimum of 1 year and a

maximum of 3 years. We also conducted a per protocol analysis to assess the robustness of the results we observed.

Weaknesses include the fact that there were differences between individual trials, in terms of outcome measures and dilation regimens used in the PD arms of the studies, as well as criteria used to define perforation in trials of LHM and POEM. Wherever possible, we did not classify either mucosal tears occurring at the time of LHM, or a mucosal tear that occurred separate to the initial submucosal entry point during POEM, which were repaired at the time of procedure, as perforations unless they modified the post-treatment course (i.e. led to a prolonged admission, use of antibiotics, or any other intervention, such as drains, etc. or conversion to an open procedure). However, there was a lack of clarity on this issue in some older trials. Where this was not possible, we extracted all data, including mucosal tears, if reported by the authors as perforations. Excluding the trial by Boeckxstaens *et al.*,[43] there were only seven perforations with LHM, and one with POEM, and even accounting for this variation in reporting, there were no significant differences between individual treatments in terms of perforation in the network. There was moderate heterogeneity between studies in our main analyses. Confidence intervals around the estimates of efficacy were wide, presumably because of the relatively small number of patients, in total, assigned to each of the interventions. The smallest number were assigned to POEM. In addition, all of the included RCTs were at high risk of bias, due to the nature of the intervention, which meant that blinding of participants or investigators was not possible, although one trial specifically mentioned that outcome assessors were different from the person undertaking the intervention. [39] This means our conclusions are based on very low quality evidence. Finally, although the ACG guideline for management of achalasia suggests that all three treatments can be used in type I or II achalasia, and that POEM may be better for type III,[28] analysis by subtype of achalasia was not possible due to incomplete reporting of outcomes among patients according to the Chicago classification.

Assumption of transitivity is fundamental to network meta-analysis, as indirect comparisons are built on the assumption that any patient included in the network could have, theoretically, been recruited to any of the trials and assigned to any of the treatments. [33] This was why RCTs of Botox were not considered, as patients entering these trials are unlikely to have been suitable for POEM or LHM because of comorbidities or risk. All included studies, except one, [41] stated that the diagnosis was made based on clinical symptoms and patients underwent upper gastrointestinal endoscopy and manometry to facilitate the diagnosis. Although not stated explicitly in this trial, [41] we believe it is likely that manometry and upper gastrointestinal endoscopy were carried out as part of the diagnostic work-up, based on the reported outcomes. In our analysis all studies, except two, [44,45] excluded patients with previous endoscopic intervention. In one study, [44] prior Botox therapy was allowed >3 months before randomisation, whereas in the other, [45] prior dilation was allowed, as this study compared LHM with POEM. In the latter trial, approximately 25% of patients in each group had undergone prior dilation. This small difference in sample population is unlikely to have had an impact on the overall results of our meta-analysis. In fact, excluding such patients is likely to increase the overall efficacy of POEM and LHM, compared with PD. The main technical aspects of POEM and LHM were similar across all included studies, as detailed in Supplementary Table 1. Although there were differences in dilation regimens between the studies, seven of the nine studies used at least two predefined serial dilations as the primary treatment and were similar in terms of their principles (see Supplementary Table 1). In the other two studies, [18,41] a single dilation regimen was used as the primary treatment. We do not believe that these two studies will have affected the overall results of the network meta-analysis, as one reported equivalent efficacy between PD and LHM, [18] and the number of patients in the other study was small. [41] The dilation regimens used in all studies seemed to reflect, pragmatically, what was feasible in routine clinical practice.[46] Further dilations over and above these regimens would reflect poor efficacy, and this would be consistent with treatment failure, as defined in the individual trials.

One previous failed attempt at meta-analysis between LHM and PD, [23] suggested this variability in PD regimens as one of the reasons for failure. Based on the argument above, we do not believe that this variability would influence the conclusions of our study. Again, variability in outcome measures was argued as one of the reasons for failed meta-analysis, but we only included studies with clearly defined outcomes using dichotomous measures of improvement or non-improvement. Five of the nine studies used the Eckardt score, [41-45] and the outcome measures in the other four studies followed similar principles to this scoring system. Due to a lack of blinding, the overall direction of any bias would seem to favour POEM or LHM, in studies that compared either of these to PD, [18-20,39,41,43,44] but in studies of POEM versus LHM this is less predictable. [42,45] However, due to the nature of the outcome measures used, including the Eckardt score and the need for re-intervention, the potential influence of this bias is low. The Eckardt score has been previously validated as a measure of achalasia severity, and there is modest correlation with physiological data; [29] hence it is unlikely to be subjective. However, it has not been validated as a measure of treatment success in achalasia. All the above may create some imprecision in rankings and should be taken into consideration when interpreting results. Despite these limitations, we believe our network meta-analysis provides a better understanding of the comparative efficacy of POEM, LHM, and PD in patients with achalasia.

We are aware of only one previously published network meta-analysis on interventions for achalasia. [47] However, it appears this study missed some eligible RCTs, [41,42,44,45] and it did not include any trials of POEM. In fact, most of the included studies were observational in nature, meaning its contribution to the evidence base is likely to be minimal. Some of the previously conducted pairwise meta-analyses did not include RCTs, but instead used data from retrospective or prospective observational studies. [22,48] Others included only RCTs, but may have missed eligible trials. [21,49] A previously published guideline for the management of achalasia recommended POEM,

LHM, or repetitive graded PD as being of comparable efficacy, [8] as they were unable to be ranked based on the current evidence. Our study therefore helps to address this key question.

In summary, this network meta-analysis demonstrates that both POEM and LHM were superior to PD for the treatment of achalasia, although neither were superior to each other. POEM was ranked first for efficacy, but there were no significant differences between the three treatments in terms of likelihood of perforation, need for re-intervention or surgery, or serious adverse events. POEM or LHM should be preferred for the treatment of achalasia, depending on local expertise, patient choice, and suitability for intervention, although PD should still be considered, taking comorbidity, cost, and patient preference into account. However, POEM and LHM are potentially expensive interventions.[50,51] Future studies should therefore consider in-built health economic evaluations of these treatments.

ACKNOWLEDGEMENTS

None.

AUTHOR CONTRIBUTIONS

ACF and PM conceived and drafted the study. PM, CJB, ACF, and NM analysed, and interpreted the data. PM, NM, and ACF drafted the manuscript. All authors have approved the final draft of the manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

DECLARATION OF INTERESTS

Pradeep Mundre: none. Christopher J. Black: none. Noor Mohammed: none. Alexander C. Ford: none.

ETHICS COMMITTEE APPROVAL

Not required.

REFERENCES

1. Hurst AF, Rake GW. Achalasia of the Cardia: So-Called Cardiospasm. *QJM: An International Journal of Medicine* 1930; **os-23**(92): 491-508.

2. Sadowski DC, Ackah F, Jiang B, Svenson LW. Achalasia: incidence, prevalence and survival. A population-based study. *Neurogastroenterol Motil* 2010; **22**(9): e256-e61.

3. Harvey PR, Thomas T, Chandan JS, et al. Incidence, morbidity and mortality of patients with achalasia in England: findings from a study of nationwide hospital and primary care data. *Gut* 2019; **68**(5): 790-5.

4. van Hoeij FB, Ponds FA, Smout AJ, Bredenoord AJ. Incidence and costs of achalasia in The Netherlands. *Neurogastroenterol Motil* 2018; **30**(2): e13195.

5. Duffield JA, Hamer PW, Heddle R, Holloway RH, Myers JC, Thompson SK. Incidence of Achalasia in South Australia Based on Esophageal Manometry Findings. *Clin Gastroenterol Hepatol* 2017; **15**(3): 360-5.

6. Vantrappen G, Hellemans J, Deloof W, Valembois P, Vandenbroucke J. Treatment of achalasia with pneumatic dilatations. *Gut* 1971; **12**(4): 268-75.

7. Boeckxstaens GE, Zaninotto G, Richter JE. Achalasia. *The Lancet* 2014; **383**(9911): 83-93.

8. Oude Nijhuis RAB, Zaninotto G, Roman S, et al. European guidelines on achalasia: United European Gastroenterology and European Society of Neurogastroenterology and Motility recommendations. *United European Gastroenterol J* 2020; **8**(1): 13-33.

9. Wen ZH, Gardener E, Wang YP. Nitrates for achalasia. *Cochrane Database Syst Rev* 2004; **2004**(1): Cd002299.

10. Storr M, Born P, Frimberger E, et al. Treatment of achalasia: the short-term response to botulinum toxin injection seems to be independent of any kind of pretreatment. *BMC gastroenterology* 2002; **2**.

11. Kahrilas PJ, Pandolfino JE. Treatments for achalasia in 2017: how to choose among them. *Curr Opin Gastroenterol* 2017; **33**(4): 270-6.

12. Park CH, Jung DH, Kim DH, et al. Comparative efficacy of per-oral endoscopic myotomy and Heller myotomy in patients with achalasia: a meta-analysis. *Gastrointest Endosc* 2019; **90**(4): 546-58.e3.

13. Schlottmann F, Allaix ME, Patti MG. Laparoscopic Heller Myotomy for Achalasia Technical Aspects. *Am Surg* 2018; **84**(4): 477-80.

14. Oelschlager BK. Surgical options for treatment of esophageal motility disorders. *Gastroenterol Hepatol (N Y)* 2007; **3**(9): 687-9.

15. Pasricha P, Hawari R, Ahmed I, et al. Submucosal Endoscopic Esophageal Myotomy: A Novel Experimental Approach for the Treatment of Achalasia. *Endoscopy* 2007; **39**: 761-4.

16. Inoue H, Minami H, Satodate H, Kudo S-E. First Clinical Experience of Submucosal Endoscopic Esophageal Myotomy for Esophageal Achalasia with No Skin Incision. *Gastrointestinal Endoscopy* 2009; **69**(5): AB122.

17. Harvey PR, Coupland B, Mytton J, Evison F, Patel P, Trudgill NJ. Outcomes of pneumatic dilatation and Heller's myotomy for achalasia in England between 2005 and 2016. *Gut* 2019; **68**(7): 1146-51.

18. Borges AA, Lemme EM, Abrahao LJ, et al. Pneumatic dilation versus laparoscopic Heller myotomy for the treatment of achalasia: variables related to a good response. *Diseases of the esophagus : official journal of the international society for diseases of the esophagus* 2014; **27**(1): 18-23.

19. Chrystoja CC, Darling GE, Diamant NE, et al. Achalasia-Specific Quality of Life After Pneumatic Dilation or Laparoscopic Heller Myotomy With Partial Fundoplication: a Multicenter, Randomized Clinical Trial. *American journal of gastroenterology* 2016; **111**(11): 1536-45.

20. Hamdy E, El Nakeeb A, El Hanfy E, et al. Comparative Study Between Laparoscopic Heller Myotomy Versus Pneumatic Dilatation for Treatment of Early Achalasia: a Prospective Randomized Study. *Journal of laparoendoscopic & advanced surgical techniques Part a* 2015; **25**(6): 460-4.

21. Bonifacio P, de Moura DTH, Bernardo WM, et al. Pneumatic dilation versus laparoscopic Heller's myotomy in the treatment of achalasia: systematic review and meta-analysis based on randomized controlled trials. *Diseases of the Esophagus* 2019; **32**(2): 01.

22. Marano L, Pallabazzer G, Solito B, et al. Surgery or Peroral Esophageal Myotomy for Achalasia: A Systematic Review and Meta-Analysis. *Medicine (Baltimore)* 2016; **95**(10): e3001.

23. de Heer J, Desai M, Boeckxstaens G, et al. Pneumatic balloon dilatation versus laparoscopic Heller myotomy for achalasia: a failed attempt at meta-analysis. *Surg Endosc* 2020.

24. Black CJ, Yuan Y, Selinger CP, et al. Efficacy of soluble fibre, antispasmodic drugs, and gut-brain neuromodulators in irritable bowel syndrome: a systematic review and network meta-analysis. *Lancet Gastroenterol Hepatol* 2020; **5**(2): 117-31.

25. Luthra P, Camilleri M, Burr NE, Quigley EMM, Black CJ, Ford AC. Efficacy of drugs in chronic idiopathic constipation: a systematic review and network meta-analysis. *Lancet Gastroenterol Hepatol* 2019; **4**(11): 831-44.

26. Black CJ, Burr NE, Camilleri M, et al. Efficacy of pharmacological therapies in patients with IBS with diarrhoea or mixed stool pattern: systematic review and network meta-analysis. *Gut* 2020; **69**(1): 74-82.

27. Black CJ, Burr NE, Quigley EMM, Moayyedi P, Houghton LA, Ford AC. Efficacy of Secretagogues in Patients With Irritable Bowel Syndrome With Constipation: Systematic Review and Network Meta-analysis. *Gastroenterology* 2018; **155**(6): 1753-63.

28. Vaezi MF, Pandolfino JE, Yadlapati RH, Greer KB, Kavitt RT. ACG Clinical Guidelines: Achalasia. *Am J Gastroenterol* 2020.

29. Taft TH, Carlson DA, Triggs J, et al. Evaluating the reliability and construct validity of the Eckardt symptom score as a measure of achalasia severity. *Neurogastroenterol Motil* 2018; **30**(6): e13287.

30. Gockel I, Junginger T. The value of scoring achalasia: a comparison of current systems and the impact on treatment--the surgeon's viewpoint. *Am Surg* 2007; **73**(4): 327-31.

31. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj* 2019; **366**: l4898.

32. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015; **162**(11): 777-84.

33. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014; **9**(7): e99682.

34. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011; **64**(2): 163-71.

35. Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011; **343**: d4002.

36. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**(7414): 557.

37. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Medical Research Methodology* 2015; **15**(1): 58.

38. Morton SC MM, O'Connor E, Lee CS, Booth M, Vandermeer BW, Snowden JM, D'Anci KE, Fu R, Gartlehner G, Wang Z, Steele DW. Quantitative Synthesis—An Update. Methods Guide for

Comparative Effectiveness Reviews. (Prepared by the Scientific Resource Center under Contract No. 290-2012-0004-C). AHRQ Publication No. 18-EHC007-EF. Rockville, MD. *Agency for health Care Research and Quality (US)* 2018.

39. Persson J, Johnsson E, Kostic S, Lundell L, Smedh U. Treatment of achalasia with laparoscopic myotomy or pneumatic dilatation: long-term results of a prospective, randomized study. *World journal of surgery* 2015; **39**(3): 713-20.

40. Kostic S, Kjellin A, Ruth M, et al. Pneumatic dilatation or laparoscopic cardiomyotomy in the management of newly diagnosed idiopathic achalasia. Results of a randomized controlled trial. *World journal of surgery* 2007; **31**(3): 470-8.

41. Linghu E, Li H. Randomized study comparing peroral endoscopic myotomy, botulinum toxin injection and balloon dilation for achalasia: one-year follow-up ACG international award. *American journal of gastroenterology* 2013; **108**: S602.

42. Moura ET, Farias GF, Coutinho LM, et al. A randomized controlled trial comparing peroral endoscopic myotomy (poem) versus laparoscopic heller myotomy with fundoplication in the treatment of achalasia. *Gastrointestinal Endoscopy* 2019; **89**(6): AB84.

43. Boeckxstaens GE, Annese V, des Varannes SB, et al. Pneumatic dilation versus laparoscopic Heller's myotomy for idiopathic achalasia. *New England journal of medicine* 2011; **364**(19): 1807-16.

44. Ponds FA, Fockens P, Lei A, et al. Effect of Peroral Endoscopic Myotomy vs Pneumatic Dilation on Symptom Severity and Treatment Outcomes Among Treatment-Naive Patients With Achalasia: A Randomized Clinical Trial. *JAMA* 2019; **322**(2): 134-44.

45. Werner YB, Hakanson B, Martinek J, et al. Endoscopic or Surgical Myotomy in Patients with Idiopathic Achalasia. *New England journal of medicine* 2019; **381**(23): 2219-29.

46. Sami SS, Haboubi HN, Ang Y, et al. UK guidelines on oesophageal dilatation in clinical practice. *Gut* 2018; **67**(6): 1000-23.

47. Aiolfi A, Bona D, Riva CG, et al. Systematic Review and Bayesian Network Meta-Analysis Comparing Laparoscopic Heller Myotomy, Pneumatic Dilatation, and Peroral Endoscopic Myotomy for Esophageal Achalasia. *J Laparoendosc Adv Surg Tech A* 2020; **30**(2): 147-55.

48. Park CH, Jung DH, Kim DH, et al. Comparative efficacy of per-oral endoscopic myotomy and Heller myotomy in patients with achalasia: a meta-analysis. *Gastrointestinal Endoscopy* 2019; **90**(4): 546-58.e3.

49. Yaghoobi M, Mayrand S, Martel M, Roshan-Afshar I, Bijarchi R, Barkun A. Laparoscopic Heller's myotomy versus pneumatic dilation in the treatment of idiopathic achalasia: a meta-analysis of randomized, controlled trials. *Gastrointestinal Endoscopy* 2013; **78**(3): 468-75.

50. Moonen A, Busch O, Costantini M, et al. Economic evaluation of the randomized European Achalasia trial comparing pneumodilation with Laparoscopic Heller myotomy. *Neurogastroenterol Motil* 2017; **29**(11).

51. Miller HJ, Neupane R, Fayezizadeh M, Majumder A, Marks JM. POEM is a cost-effective procedure: cost-utility analysis of endoscopic and surgical treatment options in the management of achalasia. *Surg Endosc* 2017; **31**(4): 1636-42.

Figure 1. Flow Diagram of Assessment of Studies Identified in the Systematic Review.

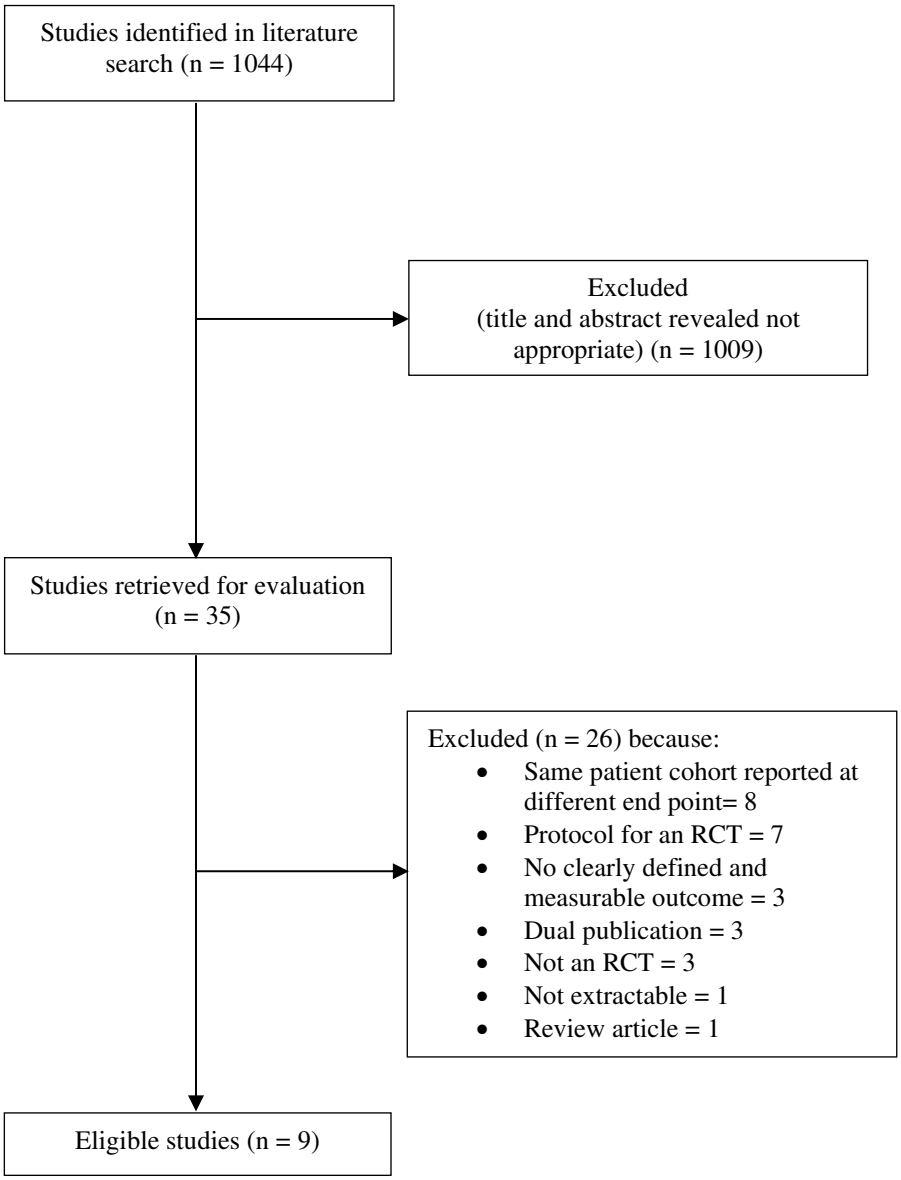
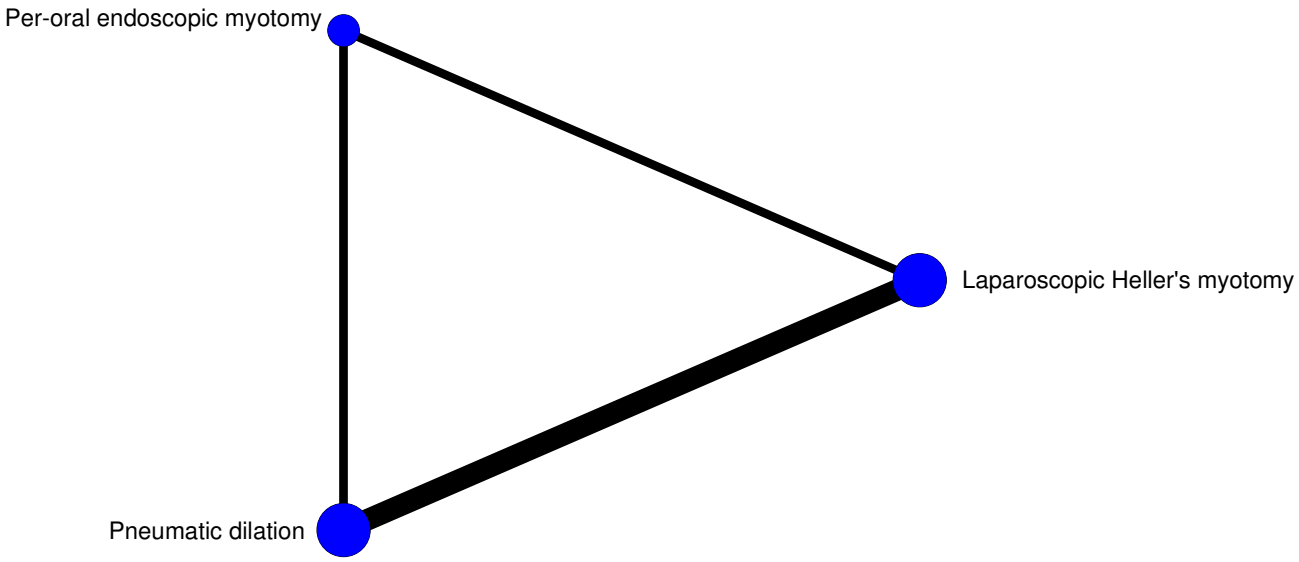
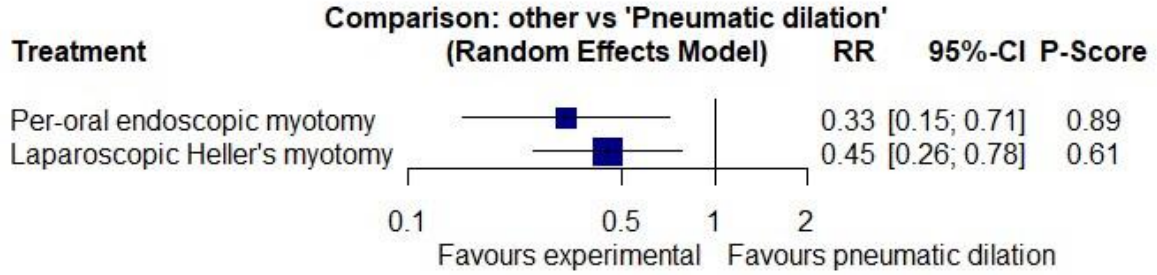


Figure 2. Network Plot for Likelihood of Failure of Therapy According to Intention-to-treat Analysis at the Last Point of Follow-up.



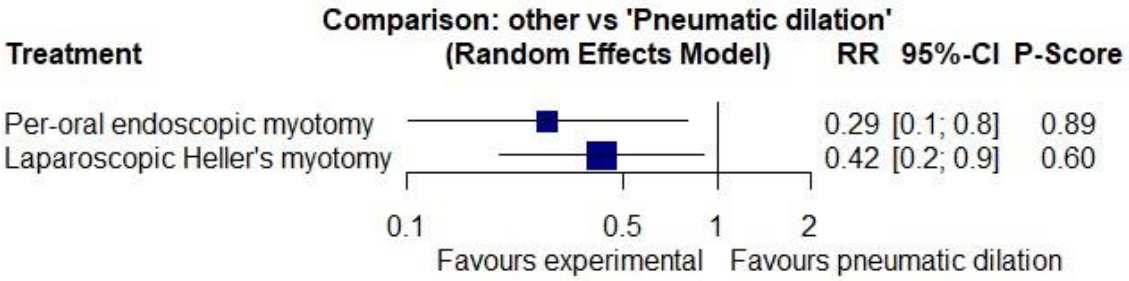
Note: Nine separate studies, containing 911 participants. Circle (node) size is proportional to the number of study participants assigned to receive each intervention. The line width (connection size) corresponds to the number of studies comparing the individual treatments.

Figure 3. Forest Plot for Likelihood of Failure of Therapy According to Intention-to-treat Analysis at the Last Point of Follow-up.



Note: The P-score is the probability of each treatment being ranked as best in the network.

Figure 4. Forest Plot for Likelihood of Failure of Therapy According to Per Protocol Analysis at the Last Point of Follow-up.



Note: The P-score is the probability of each treatment being ranked as best in the network.