

## *Comparative Effectiveness Research Review Disposition of Comments Report*

### **Research Review Title:** *Surgical Options for Inguinal Hernia: Comparative Effectiveness Review*

Draft review available for public comment from November 1, 2011 to November 29, 2011.

**Research Review Citation:** Treadwell J, Tipton K, Oyesanmi O, Sun F, Schoelles K. Surgical Options for Inguinal Hernia: Comparative Effectiveness Review. Comparative Effectiveness Review No. 70. (Prepared by the ECRI Institute Evidence-based Practice Center under Contract No. 290-2007-10063.) AHRQ Publication No. 12-EHC091-EF. Rockville, MD: Agency for Healthcare Research and Quality. August 2012. Available at: [www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm).

### **Comments to Research Review**

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each comparative effectiveness research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Commentator	Section	Comment	Response
Peer reviewer #1	No particular section	a. General Comments: Report is meaningful for key questions. The questions are appropriate and explained.	Thank you for your feedback.
Peer reviewer #1	Introduction	b. Introduction: Introduction needs to highlight that open repair can and is usually performed under local anaesthetic in some centres and countries. Laparoscopic repair on the other hand requires a general anaesthetic with muscle relaxation. As there are no trials comparing local anaesthetic open mesh repair with laparoscopic repair comparison is not possible but should be considered as a future study.	We added the following to the Introduction: “Laparoscopic repair invariably involves general anesthesia, whereas open mesh repair often involves local or regional anesthesia”
Peer reviewer #1	Methods	c. Methods: Inclusion and exclusion criteria are justifiable, search strategies, diagnostic criteria and statistical methods are appropriate.	Thank you for your feedback.
Peer reviewer #1	Results	d. Results: Detail in results appropriate, clinical help in interpretation would be helpful. The figures and tables are adequate. To my knowledge no major study has been excluded and included studies are appropriate.	Thank you for your feedback.
Peer reviewer #1	Discussion/ Conclusion	e. Discussion/ Conclusion: Implications of major findings are clearly stated and limitations of studies rightly highlighted. Future research needs to examine issue of local versus general anaesthetic inguinal hernia repair.	We added the following to the Future Research section: “Another issue for open and laparoscopic repair concerns the mode of anesthesia (local or general). Laparoscopic repair invariably involves general anesthesia, whereas open mesh repair can be local anesthesia. This difference could potentially explain any short-term differences in postoperative pain, if the anesthesia mode has any lingering effects. Future studies should consider comparing modes of anesthesia to determine its impact.”
Peer reviewer #1	No particular section	f. Clarity and Usability: Report is well presented and main points are clear. Conclusions can help inform policy and practice decisions.	Thank you for your feedback.
Peer reviewer #2	No particular section	a. General Comments: This is an exhaustive report that in general is well conducted and written. There are a couple of places where I question the ability of the authors to make the conclusions they do, in particular when they only have one qualified study that addresses the question.	See specific responses below.

Commentator	Section	Comment	Response
Peer reviewer #2	No particular section	The other particular methodologic issues I don't believe they have addressed or mentioned as limitations are related to 1) how hernia recurrence was detected in the studies reporting this (there is no gold standard) and what proportion of patients in the studies were available for followup,	<p>Regarding how studies measured hernia recurrences, one of our risk-of-bias items was used to address this. Most studies (59 of 72 studies that reported recurrence rates in the draft you reviewed) had patients visit the clinic for a physical examination, and so the determination was made by a clinician. In the other 13 studies, the method of determination was either not reported (8 studies) or clearly did not involve a clinic visit (5 studies). We added these observations to the discussion section in the context of our discussion of registries: "Another problem with registry data is the difficulty users would have to determining whether the assessment of hernia recurrence involved a patient visiting a clinic or simply involved self-report via a telephone interview or questionnaire. Most of the studies we reviewed (i.e. not of registries) had patients come into the clinic for a physical assessment, rather than rely on patient reports of recurrence."</p> <p>Regarding the proportion of patients who were available for followup, for short term outcomes this was obviously high, but for &gt;6 month hernia recurrence, we measured this with another risk-of-bias item. This latter item found that 60 of 72 studies had good followup (defined as &gt;85% of patients enrolled) for all of their recurrence datapoints, four others had good followup for some but not all timepoints, four others had good followup for none of the recurrence timepoints, and three others did not report sufficient information. We also added these observations to the discussion section about low precision: "The problem was insufficient enrollment, not a lack of follow-up of enrolled patients, because most studies did report data on at least 85% of enrolled patients."</p>
Peer reviewer #2	No particular section	and 2) what instructions patients were given as far as RTW/RTDA. There have been a couple of attempts to quantify whether the RTW/RTDA is really a function of the instructions given by the surgeon, and the patient's work status (self employed patients return to work much sooner after inguinal hernia repair than those who have workman's compensation) than the patient's ability.	Studies did not report what instructions were given to patients in defining RTW/RTDA. The specific manner of reporting varied among studies, and these are included in the evidence tables.
Peer reviewer #2	No particular section	It should be more explicitly stated (perhaps in the title) that these results do not apply to women as so many of the studies excluded women.	It is true that many studies excluded women, however some did not. We do note in the discussion that the results apply mostly to middle-aged men.

Commentator	Section	Comment	Response
Peer reviewer #2	Executive summary and introduction	Lastly, there are some pretty major errors (the goal of surgery, how lap repairs are done) in the executive summary and full introductions that will lead many surgeons to stop reading this review because these errors give the impression the authors of the review really don't understand hernia repair at all.	The review team was not comprised of hernia surgeons, but instead experts in the systematic review of evidence. We ensured that the review contained no errors that we were aware of. We added that the goal of surgery is preventing incarceration/strangulation.
Peer reviewer #2	Introduction	b. Introduction: In the abstract introduction, the sentence "in children, incidence..." follows right after description of recurrence in adults. This makes the reader think the authors are referring to recurrence in children. This is not a problem in the full introduction because the sentence starts a new paragraph.	We corrected this problem by moving the sentence about recurrence further down.
Peer reviewer #2	Executive summary	On page ES-1 to ES-2 and on page 2 the authors completely misrepresent how laparoscopic repairs are performed. Please find a surgeon familiar with these to rewrite these paragraphs for you. Your inability to describe them appropriately (they aren't accurate using lay language) undermines your credibility.	See above response
Peer reviewer #2	Executive summary	On page ES-1, line 28 the authors state "The primary goals of surgery include preventing recurrence of the hernia, returning the patient to normal activities quickly,..." when really the primary goal of surgery is to prevent hernia accidents (incarceration and strangulation).	We fixed the sentence accordingly.
Peer reviewer #2	Executive summary	On page 2, lines 53 to 55 are written poorly. Why are you even talking about prophylactic antibiotics when you don't have this as a key question? The infection rate for inguinal hernia repairs is exceedingly low anyway so I believe you could delete this entirely.	We deleted the sentence.
Peer reviewer #2	Methods	c. Methods: The methodology and explanations of such are detailed and easy to understand. While I agree with the methodology described, I believe on occasion the authors stretched things and reached conclusions that shouldn't have been reached (see below)	See specific responses below
Peer reviewer #2	Results	d. Results: The amount of detail is very good. The problem is that in several situations, there is only one study that has been done, measuring the particular outcome and yet you make a conclusion. Maybe I am confused, but I believe the document you are preparing will be used by many patients. They will not have the stamina to read the entire thing and understand for instance for KQ 1 that QoL was not the primary outcome of the Fitzgibbons WW trial and while QoL increase is an interesting significant finding, because it wasn't the study's primary outcome and it wasn't measured in the other WW trial, I think listing it as a finding is premature	It is true that QOL was not the "primary outcome" of the Fitzgibbons trial, however the evidence did show an effect, and in our view the authors' intentions and the lack of replication of the finding, taken together, were not sufficient to deny the finding of an advantage in the operated group.

Commentator	Section	Comment	Response
Peer reviewer #2	Results	In addition, several times in the sections referring to KQ 1 you state that pain was less in the surgery group vs ww (2.2 vs 5.1) which was a NON SIGNIFICANT finding (and the primary outcome measure) of the study.	We changed the wording to avoid any suggestion that there was a difference when in fact there was a statistically nonsignificant finding.
Peer reviewer #2	Results	Another example of using one study and making conclusions from it is for KQ4 where you state RTW favors TAP vs. TEPP. This was drawn from one study and with the caveats I have already listed for problems with RTW/RTDA, I believe this too is an unsupported conclusion.	We actually have two studies comparing TAPP vs. TEP and measuring RTDA, and four studies comparing TAPP vs. TEP and measuring RTW.
Peer reviewer #2	Results	KQ2c (open vs lap repair of recurrent hernias), page 35, you list the number of patients in the studies ranging from 50-2,164. The 2,164 was the Neumayer trial and that study did not have 2,164 RECURRENT hernias; just under 10% of the 2,164 repairs were for recurrent hernias. Using this total number is misleading and implies there has been a study of over 2,000 re-repairs.	We fixed the error accordingly.
Peer reviewer #2	Results	When discussing open vs lap in kids, you should probably describe somewhere how the lap repair is done since it is not done with mesh AND you have previously so inaccurately described the lap repairs in adults, the reader will be very confused.	We added text to clarify these points.
Peer reviewer #2	Discussion/ Conclusion	e. Discussion/ Conclusion: The implications are clearly stated (see above for the ones I think are not clearly SUPPORTED). The limitations are not described adequately (see above for my comments on limitations of determination of hernia recurrence and of RTW/RTDA issues). I don't believe any important literature was omitted. The future research section is clear and parts could be easily translated into new research (with appropriate funding). One big overarching issue is the need for standardized measures and definitions that should be used when looking at hernia repairs. Many of these exist due to the large studies that have been done. Explicitly stating the need for such standardized definitions/measures in future studies (whether RCTs or registries) is probably worthwhile.	We added a paragraph to the Future Research on the need for standardization with regard to how to measure various outcomes after hernia repair, including outcomes such as return to work and return to activities of daily living.
Peer reviewer #2	No specific section	f. Clarity and Usability: See my comments above. My biggest concern is that only parts of this report will be used (i.e. the written recommendations) without understanding of the strength of evidence or potential bias. The non-sophisticated reader will not be able to adequately interpret this document. The experienced surgeon will discredit it because of the major errors in description of the procedures (see above).	See above responses

Commentator	Section	Comment	Response
Peer reviewer #3	No specific section	a. General Comments: The study is extremely difficult to follow and comprehend. The authors have collected so much data, oftentimes confusing, and they have become "lost" in their own data.	Given the volume of information, we parsed as well as we could, and created overall summaries where appropriate.
Peer reviewer #3	Introduction	b. Introduction: Too concise, fails to truly introduce the reader into the real problem of hernia.	We wanted the introduction to be concise because the document itself is quite long (more than 1000 pages including the appendices). The introduction does include descriptions of types of inguinal hernias, different populations, various management approaches and the reasons for their adoption. We believe this introduction is adequate for the purpose of the review.
Peer reviewer #3	Introduction	c. Methods: Very confusing, difficult to follow, and "detached" from the real problems of hernia repair.	See above response
Peer reviewer #3	Results	d. Results: Same as above.	See above response
Peer reviewer #3	Discussion/ Conclusion	e. Discussion/ Conclusion: The discussion is clearly written by non-specialists in hernia	Correct. We are experts in systematic review, not in hernia repair. Our view is that our lack of hernia expertise does not preclude us from producing a systematic review that is useful to clinicians and policymakers.
Peer reviewer #3	No specific section	f. Clarity and Usability: The review is not clear and not usable at all. All it can do is to confuse the reader. It fails to convey basic differences between methods, their inherent risks, and pros and cons. If one wants to confuse a reader, they must let him read this.	See above response
Peer reviewer #4	No particular section	a. General Comments: I thought this report was laid out nicely. I like the Key Questions and direct response to each. The compiled data is powerful and meaningful.	Thank you for your comments.
Peer reviewer #4	Introduction	b. Introduction: Very reasonable outline. My only reservations with the background and intro relate to data regarding incidence and prevalence of inguinal hernia. As this report clearly showed, we are woefully short on "good" data on a variety of topics...especially on incidence and prevalence. We have two population-based studies in press that suggest that the 770,000 hernia repairs per year is an overestimate (Olmsted County, MN data would suggest it is closer to 600,000) and that roughly 42% of all males will develop an inguinal hernia in their lifetime. This data has not yet made it to press.	Thank you for alerting us to the pending publications. We have added a footnote containing these estimates, and cited the personal communication.
Peer reviewer #4	Methods	c. Methods: I believe the answers to all of these questions are "yes". I thought the report was well-researched, stayed on task, and attempted to answer appropriate questions. It did not overstate results.	Thank you for your feedback.

Commentator	Section	Comment	Response
Peer reviewer #4	Results	d. Results: Again, well researched and broad coverage of large topic. New data becomes available daily and you can't be expected to have relevant data recently reported (our group along has or is in the process of publishing 6 papers on inguinal hernia repair this fall). I believe the graphs and tables become excessive and far favored the Executive Summary over the minutia displayed later on. If anything, I would prefer the conclusions to key questions be BOLDED when significance was found.	In such a large report, it is always a challenge to present material in a way that most satisfies the most users. All figures of meta-analyses are placed near the end so they do not disrupt the flow (however the results of meta-analyses are presented in the text). We did not add boldfacing in the executive summary because we felt that would detract from the existing boldfacing of Key Questions and other visual components such as bullets.
Peer reviewer #4	Discussion/ Conclusion	e. Discussion/ Conclusion: Globally, "yes". As above, I think the major finding were clearly stated, but they get lost a bit in the details of the report. I would prefer that any major findings be BOLDED or set apart somehow.	See above response
Peer reviewer #4	No particular section	f. Clarity and Usability: Very well structured and outlined. The findings and conclusions of this report are in agreement with my own surgical practice (23 years) and experience (~3000 inguinal hernia repairs)...not that that statement makes it right, but I believe the conclusions are well stated, appropriate, and clearly acceptable for others to read and comprehend. If policy is made off of this report, I believe the data is fair and representative of the current best surgical practice in the USA.	Thank you for your feedback.
Christine Chang, AHRQ	General	Nicely written report.	Thank you for your feedback.
Christine Chang, AHRQ	General	Methods have much detail related to how decisions were made. Very helpful. Inclusion of details from your protocol would be helpful (dual abstraction, resolution of conflicts, oversight, etc).	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	General	Were you able to account for hernia type, severity of the hernia, etc in the analysis of results?	For Key Question 2a, where we observed large differences among study results for a few outcomes, we did examine these to determine whether they could help explain differences in study findings. However, the analyses did not reveal any consistent factors explaining the differences. The corresponding Results sections state this. The evidence tables do contain this information in case a user wants to consider these explanations for themselves.

Commentator	Section	Comment	Response
Christine Chang, AHRQ	Introduction	Consider including a few words about the theoretical advantages (or disadvantages) of one surgical approach over the other (gets at some of the decisional uncertainty for TAPP, TEP) have some mention in the full report-is this sufficient?	We have added more text on different open procedures. "Lichtenstein repair involves suturing the mesh in front of the hernia defect. Mesh plug repair involves a pre-shaped mesh plug is introduced into the hernia weakness during open surgery and a piece of flat mesh is positioned on top of the hernia defect. The Lichtenstein repair, performed under local anesthesia, is generally used during the repair of primary inguinal hernia and may also be suitable during the repair of recurrent inguinal hernia where the defect is > 4cm3. The mesh plug repair may require less dissection, and may reduce patients' discomfort postoperatively thereby quickening the return to 'normal activity'. A possible disadvantage of the mesh plug repair may be related to hardening of the plug resulting in pain in the groin region."
Christine Chang, AHRQ	Introduction	Include any FDA recalls, warnings, etc. This makes a distinction between what is commonly known and what was systematically reviewed in terms of adverse events/harms. Probably most important for mesh and glue.	This change was made before the review was posted for public comment. We have also added additional FDA-recall text in the post-review draft.
Christine Chang, AHRQ	Executive Summary	Methods: Would characterize the TEP as individuals giving input on the protocol and review, not as those collaborating in the review. That might imply that they had greater involvement than they actually did.	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Executive Summary	Please include a summary table with major outcomes, direction of effect and SOE.	The executive now contains SOE ratings in a summary table, along with major outcomes and direction of effect
Christine Chang, AHRQ	Executive Summary	Please include the SOE.	See previous response
Christine Chang, AHRQ	Executive Summary	Were you able to do any meta-analysis? It isn't clear from the results section.	We added mention that meta-analyses were performed for seven KQs (2a, 2b, 2c, 3, 4, 5, 6).
Christine Chang, AHRQ	Introduction	First paragraph: I am not sure what mention of setting, and the increased number of surgeries completed in the outpatient setting adds to an understanding of the KQ, and how it relates to specialized hernia centers and surgical experience.	We agree that the mention of outpatient setting was misplaced here. We had simply wanted to make the general point that most procedures are outpatient. So we have moved this point to earlier in the intro, where we discuss inguinal hernia surgery in general.
Christine Chang, AHRQ	Methods	instead of "proposal" would recommend "key questions and scope."	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Methods	Instead of "finalization of the topic" would recommend "finalization of the protocol."	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Methods	Search strategy: Please include information about the grey literature search and SIPS.	We added text about the grey literature search
Christine Chang, AHRQ	Methods	Search strategy: Please also note that the search will be updated during the peer review period, and any additional studies will be incorporated into the final report.	This change was made before the review was posted for public comment.

Source: <http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1176>  
 Published Online: August 17, 2012

Commentator	Section	Comment	Response
Christine Chang, AHRQ	Methods	Page 11, Length of hospital stay bullet: instead of “This change was approved by the TOO before completion of the review draft” would say “This was outlined in a protocol amendment dated X/X/XX.” That is more accurate.	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Methods	SOE rating, KQ 8 bullet: “SOE was not rated because no studies met inclusion criteria.” This is a result. Were there outcomes that you would have graded had there been evidence? Also if there were no studies, you could still grade the evidence-it would have been insufficient.	We added a sentence in the methods sections involving rating of SOE: If there were no studies for a given treatment comparison or Key Question, we rated the evidence as Insufficient.
Christine Chang, AHRQ	Methods	SOE: include definitions of SOE grading (consider Table 3 of Owens paper). Peer review and public commentary “...the initial draft report was prereviewed by the TOO and AE.” Recommend “The draft report was reviewed prior to peer review by the TOO and AE. The revised draft report was then sent to invited peer reviewers....”	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Methods	Peer review and public commentary Public commentary and peer review is for 4 weeks (28 days, not 30).	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Methods	Peer review and public commentary I would also add that “the EPC responses to all comments will be documented in a disposition of comment document which will be posted on the Effective Health Care website about 3 months after web publication of the evidence report. “	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Methods	Page 12, “Owens system”: better to say that this is the EHC methods guide, and reference the Owens paper. Stock language from the new content guidance (this is FYI only) “The methods for this comparative effectiveness review (CER) follow the methods suggested in the AHRQ <i>Methods Guide for Effectiveness and Comparative Effectiveness Reviews</i> (available at <a href="http://www.effectivehealthcare.ahrq.gov/methodsguide.cfm">http://www.effectivehealthcare.ahrq.gov/methodsguide.cfm</a> ).”	Instead of the Owens system. we now say: “We used the system described in the Effective Healthcare (EHC) Methods Guide to rate the strength of the evidence (SOE) for the major outcomes for each Key Question.”
Christine Chang, AHRQ	Methods	The strength of evidence is not the same as effect. The SOE relates to the certainty of the estimate, not the estimate itself. Separating this out will be easier for our readers, especially for those who are not familiar with our methods and terminology.	We had not meant to suggest that the rating is anything other than a rating of <b>confidence</b> . We reproduced the definitions in the Owens paper that refer to “confidence.” For greater clarity, we added the sentence “The strength of evidence is defined as one’s confidence in the evidence supporting a conclusion.”

Commentator	Section	Comment	Response
Christine Chang, AHRQ	Methods	Page 12: “We first determined whether the combined evidence on that outcome was sufficiently precise to permit a conclusion about the direction of the effect (either favors treatment A, favors treatment B, or indicates approximate equivalence by ruling out the MCSD). If not, then the rating was Insufficient (abbreviated INSUFF in our SOE tables). If it was sufficiently precise, then we assigned point values to the four core domains as follows....” It is better to separate out determination of effect from SOE. This seems to imply that precision may have been more important than other domains.	We agree that what is being rated is confidence in the effect, not the effect itself. A key issue is whether the evidence permits enough confidence to yield a conclusion. A conclusion is a statement about the direction of effect (e.g., A is better than B on outcome X, or B is better than A on outcome X, or A and B yield similar outcome X). We did not mean to suggest that the only thing that matters is precision. We edited the text to avoid this impression.
Christine Chang, AHRQ	Methods	Page 12: “For the additional domains, we sometimes added 1 for a large magnitude of effect, and we sometimes subtracted 1 for potential publication bias or selective outcome reporting (e.g., if a third or fewer of the studies included for that comparison had actually reported that outcome). The other two additional domains (all plausible confounders would reduce the effect, and dose-response association), were not relevant to any of our Key Questions.” It should be noted in the results section when you have done this (bumped up the SOE for large magnitude of effect, or down for publication bias/reporting issues). If the additional domains were not relevant, then best to leave this section out.	We did occasionally use these two additional domains (large effect, and reporting bias). We made sure that in the revision, where these upgrades and downgrades occurred, they were always mentioned in the Results section.
Christine Chang, AHRQ	Methods	Effect is categorized as favors A, favors B, or equivalent. The determination of equivalence: does this depend on the intent of the study? Is there a difference between equivalence and the inability to detect a difference?	None of these studies called themselves superiority trials, equivalence trials or non-inferiority trials. Typically the authors just wanted to compare patient outcomes after two surgical options. We did not think it critical that authors state beforehand what they intended their data to show. Thus, we did not penalize them for not calling themselves superiority trials, equivalence trials or non-inferiority trials. With regard to your second question, there is a big difference between equivalence and the inability to detect a difference. The first factor is a <b>conclusion</b> based on a narrow confidence interval near the null effect. The second factor is true of any statistically non-significant difference, and therefore does not clearly state whether the evidence permits a conclusion. We added text in the Methods section that the determination of equivalence did not depend on the authors' intent, but rather on the data.
Christine Chang, AHRQ	Results	overall description of studies. Thank you for the paragraph which outlines the organization of the rest of the section. This is very helpful.	Thank you.

Commentator	Section	Comment	Response
Christine Chang, AHRQ	Results	PRISMA figure: thank you for including the number of studies and number of articles, as well as the breakdown by KQ.	Thank you for your feedback.
Christine Chang, AHRQ	Results	Many reports include a bulleted list of key points at the beginning of each section for each KQ. Sample structure: High strength of evidence for benefit of intervention X vs. intervention Y based on four good-quality RCTs with consistent results. <i>OR</i> Intervention X had better patient-centered outcomes than intervention Y based on three RCTs and five prospective cohort studies (moderate strength of evidence). I know that this report was well into development before the formal content guidance was introduced. In summary paragraph at end of each question, please include SOE.	We made sure that the summary paragraph for each Key Question discusses all of the relevant SOE ratings.
Christine Chang, AHRQ	Results	Key Question 1 section: Table 1-is mislabeled as referring to KQ 2. Please include references in the table.	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Results	KQ 2: Outcome: patient satisfaction, page 25: You indicate that you did a statistical test on the outcomes. What statistical test did you use?	We added that they were chi-square tests.
Christine Chang, AHRQ	Results	Page 23: hernia recurrence. This sort of thing may warrant discussion in the discussion section, in understanding the clinical significance-when providers or patients think about this risk difference in terms what it means to the clinically, they may see it a bit different.	You seem to be referring to the part in KQ2a where we wrote: “The difference between these rates is only 2.05%, which is less than our predefined MCSD of three percentage points. This implies that the difference between open and laparoscopy, while statistically significant, is not substantial.” I think you mean that we need to make sure something is in the Overall discussion section regarding the clinical significance of small differences in recurrence rates. The overall discussion now has the following: <i>For Key Question 2, most outcomes favored laparoscopy, with the key exception of recurrence in the repair of primary hernia, which found slightly lower risk rates after open surgery (an estimated 2.6% for open surgery vs. an estimated 3.7% for laparoscopic surgery). We considered this to be smaller than a clinically significant difference, however some patients and clinicians may consider this an important difference. Another way to describe the difference, that may lead one to believe it is an important difference, is in relative terms: an estimated 43% higher risk after laparoscopic mesh repair than after open mesh repair in the context of primary hernia. . The infrequency of the outcome is why the relative effect sounds larger than an absolute effect.</i>

Commentator	Section	Comment	Response
Christine Chang, AHRQ	Results	Page 32, Summary of KQ 2b: "The evidence was insufficient to permit conclusions for only one outcome: that bilateral hernia patients return to work sooner if they receive laparoscopic repair. This was rated as Low strength of evidence because only one of the six included studies reported information on return to work." Please clarify.	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Results	Mesh: in the methods/KQ it indicates that mesh types will be divided into lightweight and heavyweight. It appears though in this section on page 56 that individual types were compared, without knowing their categorization. We should go with an organization that makes sense clinically.	This change was made before the review was posted for public comment.
Christine Chang, AHRQ	Results	Page 65: comment on non-English studies. Should this be in the discussion section	We added the following to the general discussion section: "Interestingly, even though we required English-language publication, 76% of the studies we included were conducted in countries whose primary language is not English. Thus, many researchers probably chose to translate their work into English. It is unclear whether researchers perform translation for all of their studies, or only for the portion of their work that they believe should receive greater prominence via English language journals."
Christine Chang, AHRQ	Discussion	I am a bit concerned about the statement that since almost all study participants were men, that the results do not apply women. It may be more accurate to say that because the data were lacking in women, the applicability is uncertain. However, in the absence of evidence, it is likely that the conclusions are applicable/not applicable because of [insert biological/anatomical/other rationale here-are there reasons why there could be potentially a difference between men and women?]	We used your more accurate phrasing for the review that was posted for public comment.
Christine Chang, AHRQ	Discussion	Since the study populations were primarily middle-aged men, wouldn't a research gap be related to women and individuals who are not middle-aged?	The proportion of people who receive inguinal hernia is overwhelming middle-aged men. Thus, the studies' patients reflected the general population. Even though technically this is a research gap, the bigger picture is that women and younger men very rarely develop an inguinal hernia. Therefore we did not call this a research gap.
Christine Chang, AHRQ	Discussion	Summary table: consider including information about magnitude of effect, especially since you have this information from some meta-analysis.	This change was made before the review was posted for public comment.

Commentator	Section	Comment	Response
Christine Chang, AHRQ	Discussion	Making the link between what is statistically significant and what is clinically significant is important. What are some of the considerations clinically when deciding what approach to take for this procedure? What do the results of this review mean in light of these considerations?	This change was made before the review was posted for public comment.
Jeanne-Marie Guise, Oregon EPC	Executive Summary	The executive summary, in general is well written.	Thank you for your feedback.
Jeanne-Marie Guise, Oregon EPC	General	Overall impressions: Presentation of findings follows key questions but somehow in the writing the structure is geared more to a reviewer than your everyday clinician or patient. Cross talking of numbers of studies and data matches up across sections but the lack of framing for results leaves a clinician needing to pull the information together from several key questions to get the information they need. I think the discussion should bring it together in clinical terms.	We received assistance from one of the surgeons on our TEP to make the discussion clinical.
Jeanne-Marie Guise, Oregon EPC	General	The harms of mesh are a big issue in surgery and this does not get highlighted as I would expect. There are issues of erosion into major blood vessels bowel etc. that are important and need to be addressed. I understand that the FDA pulls are likely happening now between draft and final but because this is important saying something explicitly about this part of the process so readers of the draft know you recognize this as an issue. I believe one version of the plug was recalled. Worth checking if the plug you reviewed was recalled.	We contacted members of our TEP or KI, asking the question "Regarding mesh-related complications, what specific complications should the report be sure to mention, by way of background? For example, anecdotally, there is a concern about "erosion into major blood vessels."" One said the concern is erosion into bowel, but not erosion into blood vessel. The other said "Erosion into major blood vessels does not occur and should not be a concern" but did mention chronic pain.
Jeanne-Marie Guise, Oregon EPC	General	Not all harms and benefits are equal some discussion of the implications of these and which may be a bigger deal would be helpful. For example recurrence means repeat surgery which has increased risks due to scar tissue, erosion into vessels could be big blood vessels and be quite dangerous.	We added three paragraphs discussing various outcomes: short-term recovery outcomes, adverse events and hernia recurrence, and long-term pain and QOL.
Jeanne-Marie Guise, Oregon EPC	Methods	The group did a good job describing factors considered in evaluating articles for selection and abstraction (the what aspects)	Thank you for your comments.

Commentator	Section	Comment	Response
Jeanne-Marie Guise, Oregon EPC	Methods	Methods section is missing description of process for each step. Especially given IOM report on standards they should provide detail on the process followed for each step in methods (the how). For example, while they describe features used to select articles they do not say how many reviewed, who were the reviewers, what happened with disagreements etc. This is true for every full text article review, risk of bias assessment, abstraction, SOE, applicability. Particularly of interest is whether investigators/clinicians were part of the dual review.	<p>We added clarification of these methods. We added a new section at the beginning of the Methods section describing the four members of the review team. In various sections of the Methods we now state:</p> <ul style="list-style-type: none"> <li>- Hernia clinicians and surgeons were only involved in the determination of KQs, specific questions for consultation, and the peer review of the document (the Key Informants and the Technical Expert Panel). The review was conducted by a multidisciplinary team of four reviewers, none of whom had expertise in hernia repair, but all had prior experience with systematic review. Below we refer to specific individuals by number: #1 (PhD), #2 (MPH), #3 (MD), and #4 (MD/PhD surgeon).</li> <li>- That abstracts were all reviewed by #1, and a randomly selected 10% of the abstracts were re-reviewed by another team member, with disagreements resolved by consensus.</li> <li>- That full article inclusion decisions were first made by the team member(s) responsible for that KQ, and a 10% randomly selected subset of the articles were rescreened by a second person, with disagreements resolved by consensus.</li> <li>- The extracted datapoints were first performed by the team member(s) responsible for that KQ, and a 10% randomly selected subset of the datapoints were checked by a second person, with disagreements resolved by consensus.</li> <li>- All risk of bias category assignments (Low, Moderate, High) were performed by #1 and #3 independently, with disagreements resolved by consensus.</li> <li>- All strength of evidence category assignments (High, Moderate, Low, Insufficient) were performed by two team members independently, with disagreements resolved by consensus (#1 and the team member(s) responsible for that KQ)</li> <li>- All applicability sections (applicability was not rated on a scale based on the applicability guidance chapter) were written by #4.</li> </ul>

Commentator	Section	Comment	Response
Jeanne-Marie Guise, Oregon EPC	Methods	What they present is a decision tree rather than an analytic framework and should be labeled as such. Given the topic it is appropriate to use decision tree just needs to be labeled as such.	Surgeons do not choose operations in the manner you suggest (i.e. first choose whether open or lap, then choose procedure, then choose mesh etc.). They base it on their prior experience and the patient to be operated on, and the decisions are made simultaneously rather than sequentially. The figure we produced we called an analytic framework because it shows various aspects of the report and the KQs asked, such as populations, treatments, intermediate outcomes, and patient—oriented outcomes, just like the other analytic frameworks we have developed over the past several years.
Jeanne-Marie Guise, Oregon EPC	Results	While the authors should be commended for brief results, however, I think they need to bring some into the paper so that it can be understood on its own without appendix. I think a summary overall of the literature whether studies were conflicting or consistent would be helpful before getting into details.	<p>The paper does contain results, specifically the results of our meta-analyses when they were appropriate, or else the results of individual studies if meta-analyses were not applicable. Each KQ has its own peculiarities in terms of the specific comparisons made and the consistency of results of the studies included for that KQ, thus it would not be possible to state whether results were consistent across KQs.</p> <p>What we did, in an attempt to help the reader get a sense of the totality of the evidence in the full report, is add a new table just before KQ1 with the following headers: KQ, # studies, # patients, study designs and counts, and range of length of follow-up. Hopefully this table will help orient the reader to the totality of what we reviewed (kind of a roadmap) so that they will have a better handle on things when they reach individual KQs.</p>
Jeanne-Marie Guise, Oregon EPC	Results	The first part of each section about study characteristics sometimes is hard to follow because categories are discussed out of context of the study. What I mean by this is for example in KQ1 statements like “In one study, surgeries were performed between 1999 and 2004. The other study did not reported (TYPO) the date range of the surgeries.” ... “One study did not report the funding source, the other was funded by a manufacturer of mesh plug.” This causes the reader to look at the reference and put it all together. Especially when there are only 2 studies discussing their characteristics by study would probably be an easier read.	<p>We have fixed the section accordingly to read “In the multi-center study, surgeries were performed between 1999 and 2004 at three university hospitals and two community clinics. This study was funded by AHRQ and the American College of Surgeons, and the lead author disclosed financial ties with a manufacturer of a mesh plug. The other single-center study conducted at a university hospital did not report the date range of the surgeries or source of funding.”</p> <p>Because we included 151 studies, and many for multiple questions, a study-by-study table would have 175 rows and would be too unwieldy for the main document because it would extend over many pages.</p>

Commentator	Section	Comment	Response
Jeanne-Marie Guise, Oregon EPC	Results	Similarly, I found it somewhat distracting to have the first reference to a table in results be Tables 21-26 and all in Appendix. This was especially magnified as an issue because of above. If there was a summary table describing studies in text the other summary by feature might not have been such a struggle. The text of the paper should be understandable on its own without having to rely on Appendix. As written the reader has to refer to Appendix to get a good understanding of included studies.	See above comment
Jeanne-Marie Guise, Oregon EPC	Results	Certain patient-oriented outcomes such as quality of life should be defined in methods and then used consistently through all sections. For example, return to daily activities, work, pain are usually considered components of QOL. Yet this report separates QOL, return to work, return to daily activities, “long term” (6 months) pain. If the outcome is important but data are lacking this is important to say. Similarly if the outcome is really a subset of QOL like general health as in KQ1 this should be said rather than mod evidence of improved QOL and there should be a discussion framing what this means and if this is clinically important.	Studies almost always reported pain separately from QOL, and also reported RTW/RTDA separately from QOL. Thus, we kept them separate.
Jeanne-Marie Guise, Oregon EPC	Results	I think clinicians would expect TEP and TAPP to be considered separately. Right now it seems to combine TEP and TAPP and considers open or closed against LS in general. This doesn't seem consistent with the way surgeons think about it. They look at what options they have and they compare those. They wouldn't usually combine such different types of surgery extraperitoneal and intraperitoneal approaches. One is much more technically challenging than the other.	We contacted our TEP/KI about this question, asking “Do you think it would be helpful to add two specific comparisons: any open mesh procedure vs. TAPP, and any open mesh procedure vs. TEP? Or perhaps it is better to keep it simple and not add these two comparisons?” Both respondents advised us to keep it simple and not separate TAPP from TEP in the context of KQ2.

Commentator	Section	Comment	Response
Jeanne-Marie Guise, Oregon EPC	Results	I feel uneasy about some of the SOE ratings. For example SOE for KQ1 only 1 study granted it's an RCT and the outcome is not QOL but one aspect of QOL. It may be helpful to consider and present the interpretation of what each level of SOE means and make sure the rating matches up about this: <i>High</i> strength of evidence indicates high confidence in the estimate of effect and that the evidence reflects the true effect; further research is unlikely to change our confidence. <i>Moderate</i> strength of evidence indicates moderate confidence that the evidence reflects the true effect; further research may change our confidence in the estimate and may change the estimate. <i>Low</i> strength of evidence indicates low confidence that the evidence reflects the true effect; further research is likely to change our confidence in the estimate and is likely to change the estimate. <i>Insufficient</i> indicates that evidence is unavailable or does not permit estimation of an effect.	The SOE definitions from the Owens paper are now listed in the revised version of the paper. Our process for determining SOE ratings was detailed in the methods section, and each SOE rating was checked by a second rater. For the QOL example in KQ1, the outcome was general quality of life, not a specific aspect of it. The only one-study problem (i.e. lack of replication) resulted in a downgrade (to Low strength) due to an inability to assess consistency.
Jeanne-Marie Guise, Oregon EPC	Discussion	This section would be more helpful if all the information could be brought together in the way a clinician or patient would think about it rather than by KQ. So for a surgeon they decide to do surgery, their next question is should it be open or LS and which type of LS and then whatever surgery they do which mesh is best. As it is the clinician has to try to link KQs which isn't easy as laid out. If you used something like this even if it wasn't easy to express the relative difference between benefits and harms, clinicians would know and be able to use it (this is just an idea not required).	We received help from one of the surgeons on our TEP to make the discussion clinical.