

Evidence-based Practice Center Systematic Review Protocol

Project Title: Comparative Effectiveness of Surgical Options for Inguinal Hernia

Protocol Posting Date: April 11, 2011

Amendment Date(s) if applicable: August 9, 2011

(Amendments Details—see Section VII)

I. Background and Objectives for the Systematic Review

Inguinal hernia

An inguinal hernia is a protrusion of abdominal contents into the inguinal canal through an abdominal wall defect. A direct inguinal hernia protrudes through the deep inguinal ring, whereas an indirect inguinal hernia protrudes through the internal inguinal ring (and may descend through the inguinal canal). Direct hernias typically develop only in adulthood, and are more likely to recur than indirect hernias.¹ If the hernia is severe enough to restrict blood supply to the intestine, it is termed a strangulated hernia, and immediate corrective surgery is necessary. Most inguinal hernias, however, are less dangerous, and elective surgery is often performed to correct the defect. Symptoms include abdominal pain and a lump in the groin area, which is most easily palpable during a cough. Some inguinal hernias, however, are asymptomatic.²

The lifetime rate of inguinal hernia is 25% in males and 2% in females.³ The risk of inguinal hernia increases with age, and the annual incidence is around 50% by the age of 75.⁴ Approximately two-thirds of inguinal hernias are indirect, and one third are direct.³ Approximately 10% of cases are bilateral.⁵ Recurrence occurs in approximately 1%-5% of cases.⁶

In children, the incidence ranges from 0.8% to 4.4%.⁷ It is ten times more common in boys, and also more common in infants born before 32 weeks gestation (13% prevalence) and infants weighing less than 1,000 grams at birth (30% prevalence).⁷ As in adults, about 10% of cases involve bilateral hernia.⁸

Numerous classification systems have been proposed for groin hernias.⁹ One commonly used system was introduced by Nyhus in 1993.¹⁰ This system employs several clinical factors including direct/indirect, degree of enlargement of the internal inguinal ring, and degree of posterior wall weakness. Specifically, it comprises six types of increasing severity: 1) indirect inguinal hernia with a normal internal ring; 2) indirect inguinal hernia with an enlarged internal ring; 3a) direct inguinal hernia; 3b) indirect inguinal hernia causing posterior wall weakness; 3c) femoral hernia; 4) recurrent hernia.¹⁰ (This review will not involve femoral hernias due to the different patient populations and pertinent treatments) Stoppa (1998)¹¹ proposed that aggravating factors (such as obesity or abdominal distension) should upgrade the patient by one Nyhus level.¹¹ Higher

severity generally means a higher risk of recurrence, and an appropriate classification may support the management approach.

Surgical repair of inguinal hernia

Surgical repair of hernias is the most commonly performed general surgical procedure in the United States.¹² In 2003, an estimated 770,000 surgical repairs of inguinal hernia were performed.¹² These repairs are typically performed on an outpatient basis (87% in 1996).¹² Such a large volume of procedures suggests that even modest improvements in patient outcomes would have a substantial impact on population health.¹³

The primary goals of surgery include preventing recurrence of the hernia, returning the patient to normal activities quickly, and minimizing postsurgical discomfort and the adverse effects of surgery. The various surgeries present different constellations of benefits and risks, which presents some clinical uncertainty in the choice among approaches. Balancing these factors is a difficult yet critical process in an effort to make the best possible medical decisions.

Some patients with inguinal hernias may not be in pain or limited in any way by the hernia. For these patients, surgery may not be necessary. One of the Key Questions in the evidence review will be a comparison between surgical and non-surgical approaches to the management of pain-free inguinal hernias.

Surgical procedures for inguinal hernia repair generally fall into three categories: open repair without the use of mesh (i.e., sutured), open repair with a mesh, and laparoscopic repair with a mesh. Within each of these categories, several specific procedures have been employed. Until the 1980s, open suture repair was the standard, however, the resulting tension along the suture line yielded relatively high rates of recurrence and patient discomfort. Non-sutured “tension-free” surgical mesh gained in popularity, and many specific open procedures were used, such as Lichtenstein repair and mesh plug repair. One author estimated that in 2003, 93% of groin hernia repairs involved the use of a mesh, and of these, about three-fourths of these repairs involved either Lichtenstein or mesh plug.¹² Thus, the most important relevant questions about current hernia repair all involve comparisons between different mesh-based approaches.

Almost all surgical procedures in adults involved the use of a synthetic mesh to cover the defect. However, mesh is not recommended for use in pediatric inguinal hernia due to concerns about the risk of inflammatory reactions, damage to the vas deferens and/or testes, and infertility.⁷

More recently, two laparoscopic approaches using a mesh (transabdominal pre-peritoneal repair or TAPP, and totally extraperitoneal repair or TEP) have seen increased use.¹⁴ Laparoscopic approaches have the potential for shortening recovery time and reducing some postoperative morbidities.¹⁵ They also may be associated with longer operation times and a relatively long learning curve. The TEP was introduced after the TAPP due to concerns about a possible increase risk of internal organ damage within the peritoneum.

Previous research has shown that the repair of a recurrent inguinal hernia is subject to a greater risk of additional recurrence.¹⁶ Further, bilateral inguinal hernia is subject to a greater recurrence risk than unilateral inguinal hernia.¹⁷ These increased risks may be due to certain anatomical difficulties that complicate the surgical approach in these types of patients.¹⁸ Some clinicians have suggested that laparoscopic approaches are better suited to recurrent and bilateral hernias, and in Key Question 2 (see below) we delineate separate comparisons for primary, bilateral and recurrent hernia.

Specific aspects about mesh repair that may influence outcomes are the type of mesh (e.g., lightweight or heavyweight), whether the mesh fixation is used, and if so, whether fixation is accomplished with sutures or glue. These mesh-specific issues are covered by specific Key Questions (see Key Questions below).

Different procedures often require different methods for anesthesia. Some forms of open mesh repair can be performed with local anesthesia, whereas laparoscopic techniques such as TAPP often require general anesthesia.¹⁹ Two key post-surgical morbidities are surgical site infection²⁰ and chronic pain.²¹ Sanabria et al. (2007)²⁰ found that without prophylactic antibiotics, surgical site infections occurred in 2.9% of patients having undergone mesh repair, as compared to only 1.4% among those who did receive prophylactic antibiotics. Regarding chronic pain, Nienhuis et al. (2007)²¹ estimated that pain lasting beyond three months postoperatively occurs in 11% of patients undergoing mesh repair.

In terms of settings, most hernia surgeries are performed not in specialized hernia centers, but rather by general surgeons who also perform many other types of surgeries.¹⁹ It is generally recognized that the surgical repair of inguinal hernia is a highly specialized skill, and patients receiving care from more experienced surgeons may be better than patients receiving care from less experienced surgeons. The evidence review will specifically examine evidence on the association between surgical experience and hernia recurrence (see Key Questions below).

Given the clinical uncertainty, a systematic review of the existing evidence on comparative effectiveness will help inform important medical decisions about surgical options for inguinal hernia. The findings of the review may impact clinical decisions by patients and surgeons, treatment recommendations by professional societies, purchasing decisions by hospitals, and coverage decisions by payers.

II. The Key Questions

Ten initial Key Questions had been posted for public comment on the Web site of the Effective Health Care Program from 10/25/2010 to 11/22/2010. Several changes were made to the Key Questions based on input from Key Informants and public comment.

Initially there had been two questions pertaining to non-mesh procedures, but given the relative rarity of such procedure in current clinical practice, as well as the existence of other substantive issues that are more important to address (i.e., open mesh vs. laparoscopic mesh in bilateral hernia or, recurrent hernia), we removed these two questions. Instead, we chose to expand the question comparing open and laparoscopic mesh procedures to specifically address three separate populations: primary hernia,

bilateral hernia, and recurrent hernia. Also, a previous question had combined type-of-mesh comparisons and type-of-fixation-method questions, and we decided to separate these into two Key Questions.

The final list of Key Questions is below. The subsequent table (Table 1) clarifies the scope of each Key Question using the PICOTS format (Patients, Interventions, Comparisons, Outcomes, Timing, Settings).

Among adults with pain-free primary inguinal hernias:

1. Does hernia repair differ from watchful waiting in patient-oriented effectiveness outcomes and/or adverse events?

Among adults with painful inguinal hernias without incarceration/strangulation:

2. Does open hernia repair with a mesh differ from laparoscopic hernia repair with a mesh in patient-oriented effectiveness outcomes and/or adverse events?
 - a. For primary hernias?
 - b. For bilateral hernias?
 - c. For recurrent hernias?
3. Do different open mesh-based repair procedures (e.g., Lichtenstein, mesh plug) differ in patient-oriented effectiveness outcomes and/or adverse events?
4. Do different laparoscopic mesh-based repair procedures (e.g., transabdominal pre-peritoneal repair, totally extraperitoneal repair) differ in patient-oriented effectiveness outcomes and/or adverse events?
5. Do different mesh products (lightweight, heavyweight) differ in patient-oriented effectiveness outcomes and/or adverse events?
6. Do different mesh fixation methods (e.g., no fixation, sutures, glue) differ in patient-oriented effectiveness outcomes and/or adverse events?
7. For each type of laparoscopic mesh repair, what is the association between surgical experience and hernia recurrence?

Among pediatric patients (age 21 or under):

8. For a possible contralateral hernia, does same-operation repair/exploration differ from watchful waiting in patient-oriented effectiveness outcomes and/or adverse events?
9. Does open hernia repair without a mesh differ from laparoscopic hernia repair without a mesh in patient-oriented effectiveness outcomes and/or adverse events?

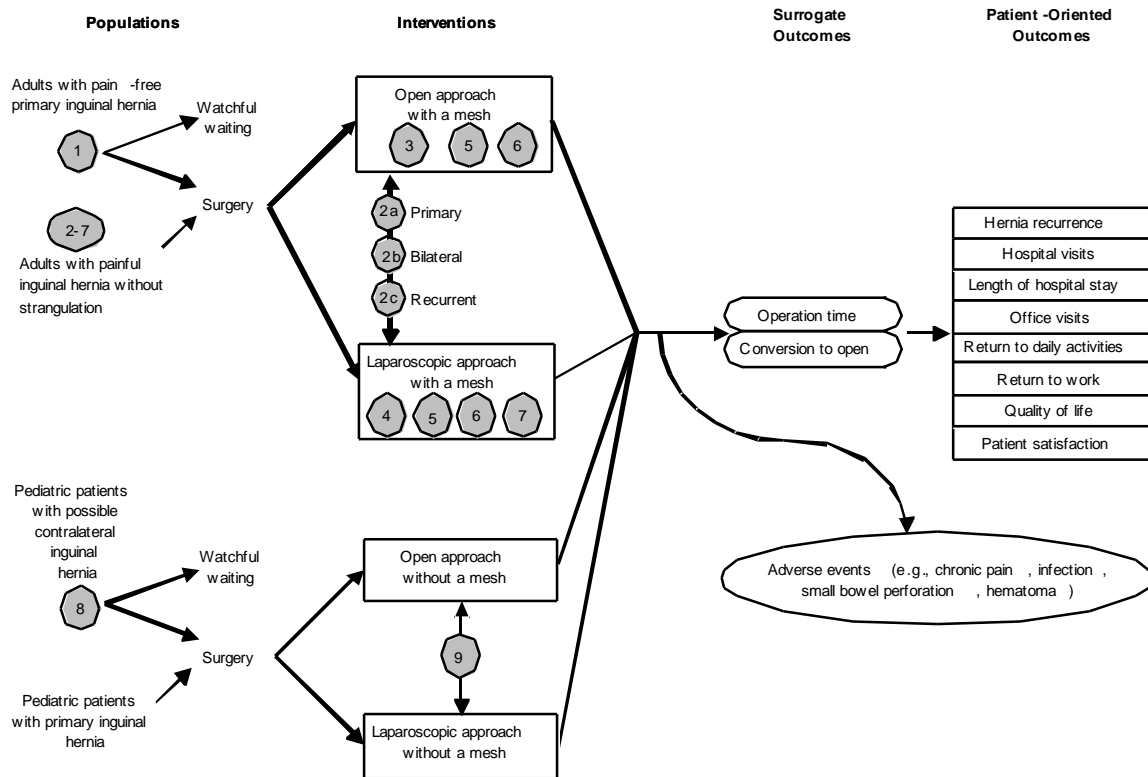
Table 1. Scope of Key Questions: Patients, Interventions, Comparators, Outcomes, Timing and Settings

For each Key Question, the table below depicts the population(s), intervention(s), comparator(s), outcome(s), timing, and setting.

	Key Question											
	1	2a	2b	2c	3	4	5	6	7	8	9	
Age	Adults									Pediatric patients		
Condition	Inguinal hernia without incarceration/strangulation									Possible contra-lateral inguinal hernia without incarceration/strangulation	Inguinal hernia without incarceration/strangulation	
Pain status	Pain-free	Painful									Pain-free or painful	
Unilateral or bilateral	Unilateral or bilateral		Bilateral	Unilateral or bilateral								
Primary or recurrent	Primary		Primary or recurrent	Recurrent	Primary or recurrent							
Intervention	Watchful waiting	Open mesh			Open mesh variant 1	Lap. mesh variant 1	Mesh variant 1	Fixation variant 1	Lap. mesh, with different levels of surgical experience	Watchful waiting	Open no mesh	
Comparator	Any surgery	Lap. mesh			Open mesh variant 2	Lap. mesh variant 2	Mesh variant 2	Fixation variant 2		Surgical exploration or repair	Lap. no mesh	
Outcomes	Patient-oriented effectiveness outcomes and adverse events (see list of outcomes at the bottom of this table)											
Timepoints	For hernia recurrence, quality of life and patient satisfaction, we will include follow-up data of at least 6 months+ on at least 50% of enrolled patients will be included. For other outcomes, we will include any timepoint for which data were reported on at least 50% of enrolled patients.									Any timepoint for which data were reported on at least 50% of enrolled patients.		
Setting	Any setting											

Patient-oriented effectiveness outcomes include: hernia recurrence, hospital visits, length of hospital stay, office visits, return to daily activities, return to work, quality of life, patient satisfaction. Adverse events include any unintended effects such as chronic pain, infection, small bowel perforation, hematoma, etc. We did not examine certain surrogate outcomes, namely operation time and the rate of conversion from laparoscopic to open surgery, because of their reduced importance in addressing the Key Questions.

III. Analytic Framework



IV. Methods

A. Criteria for Inclusion/Exclusion of Studies in the Review

The inclusion criteria are listed below in separate categories pertaining to 1) publication type, 2) study design, 3) patient characteristics, 4) treatment characteristics, and 5) data.

Publication criteria

1. Publication must be a full article; abstracts alone will not be included, because they do not include sufficient details about experimental methods to permit an evaluation of study design and conduct, and they also may only contain a subset of the measured outcomes.^{22,23} Abstracts of randomized studies that did not subsequently appear as full articles will be flagged for possible evidence of publication bias.
2. To capture the most relevant data, we will include studies published on or after 1/1/1990. Studies published before 1990 likely describe procedures no longer being used commonly, or outcomes that are not likely to be predictive of current outcomes. For our evidence searches, we anticipate that the last search will be conducted on 4/1/2011, and an updated search will be conducted while the report is under review.
3. To avoid double-counting of patients, when several reports of overlapping patients were available, only outcome data from the report with the largest number of patients will be included. We did include the data when a smaller report provided data on an outcome that was not provided by the largest report.
4. Studies must be published in English. Moher et al. have demonstrated that exclusion of non-English language studies from meta-analyses has little impact on the conclusions drawn. Juni et al. found that non-English studies typically were of lower methodological quality and that excluding them had little effect on effect size estimates in the majority of meta-analyses they examined. Although we recognize that in some situations exclusion of non-English studies could lead to bias, we believe that the few instances in which this may occur do not justify the time and cost typically necessary for translation of studies to identify those of acceptable quality for inclusion in our reviews.^{24,25} Due to the prevalence of non-English-language studies of inguinal hernia repair, however, we will examine the English abstracts of these studies in an attempt to assess to degree of bias resulting from their exclusion.

Study design criteria

5. For questions comparing interventions (i.e., all questions except the one on surgical experience), the study must either randomize patients to

treatments, or use an analytic method to address selection bias, such as baseline matching on multiple characteristics, propensity scoring, or other analytic approach. Studies with large differences at baseline between groups (regardless of whether they were randomized), or entailed confounding by indication, will be excluded. Studies comparing meshes or mesh fixation methods must not confound results by differences in surgical procedures for inserting the mesh. For the Key Question on surgical experience, a control group will not be required, however, the study must provide data on the relation between surgical experience and outcomes. The definition of surgical experience must be specific to laparoscopic mesh hernia repair, not simply a measure of general experience such as surgeon age.

6. Study can be prospective or retrospective, but retrospective studies must use consecutive enrollment (or enrollment of a random sample of eligible participants)
7. The treatments being compared must be administered during the same time period, so that any observed difference between treatment outcomes are not attributable to differential timeframes.

Patient criteria

8. To be included for a given Key Question, the study must have provided data for which at least 85% of the patients had the condition specified in the Key Question. For example, for Key Question 2a, we only included datapoints for which at least 85% of the patients were adults with painful primary inguinal hernia without incarceration/strangulation.
9. We used a flexible definition of “adulthood”; we defined “adults” as anyone aged 18+, and we defined the “pediatric population” as anyone aged 21 or less. This means that studies enrolling those aged 18/19/20 years could have been included as either an adult study or a pediatric study, depending on the average age of those enrolled.

Treatment criteria

10. The study must provide sufficient information about the treatments for one to determine that the data address one of the Key Questions.
11. The study must not describe a specialized and novel hernia repair that has not been widely practiced by other surgeons. This is to maintain the focus of the report on the most common types of repair.
12. The hernia repair must not be performed simultaneously with another operation (e.g., prostatectomy). Surgical complications of combined operations make it difficult to isolate aspects of the hernia repair itself.

Data criteria



13. The study must report data at least one of the included outcomes for at least one of the Key Questions.
14. Outcome data must not rely on retrospective recall (for example, in an interview long after the procedure had been performed) because such outcomes may not accurately reflect patients' experiences.
15. For some outcomes in the adult population, we will include datapoints at least six months after treatment (hernia recurrence, quality of life, and patient satisfaction). For all other outcomes (and in the pediatric population), there will be no minimum follow-up.
16. We will include datapoints capturing at least 10 patients with the condition of interest who represented at least 50% of eligible enrolled patients.

B. Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions

Literature searches will be performed by Medical Librarians within the EPC Information Center, and will follow established systematic search protocols. The searches will be led by the Director of Health Technology Assessment/EPC Information Center.

Consistent with our evidence-based searching protocol, for all key questions, the following databases will be searched on the OVID SP platform, utilizing the one search and de-duplication features: MEDLINE, PreMEDLINE, and Embase. The Cochrane Library, including the Central Register of Controlled Trials, the Cochrane Database of Methodology Reviews, and the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects, Health Technology Assessment Database, and the U.K. National Health Service Economic Evaluation Database, will also be searched for unique reviews, trials, economic analyses, and technology assessments.

Search terms will be identified by: 1) reviewing relevant systematic reviews on similar topics that are identified by members of the research staff, 2) reviewing how other relevant studies are indexed, their subject heading terms, and their keywords, and 3) reviewing MeSH and Emtree indexes for relevant and appropriate terms. After reviewing these, a combination of subject headings and keywords will be identified. Search strategies will be developed using these terms. Once developed, search strategies will be reviewed by senior research analyst(s) and the Director of the Health Technology Assessment/EPC Information Center. Per the request of technical staff, no limits on language will be applied by the search (as recommended on page 9 of the Relevo and Balshem chapter of the EPC guide)²⁶ and search dates have been established as 1990-2011. A study design filter will be applied to retrieve systematic reviews and clinical trials.

Literature search results will initially be reviewed by the Medical Librarian. Using the key questions and inclusion/exclusion criteria identified

by senior research analyst(s), the Medical Librarian will assess relevancy and retrieve results. Feedback from the senior research analyst(s) and the Director of the Health Technology Assessment/EPC Information Center – including details regarding gaps in the search strategy, as well as articles [identified by the senior research analyst(s)] not retrieved by the searches – will be integrated into the search strategy using key terms and subject headings. The updated strategy will be re-run in all identified databases. Additional results will be scanned and relevancy will be assessed by the Medical Librarians. New results will be downloaded and forwarded to senior research analyst(s) for review. Hand searches of reference lists in identified articles will also be reviewed for possible inclusion. The search will be updated during the peer review period of the draft report.

Articles will be reviewed first at the abstract level by a single person, and any articles possibly meeting the inclusion criteria for at least one Key Question will be obtained for full review. A randomly selected 10% of the articles excluded at the abstract level will be screened by a second person to ensure that no articles are excluded inappropriately. If this process reveals any studies that the team agrees were mistakenly excluded, then all of the other articles excluded at the abstract level will be screened by a second person.

Full articles meeting the inclusion criteria will be retained for abstraction of information on general study characteristics, patient characteristics, treatment characteristics, risk-of-bias items, and outcome data (see next section). Separately for each person who screened full articles, we will randomly select 10% of the articles excluded at the full article level and have them screened by a second person to ensure that no articles are excluded inappropriately. If this process reveals any studies that the team agrees were mistakenly excluded, then all of the other articles excluded at the full article level will be screened by a second person.

C. Data Abstraction and Data Management

We plan to extract study information into spreadsheets in Microsoft Excel.TM An initial period of pilot testing will enhance accuracy, which will be assessed by three team members. After the initial period of pilot testing for data abstraction, we will do single reviewer abstraction of data. A randomly selected 10% of the abstracted datapoints will be abstracted by a second person (separately for each person who abstracted data). Also, due to the possibility of subjective interpretation, the risk-of-bias items will be judged in duplicate, and discrepancies will be resolved with discussion. The overall categories of information to be obtained from each study will include:

- General study characteristics. Author, publication year, country, study design, dates of patient enrollment, length of follow-up, funding source, which Key Question(s) the study addressed

- Patient characteristics. Number of enrolled patients, age, sex, hernia type(s), pre-surgical pain level, pre-surgical quality-of-life scores, pre-surgical functional activity scores, unilateral/bilateral, primary/recurrent, prior hernia repairs, hernia classification (e.g., Nyhus classification, if used).
- Treatment characteristics. Procedure category, specific procedure, specific mesh (if applicable), fixation method (if applicable), number of surgeons performing the procedures in the study, surgeons' prior experience with the repair procedures performed, surgical setting (i.e., specialized hernia center or general surgery), type of anesthesia
- Risk-of-bias items. See the next section.
- Outcome data. Study methods of follow-up for data collection will be extracted, as well as the timepoint(s) of evaluation. For each included outcome, we will extract the number of patients contributing data to each included timepoint. We will extract the numerical data necessary for us to compute an effect size and its standard error for all included outcomes for each study. These may include means, standard deviations, counts, proportions, results of authors' statistical tests, or other statistical details, depending on what is reported.

Multiple publications of the same study (e.g., publications reporting subgroups, other outcomes, or longer follow-up) will be identified by examining author affiliations, study designs, enrollment criteria, and enrollment dates.

D. Assessment of Methodological Risk-of-Bias of Individual Studies

As stated above, due to the possibility of subjective interpretation, this step will be performed by two extractors for each study, and discrepancies will be resolved by consensus. We will assess the risk-of-bias (i.e., internal validity) separately for each outcome and each timepoint of each study. The reason for outcome specificity is that some subjective outcomes are more susceptible to bias than other outcomes. The reason for timepoint specificity is that longer follow-up often results in attrition or right-censoring, which may yield patients who are somewhat different than the full set of enrolled patients, and also may introduce a systematic difference between the groups being compared.

For all studies with control groups (regardless of whether patients were randomly assigned to groups), we will assess risk-of-bias using the items in the table below. All but one of these items were selected from a pool of items typically used by this EPC for technology assessments. The last question (on surgeon's prior number of procedures) was devised specifically for this project.

Item	Comments
------	----------

Item	Comments
#1. <i>Were patients randomly assigned to the study's groups?</i>	
#2. <i>Was there concealment of group allocation?</i>	
#3. <i>For non-randomized trials, did the study employ any other methods to enhance group comparability?</i>	All included studies will have a Yes for this question, because non-randomized studies are required to use an analytic control to address selection bias.
#4. <i>Was the process of assigning patients to groups made independently from physician and patient preference?</i>	
#5. <i>Did patients in different study groups have similar levels of performance on the outcome of interest at the time they were assigned to groups?</i>	This will only be used for outcomes that have a baseline measure (e.g., quality-of-life).
#6. <i>Were the study groups comparable for all other important factors at the time they were assigned to groups?</i>	
#7. <i>Was the comparison of interest prospectively planned?</i>	
#8. <i>If patients received ancillary treatment(s), was there a $\leq 5\%$ difference between groups in the proportion of patients receiving each specific ancillary treatment?</i>	If ancillary treatments differed substantially between groups, then the study will be excluded.
#9. <i>Were the two groups treated concurrently?</i>	All included studies will have a Yes for this question, because it is a requirement for inclusion.
#10. <i>Were those who assessed the patient's outcomes blinded to the group to which the patients were assigned?</i>	
#11. <i>Was the outcome measure of interest objective and was it objectively measured?</i>	<p>The following will always be considered objective outcomes: adverse events, hospital stay, office visits, and hospital visits</p> <p>The following will always be considered subjective outcomes: return-to-work, return-to-daily activities, quality of life (QOL), functional status (FCT), and patient satisfaction</p> <p>For hernia recurrence, we will consider it objective if the patient was physically assessed by a clinician for recurrence. If instead the recurrence assessment is based on patient report, this will be considered subjective.</p>
#12. <i>Was there $\leq 15\%$ difference in the length of follow-up for the two groups?</i>	This item will only be used for the following outcomes: hernia recurrence, adverse events, QOL, FCT, office visits, hospital visits
#13. <i>Did $\geq 85\%$ of enrolled patients provide data at the timepoint of interest?</i>	

Item	Comments
#14. Was there a $\leq 15\%$ difference between groups in the percentage of patients provided data at the timepoint of interest?	
#15. For questions comparing two procedures, did the two groups' surgeons have similar numbers of prior operations performing the procedure they performed in the study?	

Each of these items will be answered Yes, No, or Not reported (NR).

Note that funding source is not listed as a risk-of-bias item. This is because funding source is an indirect measure of risk-of-bias, and also because we will consider it at a later stage in the assessment of the strength of evidence, specifically when assessing the possibility of publication bias or selective outcome reporting bias (see the section below entitled “Grading the Evidence for Each Key Question”).

We will categorize each study as Low risk-of-bias, Medium risk-of-bias, or High risk-of-bias using the following method:

- In order to be considered Low risk-of-bias, the study must meet ALL of the following conditions:
 - Randomized (#1)
 - Concealment of allocation (#2) OR blinded outcome assessors (#10) or both
 - Good baseline comparability for both outcome (#5) and other patient characteristics (#6)
 - Good baseline comparability on surgeons' number of prior operations performing the compared procedures (#15)
 - If NOT blinded outcome assessors (#10) (or NR blinded outcome assessors), then it was an objective outcome (#11)
 - $\leq 15\%$ difference in length of follow-up between groups (#12)
 - $\geq 85\%$ of enrolled patients provided data to this timepoint (#13)
 - $\leq 15\%$ difference in data provision rates to this timepoint (#14)
- In order to be considered High risk-of-bias, the study must meet AT LEAST TWO of the following criteria:
 - Process of assigning patients to groups NOT made independently from physician and patient preference (#4)
 - Not good baseline comparability for either the outcome (#5) or other patient characteristics (#6)
 - Retrospective (#7)
 - Difference in ancillary treatments $\geq 5\%$ (#8)

- Not a blinded outcome assessor (#10), AND a subjective outcome (#11)
- In order to be considered Medium risk-of-bias, the study neither meets the criteria for Low risk-of-bias nor the criteria for High risk-of-bias.

E. Data Synthesis

We plan to perform meta-analysis wherever appropriate and possible. This decision will depend on the judged clinical homogeneity of the different study populations, co-interventions, and outcomes as well as what is reported by those studies. Subgroup analyses may be performed based on key identified aspects of the studies, such as the specific procedure (e.g., TAPP or TEP), the specific mesh (e.g., lightweight or heavyweight), or study design (e.g., randomization or not).

For each outcome in the review, an important consideration is the smallest difference between groups that can still be considered clinically significant (MCSD or minimum clinically significant difference). This definition aids interpretation in two main ways: 1) to determine whether a statistically significant difference is clearly clinically significant, and 2) to determine whether a statistically nonsignificant difference is small enough to exclude the possibility of a clinically significant difference. Ideally, the MCSD is decided for each outcome before the results of the studies are known.

After hernia repair, a key outcome is hernia recurrence, and for this outcome, we define the MCSD as 3 percentage points (e.g., 1% vs 4% for two separate treatments). This was based on statements in two multicenter trials (the VA trial²⁷ and the MRC trial)²⁸ that such a difference is clinically meaningful. For other anticipated outcomes, we used the following approaches concerning the definitions of minimum clinical significance:

- Length of hospital stay: 1 day difference between groups
- Number of hospital visits/number of office visits: 20% difference between groups (e.g., means of 5 visits and 4 visits)
- Time to return to daily activities/return to work: If reported as a continuous measure, 1 week difference between groups. If reported as a dichotomous measure, no a priori definition, and we may need to consult with the TEP to determine the MCSD
- Quality of life: 5% of the range of the scale (e.g., 5 points on the SF-36 which ranges from 0-100)²⁹
- Patient satisfaction: No a priori definition used, because it is not clear how studies will report this outcome, and we may need to consult with the TEP to determine the MCSD.
- Pain: If reported as a continuous measure, 20% of the range of the scale (e.g. 20 points on the VAS which ranges from 0-100).³⁰ If

reported as a dichotomous measure, we did not set an a priori definition, and we may need to consult with the TEP to determine the MCSD.

- Other adverse events: No a priori definition, and we may need to consult with the TEP to determine different MCSDs for different adverse events

If meta-analysis is deemed appropriate and possible for a given comparison and a given outcome, we will compute effect sizes and standard errors using standard methods, and we will perform DerSimonian and Laird random-effects meta-analysis³¹ using Comprehensive Meta-Analysis (CMA) software (Biostat Inc., Englewood, NJ).³²

For rare events (rates <1%), we will use either the Peto or Mantel-Haenszel odds ratios, as recommended by page 8 of the chapter entitled “Conducting Quantitative Synthesis When Comparing Medical Interventions” in the EPC guide³³ and supported by the simulation studies of Sweeting³⁴ and Bradburn.³⁵

To best interpret the findings, we will define the minimum clinically significant difference for each outcome using information in the hernia literature, general information on that outcome, input from the TEP, and/or standards suggested by Cohen.³⁶ This will help determine whether the between-groups difference is clinically significant, or whether a statistically non-significant finding is sufficiently precise to rule out the possibility of a clinically significant difference.

To measure heterogeneity, we will use I^2 and tau. Both are used because I^2 can increase simply by increasing the numbers of patients in the studies (whereas tau is a more direct measure of heterogeneity),³⁷ but tau is more difficult to interpret because its scale is different for different effect sizes. Heterogeneity measures will be used to determine whether a meta-regression will be attempted. If heterogeneity is present and there are at least five studies, we will perform a meta-regression in an attempt to explain the heterogeneity. To control for spurious findings, meta-regressions will employ the permutation test³⁸ in Stata (Stata Corporation, College Station, TX).³⁹

In determining the factors that may explain heterogeneity, we plan to focus primarily on meta-regressions to study-level predictors, rather than patient-level predictors, because the latter are subject to the ecological fallacy, as suggested in the chapter entitled “Conducting Quantitative Synthesis When Comparing Medical Interventions” in the EPC guide.³³ However, given their importance in this context, we may utilize some aggregated patient-level variables, and plan to interpret these results with caution. Examples of predictors to be included are surgical setting (specialized hernia center or not), number of prior procedures performed by the study’s surgeons, country of origin, mean age, percentage of patients with bilateral hernia, percentage of



patients with recurrent hernia, and risk-of-bias variables such as randomization to groups and baseline group comparability.

The final report will present findings in the following manner, organized by Key Question. First, if the evidence permits an answer to the Key Question, this answer will be stated first, in boldface, along with the rating of the evidence supporting that conclusion. If the evidence is insufficient to answer the question, this will be stated first. The outcomes will be presented in the following order: hernia recurrence, hospital visits, length of hospital stay, office visits, return to daily activities, return to work, quality of life, patient satisfaction, and adverse events (which includes effects such as chronic pain after surgery, infection, etc.). If space constraints require a prioritization of these outcomes, the priority list will include four outcomes: hernia recurrence, quality of life, patient satisfaction, and adverse events. However, full results of all included outcomes will be available in appendices.

F. Grading the Evidence for Each Key Question

We will implement the rating system described by Owens et al. (2009)⁴⁰ (also part of the EPC guide)²⁶ and use many of the principles described by Treadwell et al. (2006).⁴¹ This will involve a rating for confidence in the conclusion about the *direction* of the effect. Each evidence rating will be either High, Moderate, Low or Insufficient, and the process for combining domain scores into an overall rating will be systematic, thereby precluding the need for duplication of effort (as noted by Owens et al. (2009)).⁴⁰ A rating of Insufficient will apply to evidence bases that involve no studies, or where the evidence does not support a conclusion (e.g., when the uncertainty is so high that the evidence cannot discriminate among conflicting conclusions). We will be fully transparent about the process that produces the ratings, so that readers can replicate the process, or use the individual domains to arrive at their own overall ratings.

The Owens et al. (2009)⁴⁰ article describes eight domains to be considered, included risk-of-bias, consistency, directness, precision, magnitude of effect, publication bias, dose-response association, and all plausible confounders would reduce the effect. Judgments concerning applicability are not included in the rating of the strength of evidence (but see the next section on how we will assess applicability in this report).

We plan to provide evidence ratings for hernia recurrence, quality-of-life, patient satisfaction, and adverse events. Depending on time constraints of the review, we may grade additional outcomes as well. For the question on the association between surgical experience and hernia recurrence, no grading system yet exists for grading this type of evidence (i.e., on association), therefore we do not plan to rate that evidence.



G. Assessing the Applicability of the Evidence for Each Key Question

The applicability of the evidence involves five aspects: 1) patients, 2) interventions, 3) comparisons, 4) outcomes, and 5) settings. In order to define optimally the target population, we will search for any large population-based studies conducted of patients receiving treatment in the United States. The typical interventions, comparisons, outcomes (e.g., typical hernia recurrence rates), and settings of care will also be used to more clearly specify the most applicable study characteristics (i.e., most typical of hernia care in the United States). A large portion of the inguinal hernia literature has been conducted outside the United States, and this literature may have limited applicability, because other countries may differ from the U.S. in any of the five aspects listed above.

Regarding patients, this report has three general target populations:

- 1) adults with a pain-free inguinal hernia not strangulated/incarcerated;
- 2) adults with a painful inguinal hernia not strangulated/incarcerated;
- 3) pediatric patients with an inguinal hernia not strangulated/incarcerated.

We hope to define this more precisely with respect to adult ages, pediatric ages, and specific types of inguinal hernia (i.e., direct or indirect). For interventions and comparisons, a key consideration is the prior experience of the surgeons performing the procedures. If the surgeons performing the procedures had more prior experience with those procedures than typical surgeons in clinical practice, this would limit the applicability of the findings. For outcomes, if the study reported hernia recurrence rates that were much different from typical rates (e.g., 10% instead of 2%), the between-group comparison (e.g., 8% vs. 12%) may be inapplicable to typical practice. Finally, regarding settings, if the study settings were atypical (e.g., the study surgeons practiced in specialized hernia care centers), the findings may not be applicable either.

V. References

1. Simons MP, Aufenacker T, Bay-Nielsen M, et al. European Hernia Society guidelines on the treatment of inguinal hernia in adult patients. *Hernia* 2009;13:343-403.
2. Fitzgibbons RJ Jr, Giobbie-Hurder A, Gibbs JO, et al. Watchful waiting vs repair of inguinal hernia in minimally symptomatic men: a randomized clinical trial. *JAMA* 2006 Jan 18;295(3):285-92. PMID: 16418463
3. Nicks BA, Askew K. Hernias. In: eMedicine [database online]. Omaha (NE): eMedicine.com; 2010 Jan 25 [accessed 2010 Jul 14]. Available: <http://emedicine.medscape.com/article/775630-overview>.
4. Inguinal hernia: epidemiology. [database online]. San Mateo (CA): Epocrates, Inc.; 2010 [accessed 2010 Jul 14]. [1 p]. Available: <https://online.epocrates.com/noFrame/showPage.do?method=diseases&MonographId=723&ActiveSectionId=23>.
5. Schneider E. Inguinal hernia: excerpt from The 5-minute pediatric consult. [internet]. Health Grades Inc.; 2008 [accessed 2011 Jan 26]. [6 p]. Available: http://www.wrongdiagnosis.com/i/inguinal_hernia/book-diseases-20a.htm.
6. Sherwinter DA, Lavotshkin S. Hernia inguinal repair, open: treatment & medication. [internet]. eMedicine; [updated 2009 Jul 24]; [accessed 2011 Jan 26]. [23 p]. Available: <http://emedicine.medscape.com/article/1534281-treatment>.
7. Brandt ML. Pediatric hernias. *Surg Clin North Am* 2008 Feb;88(1):27-43, vii-viii. PMID: 18267160
8. Hebra A. Pediatric hernias. [internet]. eMedicine; [updated 2010 Feb 25]; [accessed 2011 Jan 26]. [23 p]. Available: <http://emedicine.medscape.com/article/932680-overview>.
9. Zollinger RM Jr. Classification systems for groin hernias. *Surg Clin North Am* 2003 Oct;83(5):1053-63. PMID: 14533903
10. Nyhus LM. Individualization of hernia repair: a new era. *Surgery* 1993 Jul;114(1):1-2. PMID: 8356511
11. Stoppa R. Hernias of the abdominal wall. In: Chevrel JP, editors. *Hernias and surgery of the abdominal wall*. Berlin: Springer; 1998. p. 171-277.
12. Rutkow IM. Demographic and socioeconomic aspects of hernia repair in the United States in 2003. *Surg Clin North Am* 2003 Oct;83(5):1045-51, v-vi. PMID: 14533902
13. Zhao G, Gao P, Ma B, et al. Open mesh techniques for inguinal hernia repair: a meta-analysis of randomized controlled trials. *Ann Surg* 2009 Jul;250(1):35-42. PMID: 19561484
14. Wake BL, McCormack K, Fraser C, et al. Transabdominal pre-peritoneal (TAPP) vs totally extraperitoneal (TEP) laparoscopic techniques for inguinal hernia repair. In: *Cochrane Database of Systematic Reviews* [internet]. Issue 1. Hoboken (NJ): John Wiley & Sons, Ltd.; 2005 [accessed 2010 Jul 2]. [Art. No.: CD004703]. Available: DOI: 10.1002/14651858.CD004703.pub2. PMID: 15674961
15. Jenkins JT, O'Dwyer PJ. Inguinal hernias. *BMJ* 2008 Feb 2;336(7638):269-72. PMID: 18244999
16. Eklund A, Rudberg C, Leijonmarck CE, et al. Recurrent inguinal hernia: randomized multicenter trial comparing laparoscopic and Lichtenstein repair. *Surg Endosc* 2007 Apr;21(4):634-40. PMID: 17364153
17. Wara P, Bay-Nielsen M, Juul P, et al. Prospective nationwide analysis of laparoscopic versus Lichtenstein repair of inguinal hernia. *Br J Surg* 2005 Oct;92(10):1277-81. PMID: 16003727
18. Kocijan R, Sandberg S, Chan YW, et al. Anatomical changes after inguinal hernia treatment: a reason for chronic pain and recurrent hernia. *Surg Endosc* 2010 Feb;24(2):395-9. PMID: 19551430
19. Jacobs DO. Mesh repair of inguinal hernias--redux. *N Engl J Med* 2004 Apr 29;350(18):1895-7. PMID: 15107484

Source: www.effectivehealthcare.ahrq.gov

Published Online: August 15, 2011

20. Sanabria A, Dominguez LC, Valdivieso E, et al. Prophylactic antibiotics for mesh inguinal hernioplasty: a meta-analysis. *Ann Surg* 2007 Mar;245(3):392-6. PMID: 17435546
21. Nienhuijs S, Staal E, Strobbe L, et al. Chronic pain after mesh repair of inguinal hernia: a systematic review. *Am J Surg* 2007 Sep;194(3):394-400. PMID: 17693290
22. Narine L, Yee DS, Einarson TR, et al. Quality of abstracts of original research articles in CMAJ in 1989. *CMAJ* 1991 Feb 15;144(4):449-53. PMID: 1993292
23. Pitkin RM, Branagan MA, Burmeister LF. Accuracy of data in abstracts of published research articles. *JAMA* 1999 Mar 24-31;281(12):1110-1. PMID: 10188662
24. Moher D, Pham B, Klassen TP, et al. What contributions do languages other than English make on the results of meta-analyses? *J Clin Epidemiol* 2000 Sep;53(9):964-72. PMID: 11004423
25. Juni P, Holenstein F, Sterne J, et al. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int J Epidemiol* 2002 Feb;31(1):115-23. PMID: 11914306
26. AHRQ EHCP methods guide for comparative effectiveness reviews of medical interventions. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2010 Nov. 207 p.
27. Neumayer L, Jonasson O, Fitzgibbons R, Henderson W, Gibbs J, Carrico CJ, Itani K, Kim L, Pappas T, Reda D, Dunlop D, McCarthy M, Hynes D, Giobbie-Hurder A, London MJ, Hatton-Ward S. Tension-free inguinal hernia repair: the design of a trial to compare open and laparoscopic surgical techniques. *J Am Coll Surg* 2003 May;196(5):743-52. PMID: 12742208
28. Laparoscopic versus open repair of groin hernia: a randomised comparison. The MRC Laparoscopic Groin Hernia Trial Group. *Lancet* 1999 Jul 17;354(9174):185-90. PMID: 10421299
29. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. *Pharmacoeconomics* 1999 Feb;15(2):141-55. PMID: 10351188
30. Hagg O, Fritzell P, Nordwall A, Swedish Lumbar Spine Study Group. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J* 2003 Feb;12(1):12-20. PMID: 12592542
31. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986 Sep;7(3):177-88. PMID: 3802833
32. Borenstein M, Hedges L, Higgins J, et al. *Comprehensive meta-analysis*. Version 2. Englewood (NJ): Biostat; 2005. Also available: http://www.meta-analysis.com/pages/about_us.html.
33. Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. In: *Methods Guide for Comparative Effectiveness Reviews*. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2010 Oct 25. Also available: <http://effectivehealthcare.ahrq.gov>.
34. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med* 2004 May 15;23(9):1351-75.
35. Bradburn MJ, Deeks JJ, Berlin JA, et al. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med* 2007 Jan 15;26(1):53-77. PMID: 16596572
36. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988. 567 p.
37. Rucker G, Schwarzer G, Carpenter JR, et al. Undue reliance on I^2 in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008 Nov 27;8(1):79. PMID: 19036172
38. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. *Stat Med* 2004 Jun 15;23(11):1663-82. PMID: 15160401
39. STATA statistics/data analysis. MP parallel edition. College Station (TX): StataCorp; 1984-2007. Single user Stata for Windows. Also available: <http://www.stata.com>.

Source: www.effectivehealthcare.ahrq.gov

Published Online: August 15, 2011



40. Owens D, Lohr K, Atkins D, et al. Grading the strength of a body of evidence when comparing medical interventions. In: Methods guide for comparative effectiveness reviews. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2009 Jul. p. 19. Also available: http://effectivehealthcare.ahrq.gov/ehc/products/122/328/2009_0805_grading.pdf.
41. Treadwell JR, Tregear SJ, Reston JT, et al. A system for rating the stability and strength of medical evidence. BMC Med Res Methodol 2006;6:52. Also available: <http://www.biomedcentral.com/1471-2288/6/52>. PMID: 17052350
42. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-58. PMID: 12111919

VI. Definition of Terms

I^2 This is a measure of heterogeneity, ranging from 0% to 100%, in which higher values suggesting greater heterogeneity. See Higgins and Thompson (2002)⁴² for more details.

Tau This is a measure of heterogeneity indicating the standard deviation of the effect sizes; it is on the scale of the effect size. For example, in a meta-analysis of log odds ratio, tau is on the scale of the log odds ratio. See Rucker et al. (2008)³⁷ for more details.

TAPP Transabdominal pre-peritoneal repair

TEP Totally extraperitoneal repair

VII. Summary of Protocol Amendments

August 9, 2011

Section	Protocol Deviation	Rationale
IV. E.	The protocol had stated that the minimum clinically significant difference (MCSD) for the outcome return to daily activities is one week. We propose to change it to one day.	The typical length of time before returning to daily activities after inguinal hernia surgery is about 10 days. This relatively short period of time suggests that shortening it by just 1 day would be considered important to patients and/or clinicians. Thus we propose setting the MCSD at one day for this outcome. Nothing in the literature clearly indicates what practicing clinicians consider the MCSD after hernia repair, so this is based on the consensus of the research team.
IV. E.	The protocol had stated that the minimum clinically significant difference (MCSD) for the outcome return to work is one week. We propose to change it to one day.	Same points as above, except for return to work.

VIII. Review of Key Questions

For all EPC reviews, key questions were reviewed and refined as needed by the EPC with input from Key Informants and the Technical Expert Panel (TEP) to assure that the questions are specific and explicit about what information is being reviewed. In addition, for Comparative Effectiveness reviews, the key questions were posted for public comment and finalized by the EPC after review of the comments.

Source: www.effectivehealthcare.ahrq.gov

Published Online: August 15, 2011

IX. Key Informants

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Experts

Technical Experts comprise a multi-disciplinary group of clinical, content, and methodologic experts who provide input in defining populations, interventions, comparisons, or outcomes as well as identifying particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicted opinions are common and perceived as health scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design and/or methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor contribute to the writing of the report and have not reviewed the report, except as given the opportunity to do so through the public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodologic expertise. Peer review comments on the preliminary draft of the report are considered by the EPC in preparation of the final draft of the report. Peer reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and will, for CERs and Technical briefs, be published three months after the publication of the Evidence report.

Source: www.effectivehealthcare.ahrq.gov

Published Online: August 15, 2011

Potential Reviewers must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than \$10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.