

Appendix 3: Evidence to decision tables

QUESTION 1

Should POEM vs. Heller be used for patients with achalasia?	
POPULATION:	achalasia in adults and children
INTERVENTION:	POEM
COMPARISON:	Heller
MAIN OUTCOMES:	Symptom resolution; postoperative relaxation pressure of lower esophageal sphincter; quality of life; grades B-D reflux esophagitis; postoperative pain; serious adverse events; return to OR for postoperative complications
PERSPECTIVE:	Patient/surgeon perspective
SETTING	International
CONFLICT OF INTERESTS:	No funding was provided for this Guideline. Individual authors all provided disclosures as listed in separate appendix within the guideline

ASSESSMENT

Desirable Effects																																									
How substantial are the desirable anticipated effects?																																									
JUDGEMENT	RESEARCH EVIDENCE					ADDITIONAL CONSIDERATIONS																																			
<ul style="list-style-type: none"> ○ Trivial ● Small ○ Moderate ○ Large ○ Varies ○ Don't know 	One low risk of bias recent randomized control trial on POEM versus laparoscopic Heller myotomy with Dor fundoplication and fourteen high risk of bias observational studies on POEM versus laparoscopic Heller myotomy were used to inform the panel's decision.					For outcomes favoring POEM, the magnitude of effect was judged to be trivial except for postoperative complications and pain which were considered important enough to judge the overall desirable effects as small. The current data are limited by the short term follow up period (up to 2 years). The panel considered these data as a proxy for long-term outcomes (5-year, 10-year, and 15-year follow-up). This is the best available proxy based on the panel's opinion. However, the panel expressed an interest in following up on any longer-term data in the future, to abolish the need for this proxy .																																			
	<table border="1"> <thead> <tr> <th>Outcomes</th> <th>Relative effect (95% CI)</th> <th colspan="3">Anticipated absolute effects* (95% CI)</th> <th>Certainty of the evidence (GRADE)</th> <th>Importance</th> </tr> <tr> <td></td> <td></td> <th>With POEM</th> <th>With Heller</th> <th>Difference</th> <td></td> <td></td> </tr> </thead> <tbody> <tr> <td>Success/Symptom resolution by Eckardt Score for dysphagia - 2 years № of participants: 221 (1 RCT)</td> <td>RR 1.02 (0.90 to 1.15)</td> <td>Study population 83.3% (73.5 to 93.9)</td> <td>81.7%</td> <td>1.6% more (8.2 fewer to 12.2 more)</td> <td>⊕○○○ VERY LOW ^{a,b}</td> <td>CRITICAL</td> </tr> <tr> <td>Post-operative relaxation pressure of LES – 2 years (mmHg)</td> <td>-</td> <td>Mean 11.3 mm HG</td> <td>Mean 11.5 mm HG</td> <td>MD 0.75 mm HG lower</td> <td>⊕○○○ VERY LOW ^{b,c}</td> <td>IMPORTANT</td> </tr> </tbody> </table>	Outcomes	Relative effect (95% CI)	Anticipated absolute effects* (95% CI)			Certainty of the evidence (GRADE)	Importance			With POEM	With Heller	Difference			Success/Symptom resolution by Eckardt Score for dysphagia - 2 years № of participants: 221 (1 RCT)	RR 1.02 (0.90 to 1.15)	Study population 83.3% (73.5 to 93.9)	81.7%	1.6% more (8.2 fewer to 12.2 more)	⊕○○○ VERY LOW ^{a,b}	CRITICAL	Post-operative relaxation pressure of LES – 2 years (mmHg)	-	Mean 11.3 mm HG	Mean 11.5 mm HG	MD 0.75 mm HG lower	⊕○○○ VERY LOW ^{b,c}	IMPORTANT												
Outcomes	Relative effect (95% CI)	Anticipated absolute effects* (95% CI)			Certainty of the evidence (GRADE)	Importance																																			
		With POEM	With Heller	Difference																																					
Success/Symptom resolution by Eckardt Score for dysphagia - 2 years № of participants: 221 (1 RCT)	RR 1.02 (0.90 to 1.15)	Study population 83.3% (73.5 to 93.9)	81.7%	1.6% more (8.2 fewer to 12.2 more)	⊕○○○ VERY LOW ^{a,b}	CRITICAL																																			
Post-operative relaxation pressure of LES – 2 years (mmHg)	-	Mean 11.3 mm HG	Mean 11.5 mm HG	MD 0.75 mm HG lower	⊕○○○ VERY LOW ^{b,c}	IMPORTANT																																			
	<table border="1"> <tbody> <tr> <td>№ of participants: 148 (1 RCT)</td> <td></td> <td></td> <td></td> <td>(2.26 lower to 0.76 higher)</td> <td></td> <td></td> </tr> <tr> <td>Quality of Life Improvement - 2 years (GIQOL, range 0 - 144) № of participants: 202 (1 RCT)</td> <td>-</td> <td>Mean improvement in GIQOL of 28.1 points</td> <td>Mean improvement in GIQOL of 24.5 points</td> <td>MD 0.14 higher (4.01 lower to 4.28 higher)</td> <td>⊕⊕○○ LOW ^b</td> <td>CRITICAL</td> </tr> <tr> <td>Presence of post-operative pain (≤ 6 months) № of participants: 269 (3 observational studies)</td> <td>RR 0.88 (0.60 to 1.29)</td> <td>Study population 18.9% (12.9 to 27.8)</td> <td>21.5%</td> <td>2.6% fewer (8.6 fewer to 6.2 more)</td> <td>⊕○○○ VERY LOW ^{b,e,f}</td> <td>IMPORTANT</td> </tr> <tr> <td>Serious adverse events [interventions, prolonged hospitalization, ICU admission, or death] № of participants: 221 (1 RCT)</td> <td>RR 0.36 (0.10 to 1.34)</td> <td>Study population 2.6% (0.7 to 9.8)</td> <td>7.3%</td> <td>4.7% fewer (6.6 fewer to 2.5 more)</td> <td>⊕⊕○○ LOW ^b</td> <td>CRITICAL</td> </tr> <tr> <td>Return to OR [for postoperative complications] № of participants: 469 (9 observational studies)</td> <td>RR 0.79 (0.28 to 2.22)</td> <td>Study population 1.8% (0.7 to 5.2)</td> <td>2.3%</td> <td>0.5% fewer (1.7 fewer to 2.9 more)</td> <td>⊕○○○ VERY LOW ^{b,g,h}</td> <td>CRITICAL</td> </tr> </tbody> </table>	№ of participants: 148 (1 RCT)				(2.26 lower to 0.76 higher)			Quality of Life Improvement - 2 years (GIQOL, range 0 - 144) № of participants: 202 (1 RCT)	-	Mean improvement in GIQOL of 28.1 points	Mean improvement in GIQOL of 24.5 points	MD 0.14 higher (4.01 lower to 4.28 higher)	⊕⊕○○ LOW ^b	CRITICAL	Presence of post-operative pain (≤ 6 months) № of participants: 269 (3 observational studies)	RR 0.88 (0.60 to 1.29)	Study population 18.9% (12.9 to 27.8)	21.5%	2.6% fewer (8.6 fewer to 6.2 more)	⊕○○○ VERY LOW ^{b,e,f}	IMPORTANT	Serious adverse events [interventions, prolonged hospitalization, ICU admission, or death] № of participants: 221 (1 RCT)	RR 0.36 (0.10 to 1.34)	Study population 2.6% (0.7 to 9.8)	7.3%	4.7% fewer (6.6 fewer to 2.5 more)	⊕⊕○○ LOW ^b	CRITICAL	Return to OR [for postoperative complications] № of participants: 469 (9 observational studies)	RR 0.79 (0.28 to 2.22)	Study population 1.8% (0.7 to 5.2)	2.3%	0.5% fewer (1.7 fewer to 2.9 more)	⊕○○○ VERY LOW ^{b,g,h}	CRITICAL	Based on expert opinion, the panel believed that POEM had additional desirable effects not assessed by the reviewed literature including lack of incisional hernia risk, lower wound infection rate, and lack of post-fundoplication side effects such as bloating, flatulence, and inability to belch or vomit. Most of the evidence was based on adults with type 1 or 2 achalasia, or on adults with unspecified achalasia type. Only one small observational study evaluated POEM versus Heller myotomy in pediatric patients (n = 18 patients) and only one study had predominantly type 3 achalasia. Based on expert opinion, the panel felt that adult evidence could be generalized to pediatric patients but not type 3 achalasia. Given their experience, they felt there were more desirable effects from POEM over laparoscopic Heller myotomy for type 3 adult or pediatric achalasia				
№ of participants: 148 (1 RCT)				(2.26 lower to 0.76 higher)																																					
Quality of Life Improvement - 2 years (GIQOL, range 0 - 144) № of participants: 202 (1 RCT)	-	Mean improvement in GIQOL of 28.1 points	Mean improvement in GIQOL of 24.5 points	MD 0.14 higher (4.01 lower to 4.28 higher)	⊕⊕○○ LOW ^b	CRITICAL																																			
Presence of post-operative pain (≤ 6 months) № of participants: 269 (3 observational studies)	RR 0.88 (0.60 to 1.29)	Study population 18.9% (12.9 to 27.8)	21.5%	2.6% fewer (8.6 fewer to 6.2 more)	⊕○○○ VERY LOW ^{b,e,f}	IMPORTANT																																			
Serious adverse events [interventions, prolonged hospitalization, ICU admission, or death] № of participants: 221 (1 RCT)	RR 0.36 (0.10 to 1.34)	Study population 2.6% (0.7 to 9.8)	7.3%	4.7% fewer (6.6 fewer to 2.5 more)	⊕⊕○○ LOW ^b	CRITICAL																																			
Return to OR [for postoperative complications] № of participants: 469 (9 observational studies)	RR 0.79 (0.28 to 2.22)	Study population 1.8% (0.7 to 5.2)	2.3%	0.5% fewer (1.7 fewer to 2.9 more)	⊕○○○ VERY LOW ^{b,g,h}	CRITICAL																																			
	<ul style="list-style-type: none"> a. Eckardt score is a blunt and imperfect measurement for the outcome of dysphagia b. Small sample size and very wide confidence interval suggest the potential for both important benefit and harm. c. On Cochrane Risk of Bias tool 2.0, there is some concern for bias due to missing outcome data. d. All studies have a high risk of bias with one or more of the following concerns for selection bias: less surgeon experience with POEM and non-comparability of multiple baseline risk factors including greater reflux/regurgitation symptoms in Heller. e. One of three studies investigated 6-month chest pain (Bhayani 2014) while the others reported post-operative pain. No statistically significant heterogeneity was observed (I2 = 0). f. 7 out of 9 observational studies are at high risk of bias due to one or more of the following concerns: incomparability of the groups at baseline on prognostic factors associated with the outcome (e.g. less surgeon experience with POEM and differences in preoperative interventions and age), attrition rate over 30%, and differential duration of follow up with shorter POEM follow-up. g. One study was pediatric which contributed 10.2% weight to the pooled estimate; test for subgroup heterogeneity was not significant with I2 = 0%. 																																								

Undesirable Effects																																							
How substantial are the undesirable anticipated effects?																																							
JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS																																					
<ul style="list-style-type: none"> ○ Large ○ Moderate ● Small ○ Trivial ○ Varies ○ Don't know 	<p>One low risk of bias recent randomized control trial on POEM versus laparoscopic Heller myotomy/Dor and twelve predominantly high risk of bias observational studies on POEM versus laparoscopic Heller myotomy were used to inform the panel's decision.</p> <p>Severe esophagitis in Werner 2019, when defined as type C and D esophagitis on EGD, was not significantly different between POEM and LHM.</p> <table border="1"> <thead> <tr> <th rowspan="2">Outcomes</th> <th rowspan="2">Relative effect (95% CI)</th> <th colspan="3">Anticipated absolute effects* (95% CI)</th> <th rowspan="2">Certainty of the evidence (GRADE)</th> <th rowspan="2">Importance</th> </tr> <tr> <th>With POEM</th> <th>With Heller</th> <th>Difference</th> </tr> </thead> <tbody> <tr> <td> Reflux esophagitis – 2 yr (Grade B-D assessed during EGD) assessed with: EGD Ne of participants: 165 (1 RCT) </td> <td> RR 1.79 (0.90 to 3.59) </td> <td> Study population 22.9% (11.5 to 46) </td> <td> 12.8% </td> <td> 10.1% more (1.3 fewer to 33.2 more) </td> <td> ⊕○○○ VERY LOW^{c,d} </td> <td> CRITICAL </td> </tr> </tbody> </table> <p>b. Small sample size and very wide confidence interval suggest the potential for both important benefit and harm. c. On Cochrane Risk of Bias tool 2.0, there is some concern for bias due to missing outcome data.</p> <p>Four observational studies also provided information on cost (in USD) for both POEM and LHM, though without any measure of variance with which meta-analysis could be performed. Despite great variation in components of cost, three of four studies stated POEM was more expensive, but the estimated cost difference was small.</p> <table border="1"> <thead> <tr> <th>Study</th> <th>POEM</th> <th>LHM</th> <th>Cost details</th> </tr> </thead> <tbody> <tr> <td>Greenleaf 2018</td> <td>10,763.21 (index admission 8629.70)</td> <td>8923.43 (index admission 7603.60)</td> <td>Includes pre-procedure (medical costs within 1 year prior to their index procedure), index admission, and post-procedure costs (within 1 year following index admission) from 2003-2016, adjusted to 2015 prices (n = 41)</td> </tr> <tr> <td>Hanna 2018</td> <td>3473</td> <td>3024</td> <td>Includes OR time and supplies, from 2011-2016.</td> </tr> <tr> <td>Khashab 2017</td> <td>14,481 (in-room 5070)</td> <td>17,782 (in-room 7616)</td> <td>Total cost includes in room charges, supplies, intraprocedural charges (in-room, supplies, drugs), and inpatient stay charges from 2012-2014.</td> </tr> <tr> <td>Miller 2017</td> <td>12,120</td> <td>11,582</td> <td>Annual cost per cure during the first-year post procedure from 2011 - 2015.</td> </tr> </tbody> </table>	Outcomes	Relative effect (95% CI)	Anticipated absolute effects* (95% CI)			Certainty of the evidence (GRADE)	Importance	With POEM	With Heller	Difference	Reflux esophagitis – 2 yr (Grade B-D assessed during EGD) assessed with: EGD Ne of participants: 165 (1 RCT)	RR 1.79 (0.90 to 3.59)	Study population 22.9% (11.5 to 46)	12.8%	10.1% more (1.3 fewer to 33.2 more)	⊕○○○ VERY LOW ^{c,d}	CRITICAL	Study	POEM	LHM	Cost details	Greenleaf 2018	10,763.21 (index admission 8629.70)	8923.43 (index admission 7603.60)	Includes pre-procedure (medical costs within 1 year prior to their index procedure), index admission, and post-procedure costs (within 1 year following index admission) from 2003-2016, adjusted to 2015 prices (n = 41)	Hanna 2018	3473	3024	Includes OR time and supplies, from 2011-2016.	Khashab 2017	14,481 (in-room 5070)	17,782 (in-room 7616)	Total cost includes in room charges, supplies, intraprocedural charges (in-room, supplies, drugs), and inpatient stay charges from 2012-2014.	Miller 2017	12,120	11,582	Annual cost per cure during the first-year post procedure from 2011 - 2015.	<p>The only undesirable effect (Reflux esophagitis on EGD) was considered to be a moderate effect as a short-term outcome but was downgraded to a small effect as the biggest difference between procedures was seen for Grade A esophagitis and not for the more clinically relevant grades B-D esophagitis. The panel felt longer-term comparative data is needed to assess reflux outcome differences between the procedures in the long term, including esophagitis and its sequelae.</p> <p>Most of the evidence was based on adults with type 1 or 2 achalasia, or on adults with unspecified achalasia type. Only one small observational study evaluated POEM versus Heller myotomy in pediatric patients (n = 18 patients) and only one study had predominantly type 3 achalasia. Based on expert opinion, the panel felt that adult evidence could be generalized to pediatric patients but not type 3 achalasia.</p>
Outcomes	Relative effect (95% CI)			Anticipated absolute effects* (95% CI)					Certainty of the evidence (GRADE)	Importance																													
		With POEM	With Heller	Difference																																			
Reflux esophagitis – 2 yr (Grade B-D assessed during EGD) assessed with: EGD Ne of participants: 165 (1 RCT)	RR 1.79 (0.90 to 3.59)	Study population 22.9% (11.5 to 46)	12.8%	10.1% more (1.3 fewer to 33.2 more)	⊕○○○ VERY LOW ^{c,d}	CRITICAL																																	
Study	POEM	LHM	Cost details																																				
Greenleaf 2018	10,763.21 (index admission 8629.70)	8923.43 (index admission 7603.60)	Includes pre-procedure (medical costs within 1 year prior to their index procedure), index admission, and post-procedure costs (within 1 year following index admission) from 2003-2016, adjusted to 2015 prices (n = 41)																																				
Hanna 2018	3473	3024	Includes OR time and supplies, from 2011-2016.																																				
Khashab 2017	14,481 (in-room 5070)	17,782 (in-room 7616)	Total cost includes in room charges, supplies, intraprocedural charges (in-room, supplies, drugs), and inpatient stay charges from 2012-2014.																																				
Miller 2017	12,120	11,582	Annual cost per cure during the first-year post procedure from 2011 - 2015.																																				

Certainty of evidence																										
What is the overall certainty of the evidence of effects?																										
JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS																								
<ul style="list-style-type: none"> ● Very low ○ Low ○ Moderate ○ High ○ No included studies 	<p>This decision was informed by the certainty of evidence for individual outcomes and based on the critical outcomes.</p> <table border="1"> <thead> <tr> <th>Outcomes</th> <th>Importance</th> <th>Certainty of the evidence (GRADE)</th> </tr> </thead> <tbody> <tr> <td>Success/Symptom resolution by Eckardt Score for dysphagia (2 years)</td> <td>CRITICAL</td> <td>⊕○○○ VERY LOW</td> </tr> <tr> <td>Post-operative relaxation pressure of LES (mmHg)</td> <td>IMPORTANT</td> <td>⊕○○○ VERY LOW</td> </tr> <tr> <td>Quality of Life Improvement - 2 years (GIQOL)</td> <td>CRITICAL</td> <td>⊕⊕○○ LOW</td> </tr> <tr> <td>Reflux esophagitis (2yr) (Grade B-D assessed during EGD) assessed with: EGD</td> <td>CRITICAL</td> <td>⊕○○○ VERY LOW</td> </tr> <tr> <td>Pain (post-operative)</td> <td>IMPORTANT</td> <td>⊕○○○ VERY LOW</td> </tr> <tr> <td>Serious adverse events</td> <td>CRITICAL</td> <td>⊕⊕○○ LOW</td> </tr> <tr> <td>Return to OR</td> <td>CRITICAL</td> <td>⊕○○○ VERY LOW</td> </tr> </tbody> </table>	Outcomes	Importance	Certainty of the evidence (GRADE)	Success/Symptom resolution by Eckardt Score for dysphagia (2 years)	CRITICAL	⊕○○○ VERY LOW	Post-operative relaxation pressure of LES (mmHg)	IMPORTANT	⊕○○○ VERY LOW	Quality of Life Improvement - 2 years (GIQOL)	CRITICAL	⊕⊕○○ LOW	Reflux esophagitis (2yr) (Grade B-D assessed during EGD) assessed with: EGD	CRITICAL	⊕○○○ VERY LOW	Pain (post-operative)	IMPORTANT	⊕○○○ VERY LOW	Serious adverse events	CRITICAL	⊕⊕○○ LOW	Return to OR	CRITICAL	⊕○○○ VERY LOW	<p>These data also serve as a proxy for long-term outcomes (5-year, 10 year, and 15-year follow-up). For long-term outcomes, the overall certainty of evidence is further downgraded based on indirectness because it is unknown how long-lasting either the desirable or undesirable effects will be.</p>
Outcomes	Importance	Certainty of the evidence (GRADE)																								
Success/Symptom resolution by Eckardt Score for dysphagia (2 years)	CRITICAL	⊕○○○ VERY LOW																								
Post-operative relaxation pressure of LES (mmHg)	IMPORTANT	⊕○○○ VERY LOW																								
Quality of Life Improvement - 2 years (GIQOL)	CRITICAL	⊕⊕○○ LOW																								
Reflux esophagitis (2yr) (Grade B-D assessed during EGD) assessed with: EGD	CRITICAL	⊕○○○ VERY LOW																								
Pain (post-operative)	IMPORTANT	⊕○○○ VERY LOW																								
Serious adverse events	CRITICAL	⊕⊕○○ LOW																								
Return to OR	CRITICAL	⊕○○○ VERY LOW																								

Values		
Is there important uncertainty about or variability in how much people value the main outcomes?		
JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
<ul style="list-style-type: none"> ○ Important uncertainty or variability ○ Possibly important uncertainty or variability ● Probably no important uncertainty or variability ○ No important uncertainty or variability 		<p>This judgement used panel expert opinion based on prior patient interactions to judge how likely patients would vary in how much they value the main outcomes AFTER an informed discussion pertaining to evidence available. Although the panel felt that there was probably no important variation, they agreed that given a lack of evidence on values, this variation may still exist. Additionally, a minority of the panel felt there may "possibly" be important variability.</p>

Balance of effects		
Does the balance between desirable and undesirable effects favor the intervention or the comparison?		
JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
<input type="radio"/> Favors the comparison <input type="radio"/> Probably favors the comparison <input checked="" type="radio"/> Does not favor either the intervention or the comparison <input type="radio"/> Probably favors the intervention <input type="radio"/> Favors the intervention <input type="radio"/> Varies <input type="radio"/> Don't know	Both desirable and undesirable anticipated effects were judged to be small.	Some panel members initially felt that the balance of desirable and undesirable effects favored the intervention (POEM) but after further deliberation among panel members the panel agreed that the balance of effects did not favor one over the other.

Acceptability		
Is the intervention acceptable to key stakeholders?		
JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
<input type="radio"/> No <input type="radio"/> Probably no <input type="radio"/> Probably yes <input checked="" type="radio"/> Yes <input type="radio"/> Varies <input type="radio"/> Don't know		Both LHM and POEM are already established procedures. The panel felt there would be clear acceptance for the recommendation to perform either procedure.

Feasibility		
Is the intervention feasible to implement?		
JUDGEMENT	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
<input type="radio"/> No <input type="radio"/> Probably no <input type="radio"/> Probably yes <input checked="" type="radio"/> Yes <input type="radio"/> Varies <input type="radio"/> Don't know		While some practitioners who are unfamiliar or inexperienced with POEM may not find a recommendation for POEM in preference to LHM to be feasible, the recommendation for either POEM or LHM was considered feasible as currently there are numerous groups that have accumulated experience with POEM.

SUMMARY OF JUDGEMENTS

	JUDGEMENT						
DESIRABLE EFFECTS	Trivial	Small	Moderate	Large		Varies	Don't know
UNDESIRABLE EFFECTS	Large	Moderate	Small	Trivial		Varies	Don't know
CERTAINTY OF EVIDENCE	Very low	Low	Moderate	High			No included studies
VALUES	Important uncertainty or variability	Possibly important uncertainty or variability	Probably no important uncertainty or variability	No important uncertainty or variability			
BALANCE OF EFFECTS	Favors the comparison	Probably favors the comparison	Does not favor either the intervention or the comparison	Probably favors the intervention	Favors the intervention	Varies	Don't know
ACCEPTABILITY	No	Probably no	Probably yes	Yes		Varies	Don't know
FEASIBILITY	No	Probably no	Probably yes	Yes		Varies	Don't know

TYPE OF RECOMMENDATION

Strong recommendation against the intervention <input type="radio"/>	Conditional recommendation against the intervention <input type="radio"/>	Conditional recommendation for either the intervention or the comparison <input checked="" type="radio"/>	Conditional recommendation for the intervention <input type="radio"/>	Strong recommendation for the intervention <input type="radio"/>
---	--	---	--	---

CONCLUSIONS

Recommendation

The panel suggests that adult patients with type I and II achalasia may be treated with either POEM or laparoscopic Heller myotomy based on surgeon and patient's shared decision-making (conditional recommendation, very low certainty evidence). Given the lack of data in children, this recommendation may also be generalized to the pediatric population. No evidence-based recommendation can be made for patients with type III achalasia.

As expert opinion given their experience, however, the panel favored POEM over laparoscopic Heller myotomy for type 3 adult or pediatric achalasia.

Justification

Both interventions are established procedures practiced in a variety of settings and environments and have proven and equivalent efficacy and safety. Importantly, the overall balance of desirable and undesirable effects does not favor one procedure over the other.

Subgroup considerations

This recommendation is primarily for adult patients because the vast majority of comparative data is based on adult populations. Given a lack of substantial data in the pediatric population, however, this evidence is the best proxy available for pediatric achalasia. In the absence of evidence that suggests that adult data are a poor proxy for the pediatric population, this panel also suggests that pediatric patients with achalasia can be treated with either POEM or laparoscopic Heller myotomy based on surgeon and patient's shared decision-making.

The subtype of Achalasia is often either not given, or outcomes are not stratified by subtype, precluding subgroup analyses on POEM versus LHM for each achalasia subtype. In those studies which report distribution of achalasia type, type 2 or 1 are usually predominant with only 1 study reporting predominantly Type 3 patients. This recommendation thus applies best to type 1 and 2 achalasia. Limited evidence is available on POEM versus LHM in type 3 achalasia patients to make an evidence-based recommendation. Based on expert opinion, however, POEM appears to perform better in type 3 patients as a longer myotomy can be performed with that approach. For type 3 achalasia patients, therefore this panel suggested consideration of POEM for its higher efficacy but deemed both POEM and LHM as safe choices.

Implementation considerations

None

Monitoring and evaluation

None

Research priorities

The panel makes multiple suggestions for future research priorities:

- More research is needed on Type 3 achalasia specific outcomes after POEM vs. Heller. This can be achieved either with type 3-only study populations, or studies with sample size large enough to perform adequately powered subgroup analysis based on achalasia subtype.
- More research is needed on pediatric populations. This can be achieved either with pediatric only studies, or studies with sample size large enough to perform adequately powered subgroup analysis based on pediatric versus adult population.
- Longer term results are needed for all outcomes given the chronic nature of achalasia. Surveillance and follow-up past 10 years is needed, especially by high quality comparative studies.
- Future studies should include better measures to determine the presence of dysphagia than Eckardt score which tends to be not very specific. More accurate, objective alternatives include manometry and timed barium swallow studies.
- More research is needed into outcomes of POEM versus Heller myotomy that relate specifically to the fundoplication component of a Heller myotomy.
- Research should be performed to establish whether there is a correlation between post-POEM LES pressure and post-POEM outcomes. Such data exist for Heller myotomy but not for POEM and the panel felt that it is not appropriate to apply the evidence from Heller to POEM as there may be substantial differences.
- While the current evidence suggests POEM leads to greater postoperative reflux, at least in the first 2 years post procedure, there is no research on the role, patient acceptance and efficacy of PPI use after POEM for this undesirable outcome. The panel recommends further investigation of strategies to address undesirable effects for both POEM and Heller myotomy and their relative efficacy for both interventions.