Comparative Effectiveness Review Disposition of Comments Report

Research Review Title:
Management of Uterine Fibroids

Draft review available for public comment from September 12, 2016 to October 11, 2016.

Research Review citation: Hartmann KE, Fonnesbeck C, Surawicz T, Krishnaswami S, Andrews JC, Wilson JE, Velez-Edwards D, Kugley S, Sathe NA. Management of Uterine Fibroids. Comparative Effectiveness Review No. 195. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290-2015-00003-I.) AHRQ Publication No. 17(18)-EHC028-EF. Rockville, MD: Agency for Healthcare Research and Quality; December 2017. www.effectivehealthcare.ahrq.gov/reports/final.cfm. doi: https://doi.org/10.23970/AHRQEPCCER195.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site or AHRQ Web site in draft form for public comment for a 3-4-week period. Comments can be submitted via the Web site, mail or E-mail. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the 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 and Affiliation	Section	Comment	Response
Peer Reviewer #2	Introduction	Good general intro to subject.	Thank you for your comment.
TEP Reviewer#5	Abstract	I believe that this sentence requires clarification: “Myomectomy reduced fibroid volume and improved quality of life, but did not improve bleeding.” If myomectomy removes the fibroid, there should be no volume or bleeding. Correct? Do you mean uterine volume? Or are the authors referring to lifetime recurrence? Please specify the time period during which these endpoints are being measured.	Myomectomy by definition reduces fibroid volume but myomectomy does not always remove all fibroids therefore fibroid volume is not uniformly zero after myomectomy. To avoid the need for this explanation we have revised to plain language: "Myomectomy removed fibroids and improved quality of life, but did not improve bleeding." For brevity we have not included the varied timing of assessment of outcomes in the abstract.
TEP Reviewer#5	Abstract	Please report a time period in this sentence: “Subsequent intervention ranged from zero to 40 percent in studies that followed women after initial fibroid treatment (give range of follow-up durations for all studies combined here).” The cumulative incidence is not meaningful without the specification of a time period.	We have noted that this is up to 24 months of follow-up. Appendix H provides results at up to 6, 12, and 24 months.
TEP Reviewer#5	Abstract	This sentence could be more complete: “Analysis of survival data suggested that use of morcellation and morcellation method were not strong predictors of overall survival among those diagnosed with sarcoma after hysterectomy for fibroids.	We have added "among those diagnosed with sarcoma after hysterectomy or myomectomy for fibroids."
Peer Reviewer #2	Executive Summary	Under the GnRH agonist section, ES-5 to ES-6 needs to summarize subsequent treatment as an outcome.	We note that we did include discussion of outcomes of interest in each section of the report if that outcome was not reported.
Peer Reviewer #2	Executive Summary	Table ES-D (Summary SOE table)--prefer the format where SOE in a third column rather than in inline text with the rest of the key findings. It is harder to see the SOE for any given outcome.	We have revised the tables in the Executive Summary to improve clarity and readability.
TEP Reviewer #2	Executive Summary	There are studies with other GnRH agonists that cover the same symptoms but using different terms e.g. amenorrhoea implies relief of HMB	We have replaced menorrhagia with "heavy menstrual bleeding" and added days of bleeding (amenorrhea = 0).
TEP Reviewer #2	Executive Summary	The term 'menorrhagia' should be avoided as it means different things in different countries. Heavy menstrual bleeding is better as it is clearer.	We have replaced the term with "heavy menstrual bleeding" in the document.
TEP Reviewer #2	Executive Summary	although these symptoms can be prevented with hormonal 'add back' therapy.	We revised this sentence revised as follows: "..., and bone loss , although some of these can be ameliorated with hormonal “add-back” therapy..
TEP Reviewer #2	Executive Summary	This is meaningless without the original uterine size. Better to use percentage change in volume.	Full details about fibroids and uterine size are provided in the full report. We prefer absolute change in size as a metric. We have added detail to clarify.
TEP Reviewer #2	Executive Summary	Bleeding worse with sub-mucous fibroids.	We have noted this in the Executive Summary.
TEP Reviewer #2	Executive Summary	Tamoxifen does not cause fibroid shrinkage.	Tamoxifen has been studied in trials so it is included. The trial, as indicated, did not report on change in fibroid characteristics.
TEP Reviewer #2	Executive Summary	There are also data from the REST trial.	REST trial data are included in this summary statement.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer #2	Executive Summary	Better to say 'provided information on adverse effects'	We use the term "harms" to imply the range of adverse events captured.
TEP Reviewer #2	Executive Summary	It should be stated that this study is not a formal RCT and the 2 groups differed significantly at baseline	We have added details that expand on the methods of the study including the fact that UAE participants were advised by a clinical protocol described to proceed to myomectomy under prescribed conditions. Technically the groups are statistically balanced at baseline though we concur this could be from being underpowered even for these comparisons.
TEP Reviewer #2	Executive Summary	Too strong for a low quality study.	We revised SOE here and later in the report as described below in the reply to comments in the more detailed portions of the full report.
TEP Reviewer #2	Executive Summary	This study is not conclusive.	Revision of this item states: "Reproductive outcomes were reported to be better after myomectomy compared with UAE among a subgroup of participants from a small study (n=66 who desired pregnancy), which was underpowered."
TEP Reviewer #2	Executive Summary	This is not true. REST and EMMY both did.	It is true of "most" studies as defined by the number of unique trials. REST and EMMY are exceptions and receive additional attention precisely because they provide better outcomes data.
TEP Reviewer #2	Executive Summary	It should be made clear that there are no data to show that pregnancy outcome is improved by myomectomy except the small study mentioned above. Studies are needed to address this point.	We have modified this sentence to note that myomectomy is an option for women desiring future fertility, though evidence is insufficient to define potential benefit.
TEP Reviewer #2	Executive Summary	but not as much as GnRH agonists.	Summaries in this section are for individual categories of intervention are not comparative, but rather provide SOE for each type of outcome assessed in the related literature.
TEP Reviewer #2	Executive Summary	possibly	We have revised this sentence to note "with the exception that pregnancy outcomes are possibly less favorable after UAE…"
TEP Reviewer#4	Executive Summary	abbreviated--short term or limited data/questions?	We revised this sentence to read: Most were judged to be of poor quality and typically reported only on technical success of the intervention. Longer term outcomes such as quality of life, improvement in symptoms and satisfaction with care were rarely described.
TEP Reviewer#4	Executive Summary	robotic approach--lap w/robotic assist?	Revised.
TEP Reviewer#4	Executive Summary	mean followup duration-range?	We have noted the range in the tables.
TEP Reviewer#4	Executive Summary	"(2 of 2)"--what???	This was a tracking note counting studies during the writing phase; it has been removed.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#4	Executive Summary	Were there other studies that reported no changes in QoL?	Several reviewer comments indicate that we were not sufficiently clear about how we have used the term "change" in the report. We have added this text in the Executive Summary and full report: "Please note, throughout this report in the text we refer to whether or not a study assessed change in specific measures include fibroid characteristics, like size or volume, bleeding, pain, and quality of life. Indicating the study assessed changed means they evaluated the characteristic or symptom at baseline and again at one or more times after treatment. Noting that a study or studies assessed change is not equivalent to noting a beneficial effect or statistical significance in changes, rather noting measurement of change in a parameter establishes the total count of studies that addressed this outcome."
TEP Reviewer#4	Executive Summary	But not others? Or they had high subsequent proc rates?	We have revised this text.
TEP Reviewer#4	Executive Summary	"All were rated as poor quality primarily due to lack of masking of participants and assessors."--to???	We have reworded this to indicate lack of masking to intervention.
TEP Reviewer#4	Executive Summary	Did others not study this or not see a change?	As above, change is a type of outcome measure. Throughout this report in the text we refer to whether or not a study assessed change in specific measures include fibroid characteristics, like size or volume, bleeding, pain, and quality of life. Indicating the study assessed changed means they evaluated the characteristic or symptom at baseline and again at one or more times after treatment. Noting that a study or studies assessed change is not equivalent to noting a beneficial effect or statistical significance in changes, rather noting measurement of change in a parameter establishes the total count of studies that addressed this outcome.
TEP Reviewer#4	Executive Summary	Did they look at the outcomes and saw no change or not look at the outcomes?	Throughout this report in the text we refer to whether or not a study assessed change in specific measures include fibroid characteristics, like size or volume, bleeding, pain, and quality of life. Indicating the study assessed changed means they evaluated the characteristic or symptom at baseline and again at one or more times after treatment. Noting that a study or studies assessed change is not equivalent to noting a beneficial effect or statistical significance in changes, rather noting measurement of change in a parameter establishes the total count of studies that addressed this outcome.
TEP Reviewer#4	Executive Summary	But how did the UAE women compare in characteristics vs. the hyst patients?	Available comparisons to hysterectomy outcomes are summarized. In these RCTs, randomization achieved groups who were comparable at baseline.
TEP Reviewer#4	Executive Summary	…within specific fibroid characteristics, symptoms and other clinical characteristics	We did not adapt this text. If evidence is insufficient overall it is implied that evidence of finer tailoring of treatment is fully insufficient.
TEP Reviewer#4	Executive Summary	"Overall, patient satisfaction…"--what time period?	This statement referred to satisfaction at the end of follow-up. We have edited the text to clarify.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#4	Executive Summary	"Too few studies…"--can you look across studies? Maybe based on criteria?	We aimed to report on the content of the literature for this KQ in order to determine if studies themselves (as opposed to aggregated in meta-regressions or other statistical approaches) are being designed with adequate power to answer these important clinical questions. The literature lacks well-designed studies to address characteristics that may be effect modifiers.
TEP Reviewer#4	Executive Summary	But can you compare among studies how women w/similar characteristics fared w/in each treatment?	Other analyses might be undertaken but were not planned in this review Further because of how data is reported and due to variation in inclusion and exclusion criteria such and analysis would be difficult to implement without primary data from the included trials.
TEP Reviewer#4	Executive Summary	Good or bad as defined by??	We have modified to lower risk or higher risk.
TEP Reviewer#4	Executive Summary	How are these known?	We have added phrase to indicate "by expert consensus."
TEP Reviewer#4	Executive Summary	ID of PCOs	Thank you for this excellent point. We have added a bullet to suggest that formal development of patient centered outcome measures for fibroid care is an important future research point.
TEP Reviewer#4	Executive Summary	What statistical methods used? CART? Latent class? Can they be improved?	We have noted that more robust statistical techniques across aggregated data are needed in order to understand the modifiers of that course and of the effectiveness of treatment, so that we can offer women an accurate account of the likely outcome of intervention choices based on their individual status.
TEP Reviewer#5	Executive Summary	Results: I believe the reader will be more interested in the magnitude of the association as opposed to statistical significance in the following sentence: "Five studies reported absence of bleeding, three noting statistical significance for clinically important reduction from baseline. One study reported reduction in days of bleeding (significance not reported), and four reported improvement in hemoglobin levels (significant in 3)."	Additional specifics about effects have been added throughout the report, including here, to help differentiate magnitude of effect from whether statistical testing suggests the effect size is significantly different given study power.
TEP Reviewer#5	Executive Summary	Results: Again, I believe the reader will want to know the direction and magnitude of the association as opposed to statistical significance. Please revise the following sentence: "Fibroid size decreased in two studies of raloxifene, was not statistically different at end of followup in one study, and was not reported in the single trial of tamoxifen."	We have revised to lay language for this executive summary statement.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#5	Executive Summary	Results-UAE: I noticed that the authors comment on fertility in the Myomectomy section but not here. Would it be worth mentioning what the outcomes are for fertility post-UAE or at least acknowledge that women who undergo UAE tend to be older and parous because UAE is not recommended to those who have not had children? Is it possible that fibroids have a different clinical course (e.g., better outcomes) in parous vs. nulliparous women?	This statement was added to the summary: "In clinical care UAE is not recommended for women who desired future pregnancy, when available pregnancy outcomes after UAE are presented in the sections that compare UAE to other procedures or surgeries."
TEP Reviewer#5	Executive Summary	Results: Please specify the time period (or range of time periods) for the following sentence: “Conversion from myomectomy to another procedure ranged from 0 to 17 percent in eight studies (n=658).”	We have edited to expand on typical definition of conversion which is "intraoperative conversion".
TEP Reviewer#5	Executive Summary	Results: Again, a time period (or range of time periods) needed here: “Rates of subsequent intervention ranged from zero to 40 percent for women in their 30s, 40s, and 50s.”	We have added the time frame.
TEP Reviewer#5	Executive Summary	Results-Expected Management: Some acknowledgement that the women opting for expectant management may have fewer symptoms than the average woman with fibroids. Also, if there are any data regarding their ages at diagnosis, symptom severity, and estimated time until menopause, that would be interesting to include. Some skeptics might argue that they are not a fair comparison group although I am glad they are included in this report. I do indeed see text on page 27 (lines 21-23) that alludes to this, but perhaps page 20 is a good place to first mention similar to what is on page 27: “No studies were appropriately powered to understand whether specific groups of patients, such as those closer to menopause or with a specific symptom pattern have outcomes that are modified by those characteristics.”	We have revised this section to address potential for differences from other women with fibroids and considerations suggested by other reviewers that covers similar concepts to these.
TEP Reviewer#5	Executive Summary	Results: “Evidence is moderate that myomectomy is associated with improved fibroid characteristics (volume/size) and quality of life.” Please clarify whether the authors mean uterine volume? By definition, wouldn’t the fibroid volume and size be zero when the fibroid is removed via myomectomy? This is unclear.	We revised this sentence to read: "Evidence is moderate that myomectomy removes fibroids and is associated with decreased uterine volume and improved quality of life.
TEP Reviewer#5	Executive Summary	Results: The word “statistically” is not needed in the following sentence: "Of note, the evidence is insufficient to determine if myomectomy statistically meaningfully improves bleeding patterns or anemia."	Revised.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#5	Executive Summary	Research recommendations. Consider mentioning that reproductive outcomes such as fertility and pregnancy complications are very important to women as they make decisions about fibroid treatment. Research in this area is very limited.	We have added a key point as follows: Likewise reproductive outcomes such as fertility and risk of future pregnancy complications are very important to women as they make decisions about fibroid treatment and the current literature is insufficient to guide choices.
TEP Reviewer#5	Executive Summary	I would be interested in knowing the range of effect sizes (i.e., magnitude of % change) in this sentence: "Statistically significant improvements from baseline were noted in these aspects across studies at one or more doses"	Throughout the report we have added more specific counts and outcome results. However here we have decided to retain this summary since a listing of differences in scores without the context of scoring methods, baseline scores and known minimally important clinical differences made the presentation cluttered but not more informative. (The instruments themselves are presented earlier for background.) The references specific to each quality of life domain presented in the literature will allow those with interest to find the details.

Commentator and Affiliation	Section	Comment	Response
Public Reviewer (Allergen)	Background	Uterine fibroids affect between 80 to 90 percent of African American women and 70 percent of Caucasian women by age 50. The condition is a leading cause of hysterectomies in the U.S., resulting in more than 250,000 such procedures annually. We commend AHRQ's Effective Health Care Program for recognizing the wide-ranging impact of uterine fibroids on women's health, including "personal and societal costs including diminished quality of life, disruption of usual activities and roles, lost work time associated with symptoms, and substantial healthcare expenditures." Currently, there are **no Food and Drug Administration (FDA)-approved medications** to specifically manage long term uterine fibroids. The only approved medication by FDA for uterine fibroids is leuprolide, which concomitantly with iron therapy is indicated for the preoperative hematologic improvement of patients with anemia caused by uterine fibroids. Due to safety issues, the use of leuprolide is limited for only 3 months. Ulipristal acetate is currently the only medication approved by the European Medicines Agency for intermittent treatment of moderate to severe symptoms of uterine fibroids in adult women of reproductive age, and for preoperative use. The latter indication also is approved in Canada but currently is limited to two, 3 month treatment cycles. To date, more than 300,000 women have been treated with Ulipristal acetate for fibroids in over 50 countries. Allergan is conducting a Phase III development program specifically in the United States with two Phase III studies. An application for FDA approval is expected to be filed in 2017. The first positive Phase III trial results have been already released and the second Phase III trial will conclude in early 2017.	Thank you for your comment. As background for the methods of the review, we included RCTs of any intervention for uterine fibroids available in the U.S., without respect to approved label indications, if the reported outcomes included patient-centered final health outcomes. Studies with only intermediate outcomes (detailed in Figure 1) were not included for comparative effectiveness, except to evaluate harms. The review "does not cover preoperative or adjunctive treatments." This accounts for exclusion of some papers included in other reviews. See below for correction of identified errors with regard to European Phase III trials. We look forward to the publication of the US Phase III trials but at present do not find them in the published literature.
Public Reviewer (Allergen)	Background	The first US Phase III trial indicates that medical therapy with Ulipristal acetate results in statistically significant reductions in uterine bleeding and statistically significant improvements in women's symptom and health-related quality-of-life indicators. As discussed in more detail below, **we urge AHRQ to incorporate the existing publications from randomized controlled studies of short term and long term intermittent use of Ulipristal** in the medical management of uterine fibroids and emerging research to assure that the final report fully reflects the evidence on non-invasive treatments options, which may offer the first medical management option for women with uterine fibroids in the US.	Thank you for the guidance. We are not fully certain without references which publications represent the full complement of "existing publications" referenced but believe the action described below with regard to Donnez et al. 2012 (PMID: 22296075 and PMID: 22296076) addresses and resolves the inclusion of the relevant European Phase III trials.

Commentator and Affiliation	Section	Comment	Response
Public Reviewer (Allergen)	Background	AHRQ Incorrectly Excluded Two Pivotal Trials Which Demonstrates the Efficacy and Safety of Ulipristal in the Treatment of Uterine Fibroids. Donnez J, Tatarchuk TF, Bouchard P, et al.; Ulipristal acetate versus placebo for fibroid treatment before surgery. N Engl J Med 2012 Feb 2;366(5):409-20. Donnez J, Tomaszewski J, Vazquez F, et al.; Ulipristal acetate versus leuprolide acetate for uterine fibroids. N Engl J Med 2012 Feb 2;366(5):421-32.	Thank you for identifying this error. These trials were originally screened out as medication use adjunctive to surgery, this was a mis-interpretation of phrases like "before surgery" and "before planned surgery". Since Table 1 in both Donnez publications noted (PMID: 22296075 and PMID: 22296076) as well as supplemental materials make it clear that only a portion of women proceded to surgery and the effects of ulipristal on the surgical procedures were not the focus, we have re-reviewed all publications retrieved in the search for RCTs of ulipristal and CB-2914. The publications noted **do contribute** useful data and **are now included**. We also identified one additional trial by Barlow and colleagues (PEARL 1; PMID: 24457604) and two additional follow-up papers from trials (Luyckx, 2014, PMID: 25241376 and Williams, 2012, PMID: 23018219)
Public Reviewer (Allergen)	Background	In their Management of Uterine Fibroids draft report, AHRQ incorrectly excludes two pivotal EU registration studies which were published in the New England Journal of Medicine in 2012. According to the AHRQ draft report; these studies were excluded for the following reasons: X-1(e) Pre-operative adjuncts to shrink fibroids or improve anemia. X- 6 Does not report an outcome of interest X- 7 Does not address a key question	As above, thank you for identifying this error. These studies were misclassified. These publications do contribute useful data and **are now included**.
Public Reviewer (Allergen)	Background	Both studies were randomized controlled trials (RCTs) that provided key efficacy and safety information for Ulipristal treatment and served as the basis for the approval of Ulipristal for the treatment for uterine fibroids in Europe and Canada. Both of these pivotal RCTs evaluated the benefits and harms of Ulipristal compared with placebo (ref 268) or active control – leuprolide (ref 269) and demonstrated improvement in bleeding and symptoms such as pain and improvement in quality of life, which are relevant patient centered outcomes.	As above, thank you for identifying this error. These studies were misclassified. These publications do contribute useful data and **are now included**.
Public Reviewer (Allergen)	Background	Only subjects who were eligible for surgery were enrolled in these studies, however the treatment outcomes observed below were both patient centered and not only related to the surgery: Reduction of uterine bleeding, defined as a PBAC score < 75, Rate of amenorrhea, Change in pain (VAS/ Visual Analog scale and SF McGill Pain Questionnaire), Change in QoL (quality of life), Change in fibroid and uterus volume, Change in Hemoglobin, Hematocrit and ferritin levels	We agree these are eligible and important outcomes. These studies were misclassified. These publications do contribute useful data and **are now included**.

Commentator and Affiliation	Section	Comment	Response
Public Reviewer (Allergen)	Background	Per information included in Appendix D of the draft report appendix document, AHRQ appears to have excluded these two European registration studies due to preoperative use. These studies met all RCT criteria and included only Ulipristal treatment related endpoints such as bleeding control and fibroid volume reduction after Ulipristal treatment. There were no post-operative endpoints in those studies. These studies were inappropriately excluded and do not meet any of the exclusion criterion cited by AHRQ in Appendix D.	We agree and they are now included.
Public Reviewer (Allergen)	Background	Recommendation: AHRQ Should incorporate two publications of well controlled pivotal EU Ulipristal Studies (Donnez et. al. 2012) to Randomized Control Trial Data in the report. Allergan believes that data from these publications will be used to support efficacy and safety assessment of Ulipristal and we strongly recommend that AHRQ should include data from both Donnez clinical trials prior to releasing their final report.	We agree these trials contribute and the data is **now included in review of the medication category, comparative effectiveness, and subsequent interventions**.
Peer reviewer #1	Introduction	Problem clearly introduced and justified	Thank you for your comment.
TEP Reviewer #2	Introduction	Our study an RCT pre TAH and so might have been usefully included.	As noted, we did include studies of preoperative medications in which all women were intending to have subsequent surgery.
TEP Reviewer#5	Introduction	I think the introduction is beautifully-written and I like how it is phrased in terms of numbers of woman affected. I would suggest adding one more sentence about how African American women are disproportionately affected by fibroids relative to other racial/ethnic populations in terms of age at first diagnosis, severity of symptoms, and health care costs. If there is still room, the age range of women typically affected by fibroids also seems appropriate for this section.	We have added additional detail added as suggested.
Peer Reviewer #2	Methods	Rather than the informal "cohorts" I think you should consistently be using "cohort studies" for clarity. The word "cohort" alone just means a group and that's not what you really need to convey. This repeats in multiple places in the report.	We have replaced cohorts with "cohort study" in the majority of locations. However for brevity in instances in which cohort study is fully implied we have retained the more telegraphic form.
Peer Reviewer #2	Methods	Seems like a big effect modifier would be size of fibroid and it seems like most studies report fibroid volume. Why is fibroid volume not included as part of metaregression to help answer the question of how fibroid characteristics relate to treatment effects for KQ 2?	KQ2 was not a planned target for meta-analysis because few studies reported key individual characteristics or common outcome metrics for meaningful adjustment.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer #2	Methods	Is this the same as GRADE?	AHRQ EPC's strength of evidence (SOE) approach is similar to GRADE but assesses some different domains. EPC's assess the domains of study limitations (which incorporates risk of bias), consistency, directness, precision, and reporting bias and may assess additional optional domains. For more information about EPC SOE methods, please see Berkman et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. *Journal of Clinical Epidemiology* 2015. http://dx.doi.org/10.1016/j.jclinepi.2014.11.023
Public Reviewer (Allergen)	Methods	Quality of Life (pg. 11) – AHRQ's Current Symptom Severity Scale Does Not Measure the Variability and Severity of Menses.	We have aligned the description to more closely match Spies et al. (PMID: 11814511) and Coyne et al. (PMID: 22867776). **Revision reads**: The Symptom Severity Scale assesses severity of fibroid related symptoms (including items that reflect bleeding characteristics, pressure, urinary frequency, and fatigue). The scale is reported as a 0 to 100 score, with a higher score representing greater severity of symptoms.
Public Reviewer (Allergen)	Methods	In its draft report, AHRQ discusses that the System Severity Scale of the UFS-QoL measures the variability and severity of menses through a zero to one hundred scale. In fact, the intent of the Symptom Severity Scale is not to measure the "variability and severity of menses" as stated in the report. It measures the "distress due to the frequency and severity of uterine leimyomata symptoms". Furthermore, the response options are not measured on a 0 to 100 scale. Instead, item response choices are presented as five-level Likert scales ranging from 'not at all' to 'a very great deal.' The response options are transformed to a summary scale that ranges from 0 to 100, where higher symptom scores signify greater distress due to symptom severity. Finally, the HRQoL is a subscale of the UFS-QOL made up of six (not seven) domains (not subscales): 1) concern; 2) activities; 3) energy/mood; 4) control; 5) self-consciousness; and 6) sexual function.	We have revised our description of this scale while retaining emphasis on interpretation of scores as above (not how items are administered and how scores are transformed). The parent paper for the UFS-QOL Questionnaire describes "subscales" (Table 2) and we have retained that language. To better acknowledge that the SSS and HRQoL stand alone within the UFS-QoL, we **revised the text** about the HRQoL to say: "The HRQoL scale includes subscales to assess: 1) concern; 2) activities; 3) energy/mood; 4) control; 5) self-consciousness; and 6) sexual function. The HRQoL Scale is also reported as scores from 0 to 100, with higher scores reflecting better quality of life."
Public Reviewer (Allergen)	Methods	Recommendation: AHRQ should amend its draft report and restate the intent of the Symptom Severity Scale contained within the UFS-QoL. They should also amend the reports description of the domains and summary score of the UFS-QoL accordingly.	We have revised our descriptions as above.
TEP Reviewer#5	Methods	Yes, the inclusion and exclusions criteria are clearly defined and justifiable. The search strategies are explicitly stated and logical. The definitions and diagnostic criteria for the outcomes are clear.	Thank you for your comments.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#5	Methods	There was no abbreviation for LMS (leiomyosarcoma?), but perhaps it can be removed and "sarcoma" can be used in its place throughout the document. The use of "sarcoma" and "leiomyosarcoma" interchangeably throughout is a bit confusing -- please consider using just one of these terms and being consistent.	We have used term leiomyosarcoma (instead of sarcoma) throughout the report. We have avoided using LMS abbreviation unless necessary for tables or figures.
TEP Reviewer#5	Methods	Yes, the statistical methods used are appropriate. At times, the authors rely too heavily on statistical significance to summarize the results of key studies. See attached document for additional specific comments.	In instances in which we comment on statistical significance we generally have also noted clinical significance. Most often we make this note to indicate inadequate study power. We have revised in several places to state the latter. Presenting both results and statistical significance is our aim as demonstrated in tables in which we make visible the direction of effect, the consistency of findings across studies, and the reported statistical significance of outcomes.
TEP Reviewer#5	Methods	Spell out LMS on first mention. Be consistent in the use of this term. I see “sarcoma” and “leiomyosarcoma” used in the same paragraph as well.	We have used term leiomyosarcoma (instead of sarcoma) throughout the report. We have avoided using LMS abbreviation unless necessary for tables or figures.
TEP Reviewer#5	Results	Outcomes: Why not include “uterine volume/size” as an endpoint of interest? I see text mentioning this outcome on page 49 of 139, line 15, for example.	We have noted that characteristics can include fibroid volume, fibroid size, uterine volume and size.
Peer reviewer #1	Results	1. Result details adequate 2. Study characteristics are poorly defined-large number of studies and inter-study variability may be the limiting factor	We have added additional detail about study populations. The reviewer is correct that while many study populations share some characteristics (e.g. symptomatic fibroids, premenopausal status), inclusion and exclusion criteria do not result in demonstrable comparability of populations on other potentially important characteristics such as total fibroid number, volume, parity, reproductive intent, etc.
TEP reviewer #1	Results	Yes, the results are clearly stated and describe relevant and concrete outcomes.	Thank you for your comment.
TEP reviewer #1	Results	On page 87 of 139, first paragraph, "Perforation of Organs" - was damage to GU organs such as bladder and ureter also evaluated?	Yes, we sought and extracted any data related to injury to organs and perforations including but not limited to bowel, bladder, and ureter.
TEP reviewer #1	Results	For KQ3, reference #149 was not limited to cases including morcellation of presumed myomas. The denominator includes all LAVH and TVH for presumed benign gynecologic conditions, regardless of indication, so includes many non-fibroid cases (~47% of cohort).	To assure we included data only from women with a pre-operative diagnosis of fibroids we used the denominator data from Table 2 (n= 435 with fibroids) and information from text and from Table 1 in the pathology column that indicates that none of the incidental cases had leiomyosarcoma, so the numerator was zero cases.

Commentator and Affiliation	Section	Comment	Response
TEP reviewer #1	Results	For KQ4 additional studies that I do not see included but which may impact the conclusions are: Raine-Bennett T, Tucker LY, Zaritsky E, Littell RD, Palen T, Neugebauer R, Axtell A, Schultze PM, Kronbach DW, Embry-Schubert J, Sundang A, Bischoff K, Compton-Phillips AL, Lentz SE. Occult Uterine Sarcoma and Leiomyosarcoma: Incidence of and Survival Associated With Morcellation. Obstet Gynecol. 2016 Jan;127(1):29-39. Cusidó M, Fargas F, Baulies S, Plana A, Rodríguez I, Tresserra F, Pascual MA, Fábregas. Impact of Surgery on the Evolution of Uterine Sarcomas. J Minim Invasive Gynecol. 2015 Sep-Oct;22(6):1068-74.	These studies have been included in our updated analysis.
Peer Reviewer #2	Results	Pg 32, lines 17-18: Sentence is unclear. Do you mean that "The proportion of subjects who did not have at least one post-treatment biopsy were low in seven studies, ranging from XX% to 27%." or do you mean "The proportion of subjects who did not have at least one post-treatment biopsy were very low in five studies (ranging from XX to YY%) and low in two others (21 and 27%, respectively.) OR do you mean something else?	We have revised this text to better specify the proportion of participants who had endometrial biopsies.
Peer Reviewer #2	Results	Table 12, pg. 32 would benefit from # of subjects in each group in addition to the # studies.	The total N in bottom row of table is total participants, entries in table cells are number of participants with the related biopsy findings. As noted, total studies with data at the related dose/time interval is indicated in the column headers.
Peer Reviewer #2	Results	Mifepristone Summary, pg 33, lines 13-27: there needs to be acknowledgement that mifepristone is really not freely available in the U.S. It is sold only to physicians who have a prescriber's agreement with the manufacturer (Danco) and not to pharmacies. In addition, it is only sold as pills of 200mg in this manner. Use in the U.S. for this indication would likely require a substantial amount of help from a compounding pharmacy along with a physician who had a dispensing arrangement with that pharmacy. Not straightforward at all and this should at least be acknowledged that using it would have some hoops that the other medical treatments reviewed do not have.	We agree that this is important to note. We have added the following sentence to the Pharmaceutical Management Overview and Nomenclature section "In the United States only physicians with a prescriber's agreement with the manufacturer (Danco) can obtain the drug and prescribing for patients with fibroids can require collaboration with a compounding pharmacy. This means mifepristone is not readily available for generalist use."
Peer Reviewer #2	Results	Page 34, lines 22-30, Ulipristal Summary. Word choice for "Moderate evidence" should be changed to something like "There is moderate SOE that ulipristal reduces the size of fibroids."	We have revised this statement to read "There is moderate strength of evidence that ulipristal reduces the size of fibroids."
Peer Reviewer #2	Results	Page 34, line 35, should include the # of subjects in that one small study for clarity.	We have added the number of subjects (n=60).
Peer Reviewer #2	Results	Page 41, Table 16: the second column "Group Participants, N" never actually has the # of subjects in that column.	We have added the number of subjects to the table.

Commentator and Affiliation	Section	Comment	Response
Peer Reviewer #2	Results	Page 45, Table 21: the next to last column "No Rx" should really be "No Treatment" (No Tx) as there are other treatments besides medications and Rx really is an abbreviation that refers to prescriptions or drugs.	We have revised the table.
Peer Reviewer #2	Results	Page 47, line 42: "preserve future fertility" needs a caveat about FSH/AMH results and that not everyone has preserved fertility.	We have revised the sentence to note that occlusion may allow some women to preserve future fertility but expected reproductive outcomes are not clear.
Peer Reviewer #2	Results	Page 49, line 24: : "(71 of 72) fibroids . . . " would read better as "71 of 72 fibroids) . . . "	We have revised the text "the authors reported that radiofrequency volumetric ablation successfully excised 98.6 percent (71 of 72) fibroids) in 25 patients"
Peer Reviewer #2	Results	Page 50, lines 6-15, and Page 51, lines 4-50: seems like this section on Endometrial Ablation really belongs in Procedures rather than Surgical Interventions.	We recognize that different groupings of interventions are possible. Endometrial ablation in retained in surgery since it is our impression that the majority of procedures in these studies took place in the operating room.
Peer Reviewer #2	Results	Page 53, line 35: unclear whether the " . . . 47.5 percent (19/40)." refers to 40 in the control group total or 40 trying to conceive. Please clarify.	We have clarified the statement to "A subset of women from a RCT comparing myomectomy to UAE among women with reproductive plans noted that women who attempted to conceive following myomectomy, 31 of 40 were pregnant at 13 months after fibroid removal and the delivery rate was 47.5 percent (19/40).{, #3052}
Peer Reviewer #2	Results	Page 55, Table 27, line 18: The note "a" doesn't appear anywhere in the table except the title. Wouldn't it just be more clear to title the table "Estimated Hysterectomy Intraoperative Blood Loss" or something similar?	We have revised presentation of tables throughout the document.
TEP Reviewer #2	Results	The data for even hysteroscopic myomectomy are not good quality.	We now included a sensitivity analysis which excludes the studies that included hysteroscopic resection. We understand the concern that retrieval of the whole mass and all specimen may be less robust for this surgical approach. Counter to our clinical instincts, excluding these studies with hysteroscopies in their counts, reduces rather than increasing the prevalence estimates. See Table X for details.
TEP Reviewer #2	Results	This should not be considered conclusive.	Is not offered as conclusive, which is why we noted a single small study. We have rephrased this sentence to state: Pregnancy outcomes were reported to be better after myomectomy than UAE among a subset of participants from a small study.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer #2	Results	You need to be less categorical as none of the studies are powered to show effects on fertility. The FSH and AMH studies are poor and you are drawing strong conclusions from them which I am concerned may not be justified.	We have added additional detail to better illustrate the considerations: "Ovarian failure, measured by follicle stimulating hormone (FSH) >40 IU/L and anti-Mullerian hormone (AMH), was reported in two trials (EMMY and REST). In EMMY trial 88 women were assigned to UAE. Their average age at baseline was 45. In this group FSH increased significantly compared with baseline (+12.1, p=0.001) by24 months after treatment with UAE. FSH >40 IU/L was reported in 12 percent and 18 percent at 12 and 24 months, respectively.{, #3175} Levels of AMH, were significantly lower indicating ovarian aging (p<0.05) at each followup up to 24 months after UAE. These changes in FSH and AMH were comparable to those randomized to hysterectomy (p=0.37), and the only predictor of becoming menopausal in each group was being older than 45 at randomization. A similar proportion of 73 women (11%) were observed to have menopausal levels of FSH at 12 months after UAE in the REST study. This was also comparable to levels in the surgical arm of their trial (p=0.47). Participants in REST also had an average age in their mid-forties at the time of randomization. The trial by Mara and colleagues included 58 women randomized to UAE. The average age of participants in their study was more than a decade younger than the other trials. In this younger study population the risk of elevated FSH >10 IU/L after intervention was higher among those with UAE (13.8%) than myomectomy (3.2%; p<0.05), though no participants became frankly menopausal. " We would also note that power was likely adequate for single SD change in these continuous measures.
TEP Reviewer #2	Results	This demonstrated that FSH also went up in the hysterectomy group as you would expect with age.	We have added more information about FSH changes in both the UAE only and UAE comparison sections for emphasis.
TEP Reviewer #2	Results	insufficiently powered.	This is explicitly noted in the text.
TEP Reviewer #2	Results	Numbers wanting to become pregnant usually unknown.	We agree this is poorly tracked in this literature, can be determined, and deserves attention.
TEP Reviewer #2	Results	Not helpful as REST and EMMY do as does FUME.	It is correct that "most" do not. Added: "With notable exceptions,..." along with references.
TEP Reviewer #2	Results	tibolone is not a GNRH agonist it is a type of HRT and is a compound with properties of oestrogen, progestogen and androgen.	These two studies have been moved to the section of the report discussing estrogens and estrogen receptor agents. We note that we did not address effects of tibolone because it is not available in the US; rather we focused on data in the comparator arm.
TEP Reviewer #2	Results	You would not expect [tibolone] to affect size.	We note that we did not address effects of tibolone because it is not available in the US; rather we focused on data in the comparator arm.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer #2	Results	I am surprised at this unless used as add back, I would have expected it to increase size certainly in some women.	HRT does not substantially spur fibroid growth. The prior SER which included review of large HRT cohorts with fibroid imagining summarized: We found five studies that provide moderate evidence (Level III) on the effects of menopausal hormone therapy on uterine fibroids. One study reports higher risks of first diagnosis of fibroids in peri- and postmenopausal women with a body mass index (BMI) less than 24 and 5 years or more of estrogen and progestogen therapy. Three of four studies reported no effect on fibroid size; one reported a higher rate of uterine growth with the percutaneous-oral schedule of hormone replacement therapy than with a single oral combination of oestradiol valerate and cyproterone acetate.
TEP Reviewer #2	Results	I think this needs rethinking as the groups in the Mar group are not well matched at baseline.	We revised the reproductive outcomes summary to note: "Because of low power to detect differences in pregnancy outcomes the evidence is insufficient to determine outcomes are proven to be better for pregnancies conceived after myomectomy compared to UAE. Likewise the limited data do not show that UAE is a safe option for women who wish to conceive. "
TEP Reviewer #2	Results	You need to think about age of women here. Most were in their 40s and so AMH meaningless in this group many of whom were perimenopausal.	We have taken age into account directly and included in the text.
TEP Reviewer #2	Results	This study is much higher quality than Mara study, also rated as Fair!.	Thank you for the query about risk of bias assessment for this study. Our assessment of risk of bias encompasses multiple factors including selection, performance, reporting, attrition, detection and other bias. We mistakenly considered attrition bias as 'medium' when in effect penalized the REST study for having an extended follow-up period. We agree overall the trial should have been scored as 'low' risk of bias as by the end of 5 years as there was only <10% lost to follow up. We have now updated the overall rating from 'fair' to 'good' quality study.
TEP Reviewer #2	Results	"Overall, fewer than half of women had another intervention within 24 months." This is worth emphasising I agree.	We agree. This information recurs in summaries.
TEP Reviewer #2	Results	I realise I am biased but I am unsure why EMMY is assessed as Good Quality and REST as fair. Happily the NEJM don't agree with you!	Thank you for the query about risk of bias assessment for this study. Our assessment of risk of bias encompasses multiple factors including selection, performance, reporting, attrition, detection and other bias. We mistakenly considered attrition bias as 'medium' when in effect penalized this study for having an extended follow-up period. We agree overall the trial should have been scored as 'low' risk of bias as by the end of 5 years as there was only <10% lost to follow up. We have now updated the overall rating from 'fair' to 'good' quality study.

Commentator and Affiliation	Section	Comment	Response
Public Reviewer (Allergen)	Results	Evidence Map for KQ 1 (pg. ES-4) - Ulipristal's pharmacologic class should be categorized as "Selective Progesterone Receptor Modulator" and not "Progestins."	We have improved the consistency of use of the full section header: "Progesterone receptor agents: anti-progestins, selective receptor modulators, and intra-uterine progesterone treatments," and the shortened section header "Progesterone Receptor Agents" with tables notes as needed, including in this location. The organizing principle of the report benefits from parsimonious grouping of the upper levels of categories of intervention. We intend to convey this grouping is interventions active in the progesterone pathways.
Public Reviewer (Allergen)	Results	Pharmacology and Clinical Safety: In the draft report, AHRQ categorizes Ulipristal as a Progestin. According to the European Medicines Agency labeling, Ulipristal acetate has a steroid structure and acts as a selective progesterone receptor modulator with predominantly inhibitory effects on the progesterone receptor.	We have changed our category name as above. Ullipristal is described as follows: Ulipristal acetate (Ella®, Esmya®) is a selective progesterone receptor modulator which binds the human progesterone but not the estrogen receptor. Ulipristal is structurally similar to mifepristone, but has less antiglucocorticoid activity, suggesting it is better alternative to mifepristone for long term use. It has been FDA approved since 2010 for emergency contraception. The European Medicines Agency granted marketing authorization for ulipristal acetate, 5 mg (Esmya, Preglem/Gedeon Richter) for long term medical management and preoperative therapy in reproductive age women with uterine fibroids.
Public Reviewer (Allergen)	Results	Recommendation: AHRQ should reclassify Ulipristal as a selective progesterone receptor modulator.	We have edited to consistently use our category name as above.
Public Reviewer (Allergen)	Results	Pharmaceutical Management- Ulipristal Acetate (pg. ES-6) The Information on Ulipristal's Hepatotoxicity is Incorrect and Is Only Based on Two Small Phase II Studies.	We summarize harms as reported in the studies meeting our criteria. This now includes the two Donnez Phase III trials discussed above.
Public Reviewer (Allergen)	Results	In their draft report, AHRQ used two relatively small Phase II studies to conclude that Ulipristal elevated liver enzymes in clinical trials. Utilizing such limited patient populations provides a substantial survey bias and does not adequately portray clinical results across broad populations. These two Phase II (Nieman and Levens) studies are studies of the symptomatic treatment for uterine fibroids. The Donnez et al studies 2012 showed no increase in liver function enzymes therefore AHRQ should include these studies in their final report.	As above, the additional publications are now taken into account.
Public Reviewer (Allergen)	Results	In addition, it is not uncommon to detect transient mildly increased in liver enzymes in sporadic cases in different studies. The information should be more focus on hepatotoxicity (when AST /ALT are concomitant with bilirubin elevation) and not only on transient AST/ALT increases. Ulipristal showed no hepatotoxic signals with repeated treatment cycles.	We reported potential harms as described in the included studies (now including Donnez). Harms are not a critical focus of the report and we do not plan to expand discussion of evaluation of this or other potential harms.

Commentator and Affiliation	Section	Comment	Response
Public Reviewer (Allergen)	Results	Recommendation: AHRQ should focus conclusions pertaining to safety, including liver enzymes, on large, well established Phase III RCTs in making their recommendations and not on small phase 2 studies with Ulipristal.	We reported potential harms as described in the included studies (now including Donnez). The revised text states: About 2 to 10 percent of women taking ulipristal experienced hot flashes. Among 978 biopsies at completion of treatment, six cases of confirmed hyperplasia (one with atypia) were reported. In two studies with six month followup after treatment, no women who had taken ulipristal had hyperplasia. The two smallest studies reported modest elevations of liver function enzymes during treatment; another larger trial documented change in liver function enzymes was comparable to those taking placebo.
Public Reviewer (Allergen)	Results	Progesterone Antagonist, Selective Receptor Modulators and Intra-uterine Progesterone Treatments (pg. 28) - Two of the Four Studies Utilized to Determine Medication Effectiveness for Ulipristal Were Phase II Trials; AHRQ does not correctly characterize Ulipristal's impact on Bleeding and Fibroid Size Reduction (pg.16)	We are interpreting these references to text with publication citations to mean we did not include the newer Phase III Trials which are **now included** as noted above. The revised summary text now states: Seven RCTs investigated treatment with ulipristal, a selective progesterone receptor modulator. All seven studies found ulipristal effective for reducing the size of individual fibroids and the overall fibroid burden as measured by total fibroid or uterine volume. Ulipristal, as intended, resulted in absent menses for the majority of women during treatment (range 62 to 100%), and the majority of studies reported improved bleeding and improved or stable hematocrit or hemoglobin. All ulipristal doses compared with placebo resulted in improved fibroid-related quality of life; and two trials also documented improvement in pain. [Please note as we audit extracted data for newly added RCTs, corrections to exact numbers in tables and text may occur.]
Public Reviewer (Allergen)	Results	In their draft report, AHRQ reviewed twelve studies designed to test the effectiveness of Ulipristal and mifepristone in the management of uterine fibroids. Two of the four Ulipristal trials consisted of small Phase II studies. Large, well-controlled studies should be used in determining efficacy and not small, biased samples. Allergan does not understand why AHRQ excluded two, pivotal clinical trials conducted by Donnez et. al which were published in the New England Journal of Medicine. Allergan strongly recommends that AHRQ include data from the published Donnez clinical trials prior to releasing their report.	As above, the additional publications are now taken into account.

Commentator and Affiliation	Section	Comment	Response
Public Reviewer (Allergen)	Results	One such randomized, double-blind, placebo- controlled trial conducted by Donnez, et. al. and reported in the New England Journal of Medicine demonstrated that oral Ulipristal acetate at a dose of 5 mg per day or 10 mg per day was effective in controlling excessive bleeding and shrinking fibroids in patients who had severe bleeding and associated anemia at baseline. Treatment with Ulipristal acetate, as compared with placebo, also resulted in clinically significant increases in hemoglobin and hematocrit levels and reductions in self-reported pain and discomfort due to fibroids.	As above, the additional publications are now included in the tables and text for all relevant outcome categories for which they contributed data including change in fibroid characteristics, anemia, and symptoms.
Public Reviewer (Allergen)	Results	Recommendation: AHRQ should not focus solely on two small Phase II studies by Levens and Nieman in their determination of efficacy and safety. Determination of efficacy and safety should focus on large, well-controlled studies such as the Donnez, et. al. studies, published in the New England Journal of Medicine in 2012 in addition to two additional long term studies from Donnez et al and published in Fertility and Sterility in 2014 and 2015. When determining the clinical effectiveness of Ulipristal in the medical management of uterine fibroids Allergan strongly recommends that AHRQ include data from the Donnez Phase III clinical studies prior to releasing their final report.	The Donnez publications are included. We however will include all literature that meets inclusion criteria with appropriate caveats about study size, power, design, and quality, meaning the two smaller trials will remain in the report.
Public Reviewer (Allergen)	Results	Progesterone Antagonist, Selective Receptor Modulators and Intra-uterine Progesterone Treatments, Ulipristal Acetate, Risk of Harms From Ulipristal (pg. 16) – AHRQ does not Correctly Characterize Treatment Related Effects of Ulipristal	We seem to have a page mis-alignment. The draft report included the largest sections on the outcomes of treatment with Ulipristal beginning on page 33 (page number at foot of the page). We assume the concern about incorrect reporting relates to the omission of the two Donnez publications. This has been corrected. The comparison to GnRH agonist is also now included in the related sections about comparisons of medications with other classes of medications.
Public Reviewer (Allergen)	Results	The data on biopsies provided on page 16 of the draft report do not match the data provided on ES-6 in the number of biopsies obtained and in the number of hyperplasia cases reported. All biopsies collected during 4 phase 3 studies (Donnez 2012, 2014, and 2015) were assessed by 3 independent pathologists and therefore all 4 publications should be included in describing biopsy findings.	We have corrected the data in the executive summary.
Public Reviewer (Allergen)	Results	Recommendation: When determining the treatment related effects of Ulipristal, Allergan strongly urges AHRQ to incorporate key safety information from two Phase III controlled studies (Donnez 2012) in addition to long term studies (Donnez, 2014 and 2015).	Donnez 2014 and 2015 were previously included and both 2012 publications are now incorporated as is Donnez 2016 which was identified in our update search.
Public Reviewer (Allergen)	Results	Appendix G, Registered Study Protocols (pg. G-5)- AHRQ Does not List the 5mg dose for Ulipristal in the Appendix	We have revised our description of this protocol.

Commentator and Affiliation	Section	Comment	Response
Public Reviewer (Allergen)	Results	Recommendation: AHRQ should appropriately list both the 5mg and 10mg dosage strengths for Ulipristal. The 5mg dose is the approved dose for Ulipristal in Canada and Europe.	We have amended the dose listing.
Public Reviewer (Allergen)	Results	There is a high level of evidence to support efficacy and safety of Ulipristal in the treatment of symptomatic uterine fibroids.	As we complete review of new studies (Donnez both assigned low risk of bias, good quality) and compilation of data, we have not made a final determination of whether the total number of participants, study quality, and length of followup will quality for strong evidence of effectiveness for all time intervals. Longer timeframes and some outcomes may remain moderate strength.
Public Reviewer (Allergen)	Results	Recommendation: AHRQ should adopt currently available published data from the literature	With correction of the exclusion of two papers and re-evaluation of the search, we now include a total of 11 publications reporting data from a ulipristal arm in an RCT. This is an increase of three distinctive RCTs and two followup publications that contribute to the synthesis of extant literature. We believe our process for audit while reports are out for review, comments from peer reviewers such as your group, and the update with audit have generated a comprehensive view of the eligible extant literature. Thank you for contributions to improving the report.
TEP Reviewer#4	Results	Any details on whether they took the cervix? Tubes? Ovaries?	These data were not reported.
TEP Reviewer#4	Results	So is race an EM of Tx effectiveness?	Data were inadequate to examine this.
TEP Reviewer#4	Results	Fertility? Some groups say its [UAE] contraindicated for fertility.	We revised this sentence to read: "It is not currently clinically recommended for women who wish to have future pregnancies."
TEP Reviewer#4	Results	So do more complications [of UAE] emerge over time?	The data illustrate complications that occurred at any time point. Studies typically did not report emergence over time.
TEP Reviewer#4	Results	How do you know if power calc. were not included?	We revised this sentence to read: "Thus post hoc estimation suggests power was to detect…was limited."
TEP Reviewer#4	Results	What is the role of the menopausal transition?	This is not directly addressed by our analysis other than grouping of outcomes by decade of life. Potential implications of age outside the findings in our results section are discussed later.
TEP Reviewer#4	Results	Figure 9: Color figures so they are easy to read in black and white too.	We have revised the figures to improve readability.
TEP Reviewer#5	Results	Figure 8 (p 101 of 139): The right hand side of graph is cut off at the bottom. We cannot see the full results for the prospective studies and the RCTs. Please expand the y-axis	We have revised the figures in the report.
TEP Reviewer#5	Results	The detail is appropriate. The characteristics of the studies are clearly described, with key messages clearly communicated. Figure 8 has some formatting issues, but the vast majority of the figures and tables are clear and beautifully presented. See attached document for additional specific comments.	Thank you. We have corrected formatting for Figure 8.

Commentator and Affiliation	Section	Comment	Response
Public Reviewer (Allergen)	Discussion/ Conclusion	Allergan strongly recommends that AHRQ include all Ulipristal randomized controlled Phase III studies that meet all specified inclusion criteria (both Donnez 2012 publications) as we firmly believe that these studies are critical to review when assessing the efficacy and safety of Ulipristal in the final report. These studies employed only Ulipristal treatment related endpoints and no adjunctive surgery related endpoints, and they, combined with long-term treatment studies (Donnez et al 2014 and Donnez et al 2015) provide key Ulipristal efficacy and safety data for uterine fibroid treatment and should be included in the final report. Excluding these publications would not accurately reflect the published data on the medical management of symptomatic uterine fibroids with Ulipristal and therefore we stress that AHRQ should not endorse any final report without the inclusion of them.	We have included the Donnez et al 2016 publication "Long term management of uterine fibroids with ulipristal acetate" in our final report (this is a followup paper to the Donnez et al trial from 2015.
Peer reviewer #1	Discussion/ Conclusion	1. No definite conclusions due to insufficient data, sample size or number of studies. 2. Very clear that there is a paucity of well defined studies and therefore future direction is clearly outlined.	We agree. Thank you for your comment.
Peer Reviewer #2	Discussion/ Conclusion	For ES section on limitations of SR (pg ER-16) need to specify that for questions of effectiveness (KQ 1-2) there was restriction to RCTs, but different design inclusion for KQ 3-4. Current language makes it sound like restricted to RCTs for all KQ and that's not correct.	We have added details: "...to restrict to randomized clinical trials for the comparative effectiveness synthesis (KQs 1 and 2), …"
Peer Reviewer #2	Discussion/ Conclusion	SOE Tables 32-38, on pages 75-79, with the modified GRADE rating details (ROB, Limitations, Directness, Consistency, Precision, Reporting Bias) are very helpful. Thank you.	Thank you for your comments.
Peer Reviewer #2	Discussion/ Conclusion	Page 81, section on Applicability: the whole issue of whether the patient populations across treatments are similar or not is important is is really not addressed. Do patients offered or enrolled in studies of medications, or procedures, or surgeries differ from others who are offered or enrolled in studies of other possible treatments? For example, are women with future fertility concerns only enrolled in myomectomy studies because clinicians would never offer anything else? Are women with large subserosal fibroids really only offered hysterectomy? Basically, how do the patient populations across these studies differ? This section starts with the statement that the findings are "widely applicable to the general population of women . . . " and yet when I look at the evidence tables I'm not really convinced of this statement unless one adopts the attitude that women are all about the same. I'd like to see more nuanced discussion and analysis of this.	The age range and fibroid characteristics of women across studies are primarily pre-menopausal and similar in terms of fibroid characteristics. At times the inclusion and exclusion criteria for the study exert influence and may exclude women for whom the treatment is available. Future fertility desires are rarely or poorly captured as part of descriptions of populations. We do feel the data for each type of intervention reflects those who are clinically considered eligible for the interventions and thus the included studies have good applicability.

Commentator and Affiliation	Section	Comment	Response
Peer Reviewer #2	Discussion/ Conclusion	Page 89. The COMPARE study funded by PCORI/AHRQ is referred to in the ES, but not in main text of report. It should be added here.	This identical sentence also appears in the main report.
TEP Reviewer #2	Discussion/ Conclusion	Not enough time to get longer term info with UPA. What happens after treatment is unknown unlike with GnRH	The newer studies now included provide follow up to 6 months after treatment. We do note that this is only for a subset of those randomized who did not go on to have surgery and those who chose surgery may differ in substantial ways from the women available for 6 month follow up.
TEP Reviewer#4	Discussion/ Conclusion	Next intervention may be affected by insurance type. I see this is mentioned on next page.	Thank you for your comment.
TEP Reviewer#4	Discussion/ Conclusion	But to determine appropriate sample sizes we need to know what the clinically meaningful differences	Added determination of minimal important difference and patient centered outcomes to list of research needs.
TEP Reviewer#4	Discussion/ Conclusion	How about CART techniques to identify important factors?	Good point, since this is now the discussion section we added: "Likewise determinants of outcomes may be examined by use of tools such as classification and regression tree (CART) analysis to partition extant date in ways that better reveal the contribution of fibroid and patient characteristics to outcomes."
TEP Reviewer#4	Discussion/ Conclusion	Willingness to agree to randomization is not addressed	We do not believe this a major hindrance given success of large scale trials in other areas of intervention, including in populations that can be challenging to recruit.
TEP Reviewer#5	Discussion/ Conclusion	The implications of the major findings are clearly stated. Yes, the future research section is clear. I would suggest adding more about fertility and pregnancy outcomes, of course, because that is an interest of mine. See attached document for additional specific comments.	Thank you for your comments.
TEP Reviewer#5	General comments	Regarding feasibility of trials, vitamin supplementation (e.g., vitamin D) might be worth adding to list as an example.	We have added nutritional supplements to list of suggestions for future research.
Peer reviewer #1	General Comments	1. Yes, clinically meaningful 2. Target population defined but not adequately captured. 3. Audience defined 4. KQ clearly stated	We have added additional detail about study populations in several locations throughout report.
Peer reviewer #1	General Comments	As the authors correctly point out uterine fibroids present a clinically significant medical, Social and financial burden, the treatment of which remains inadequately addressed. This review is a well thought out and well written article that summarizes the state of the medical knowledge as published in English journal reports.	Thank you for your comment.
Peer reviewer #1	General Comments	The strengths include addressing highly relevant questions, large volume of reviewed literature and Meta analyses, an excellent recognition and detailing of the limitations in this field of medicine as it pertains to treatment decision making for women with symptomatic uterine fibroids.	Thank you for your comment.

Commentator and Affiliation	Section	Comment	Response
Peer reviewer #1	General Comments	The following is a list of concerns, which in all fairness the authors have highlighted themselves but are worth mentioning: 1: Studies analyzed: Are the symptoms across studies comparable? Are the sample sizes significant? As it relates to outcomes there are several categories with only 1 study. What clinical observation can be reliably gleaned from this result? KQ1: Low / insufficient SOE no data yet support expectant management as a “safe” choice p28 contradicts an earlier comment that there were inadequate expectant management patients. Would suggest changing the language.	In follow-up of 3 to 12 months fibroid characteristics and symptoms did not significantly change. We do not equate this with "safe" and note that the overall quality of the research is poor to inform care. However we do think we can grade the fibroid change and sympton stability as low level evidence. We have added a sentence about inability to judge safety and have clarified that the SOE for expectant management is insufficient.
Peer reviewer #1	General Comments	2: Are individual patient differences stratifiable? Were they adequately stratified for the purposes of this report? KQ2: not answered since the above is unclear.	We primarily sought to determine if the literature now contained trials explicitly designed and powered to determine for which women/which type of fibroids outcomes might differ. Though it is theoretically possible to conduct meta-regression for types of intervention, this was not within the scope of our analytic work. For simple descriptions very few studies identified outcomes by participant or fibroid characteristics, combining these without ability to adjust for confounders could introduce new biases.
Peer reviewer #1	General Comments	3. KQ3: At time of surgery-does that mean no preoperative assessment proved beneficial in excluding patients? If so what percent of these had proven malignancy at time of surgery?	Other than increasing concern for risk with age, we did not identify consistently reported factors that could be used with confidence for screening. This is in part because it is difficult to achieve both high sensitivity and high specificity when the condition to be detected is very rare. Even age > 50 would have poor diagnostic characteristics.
Peer reviewer #1	General Comments	4. Any specific modality of screening recommended to providers to minimize risk of operating on a patient presumed to have benign lesions demonstrating postoperative malignancy?	Other than increasing concern for risk with age, we did not identify consistently reported factors that could be used with confidence for screening. This is in part because it is difficult to achieve both high sensitivity and high specificity when the condition to be detected is very rare. Even age > 50 would have poor diagnostic characteristics.
Peer reviewer #1	General Comments	5. KQ4: How strong is this conclusion?	The literature is moving very rapidly. Between the peer review assessment that the approaches were not statistically different and the revision, the number of cases available in the literature for this model doubled as did the number of publications. The new summary is compatible with increased risk from power morcellation. However we have concerns about bias in the literature that may over estimate this risk and these will need to be weighed in context.
Peer reviewer #1	General Comments	6. re: Longitudinal f/u for UAE patients- Interval before retreatment; Limitations and patient selection; EMMY and REST trial and recurrence after UAE, very low number of studies and sample size. How about UAE and recurrence as it relates to fibroid size, location and number? Is it recurrence of same fibroid or de novo?	We aimed to consistently note the timing and overall span of follow-up which is better for UAE than for many interventions. We concur there is less literature than desirable but what is available is from good quality studies. There is not sufficient data other than the few items noted in the KQ2 section to address how fibroid or patient characteristics influence outcomes. This should be a priority for future research.

Commentator and Affiliation	Section	Comment	Response
Peer reviewer #1	General Comments	7. Why ablate the endometrium if bleeding is from a fibroid or were these patients demonstrating AUB independent of the contribution form a fibroid?	The source of the bleeding is not believed to be the fibroid. Ablation is used when the amount of the endometrium available to bleed is increased by increased cavity size and when fibroids are believed to contribute to disordered endometrial shedding during menses and poor stabilization in the early proliferative phase. These are unproven mechanisms but use of endometrial ablation to treat menorrhagia associated with fibroids is relatively common in practice.
Peer reviewer #1	General Comments	8. Any comments on hysterectomy and pelvic floor dysfunction? Supravcervical vs. total?	Other reviews address the relationship between hysterectomy and subsequent pelvic floor disorders. Other than mapping subsequent treatments after initial intervention for fibroids, we have not specifically sought long-term pelvic floor outcomes as they were not included as final health outcomes in our analytic framework developed a priori with expert guidance.
Peer reviewer #1	General Comments	9. Was the time to death changed by morcellation approach? P103 And is the stage and grade of tumor taken into account?	The updated analysis suggests survival time is reduced by power morcellation. Staging is not uniformly conducted or documented in this literature and the field as a whole notes that grading has low reproducibility. We did restrict the known LMS cases handled by intact removal of the uterus to Stage 1 tumor to better approximate the general status of women for who LMS is incidentally discovered. We sought to conduct sensitivity analyses adjusting for year of publication as a surrogate for secular trends in diagnosis and treatment put these models were not possible because of the heavy distribution of recent publications. Another approach that could accomplish this but which is not possible at this time would be meta-regression of primary data provided by the largest of these population based registries with relatively contemporary data.
Peer reviewer #1	General Comments	Rather bold statement to suggest that no difference between morcellation and intact removal of specimen versus non-power morcellation.	Indeed and the evolving data now suggest there is risk, though several large studies (including a Norwegian national dataset) from which related data could not be extracted suggest there is not increased risk. We have revised our findings based on an updated analysis conducted during peer review of the draft report. We have aimed for clarity and cautious interpretation of these evolving data.

Commentator and Affiliation	Section	Comment	Response
Peer reviewer #1	General Comments	Limitations as detailed by authors: RE: making decisions. The available literature has substantial gaps in collecting this information as indicated by the number of studies that addressed each of our eight primary outcomes: • Fibroid characteristics (e.g. change in size, number, volume): 51 • Symptoms status (e.g. bleeding, pain, bulk symptoms): 51 • Sexual function: 10 • Quality of life and satisfaction with outcomes: 8 • Desired fertility status: 1 • Pregnancy outcomes: 8 • Fibroid recurrence: 5 • Subsequent treatment for fibroids: 19 Little continuity exists in approaches to measuring outcomes and use of unvalidated measures are common.	Thank you for your comments.
Peer reviewer #1	General Comments	In summary a very valuable addition to the literature and highlights the need for more prospective and controlled studies to answer these questions. Conclusions need to be tempered in light of the limitations of the study acknowledging that at present time no strong treatment recommendations can be derived from these analyses despite several trends that may have been noted.	We have aimed to be consistent in descriptions of content of the literature and in providing summaries of strength of evidence. Few interventions reach high strength of evidence. That said EPC reports do not make recommendations to directly inform clinical care, rather we seek to systematically organizing information for use by stakeholders.
TEP reviewer #1	General Comments	Yes to all. The Key Questions are clearly defined and specific.	Thank you for your comment.
Peer Reviewer #2	General Comments	See multiple comments below in clarity and usability section and also in results section of this form. This report has multiple issues with poor copy editing and does not have a cohesive editorial voice. It reads like it was put together in a rush and that no one really pulled it together. The review is pretty well done and seems to be fairly accurate and complete, but my review identified several places where it could be made better.	We have revised the report extensively with attention to voice, clarity, copy edits, and consistency of terminology. Thank you for your detailed edits.
TEP Reviewer #2	General Comments	Overall this is comprehensive and provides much useful information. It is quite dense and so most will only read the summary which must be absolutely robust. It took me several hours to go through it and I am just including some of the points I felt to be most important for consideration.	Thank you for your careful review.
TEP Reviewer #2	General Comments	Care must be taken with small studies as it is so easy to draw inappropriate conclusions with wording of conclusions that is too strong which is particularly true with the fertility sparing treatments. To do a study with fertility as an endpoint requires 2000 women and is simply not possible. Many women in the studies of UAE and/or myomectomy do not actually want to conceive which complicates things further.	We have noted that information on fertility following these interventions is sparse and that studies assessing fertility outcomes are small.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer #2	General Comments	Consequently, I think some rethinking is needed in important areas such as fertility after UAE and myomectomy. Patients are being fed information without being aware of its limitations and uncertainty. I feel strongly that you should include this and reassess your view of the Mara paper in particular. Too much emphasis is being placed on conclusions from this and other lower quality studies (as assessed by most reviewers).	We have noted that information on fertility following these interventions is sparse and that studies assessing fertility outcomes are small. We considered the Mara study to have low risk of bias for selection, performance, reporting, and attrition bias. Even though allocation concealment was not explicitly elaborated, sequence generation was provided in sufficient detail and we feel that allocation is less likely to be foreseen in advance of treatment. The two areas scored as medium risk of bias were risk of detection bias since masking at time of outcome evaluations was unclear and "other" in relation to the fact that the treatment regimen in one arm defaulted to a clinical protocol if a certain treatment response was not achieved. Hence we considered this trial to be a 'fair' quality study.
TEP Reviewer #2	General Comments	I am unfamiliar with your method of grading evidence as I have been involved with using GRADE recently.	AHRQ EPC's strength of evidence (SOE) approach is similar to GRADE but assesses some different domains. EPC's assess the domains of study limitations (which incorporates risk of bias), consistency, directness, precision, and reporting bias and may assess additional optional domains. For more information about EPC SOE methods, please see Berkman et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. *Journal of Clinical Epidemiology* 2015. http://dx.doi.org/10.1016/j.jclinepi.2014.11.023
TEP Reviewer #2	General Comments	The REST study was published in NEJM and is considered a good quality study, again by most reviewers. Obviously I am biased but I would like to know why you rate this fair and EMMY which is generally considered as similar quality, as Good. As already mentioned, I appreciate this maybe a method of grading with which I am not familiar as it is neither levels of evidence or GRADE.	Thank you for the query about risk of bias assessment for this study. Our assessment of risk of bias encompasses multiple factors including selection, performance, reporting, attrition, detection and other bias. We mistakenly considered attrition bias as 'medium,' which in effect penalized this study for having an extended follow-up period. We agree overall the trial should have been scored as 'low' risk of bias as by the end of 5 years as there was only <10% lost to follow up. We have now updated the overall rating from 'fair' to 'good' quality study.
TEP Reviewer #2	General Comments	The GnRH agonist data is incomplete and I would like to have seen some of the pre-operative studies included as they provide useful information on fibroid size in particular.	Our exclusion criteria were as follows: "This review does not cover preoperative adjunctive treatments such as gonadotropin-releasing hormone (GnRH) agonists or intraoperative techniques, like use of cell savers that have established effectiveness as preoperative or adjunctive interventions to minimize blood loss or otherwise improve short-term operative outcomes." Studies of preoperative GnRH, unlike studies of ulipristal, were focused on attempting to decrease size of fibroids to avoid surgery and did not meet our inclusion criteria.
TEP Reviewer #2	General Comments	These are not the same as those discussing per-operative interventions and you comment on the use of UPA pre-operatively. This is somewhat inconsistent perhaps.	We did include some studies of ulipristal that were not focused on effect on surgical outcomes rather took advantage of populations of women who intended to have surgery in order to examine effects. These more recent studies also include follow-up of substantial proportions who did not proceed to surgery

Commentator and Affiliation	Section	Comment	Response
Public Reviewer (Allergen)	General Comments	Allergan appreciates the opportunity to provide comments on the Agency for Health Research and Quality (AHRQ) Management of Uterine Fibroids Draft Report. Allergan is a unique, global pharmaceutical company focused on developing, manufacturing, and commercializing innovative branded pharmaceuticals, and biological products for patients around the world. Our portfolio includes best-in-class products that provide valuable treatments in women's health, central nervous system, eye care, medical aesthetics, gastroenterology, urology, cardiovascular, and anti-infective therapeutic categories. Allergan is an industry leader in research and development, with one of the broadest development pipelines in the pharmaceutical industry. Allergan is committed to working with AHRQ, physicians, hospitals and patients to deliver innovative and meaningful treatments that help people around the world live longer, healthier lives. Allergan acquired the rights to develop and market Ulipristal upon approval in the United States, for the symptomatic treatment of uterine fibroids. Ulipristal (trade name Esmya®) is currently available in Canada and Europe. We appreciate the agency's willingness to review the appropriate management of uterine fibroids, which is a very common condition affecting millions of women, however, we **request that AHRQ review our recommendations and update the final report** accordingly.	Thank you for your review. Detailed responses to comments, including any revisions to the report, follow.
TEP Reviewer#3	General Comments	I believe overall the document is well researched, balanced and is a fair representation of our current knowledge in the field. Having said that, I have the following comments:	Thank you for your comment.
TEP Reviewer#3	General Comments	It is not wise to group uterine artery embolization with surgical uterine artery occlusion, or division or any other method of temporarily or permanently occluding the uterine artery before it reaches the margin of the uterus. The two types of procedures are completely different and have different outcomes. While in the short run, symptoms may be improved with both, later outcomes clearly favor uterine embolization and the imaging follow-up shows a much lower fibroid infarction rate for proximal uterine artery occlusion. Reference for this is: a. Uterine artery embolization versus laparoscopic occlusion of uterine arteries for leiomyomas: long-term results of a randomized comparative trial.Hald K, Noreng HJ, Istre O, Kløw NE.J Vasc Interv Radiol. 2009 Oct;20(10):1303-10 By putting these together, the outcomes from uterine embolization are underestimated.	The methods are grouped as a general category but the outcomes of occlusion by ligation or cautery are not included in the summaries of outcomes of UAE. UAE studies are cleary indicated and discussed as a unit. To assure this is clear we have combed the document for mentions of the term UAE and been certain we keep the procedures themselves distinctive. We have also revised the heading to save Uterine Artery Embolization and Uterine Artery Occlusion.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#3	General Comments	Uterine artery embolization does not cause a complete or permanent occlusion of the uterine arteries. The embolic material is carried to the fibroids from the main vessel by blood flow. Therefore, the embolic material gets to and occludes the end vessels to the fibroids. With surgical occlusion or division, the main vessel is severed, but this leaves room for ovarian flow, flow from the round ligament arteries and other collaterals to supply the fibroids. The infarction rates are therefore much lower, as noted in the paper above.	We revised the initial description to read: Uterine artery embolization (UAE) involves placement of a catheter through a blood vessel in the groin, using techniques similar to cardiac catheterization. Arteries serving the uterus or specific fibroids are then blocked by introducing an embolization agent to close off the blood flow to the fibroid(s). Similar but less selective techniques can be used to directly occlude the main uterine vessels with sutures or with coagulation at the time of open or laparoscopic surgeries.
TEP Reviewer#3	General Comments	The evidence clearly shows that menstrual bleeding is reduced with UAE in most studies and this is not reflected in the executive summary, (lines 11-17, page 22 of 139), and in fact in the body of the report- (lines 36-38, page 112 of 139) it states that it has modest to minimal effects on bleeding. This is in contradistinction to Table 16, on page 71 of 139 that shows where reported, menstrual bleeding was improved in all studies. Also on page 112, it should be called uterine artery embolization, not occlusion.	Thank you. We have corrected the SOE for UAE and bleeding to high and reviewed the entire report to assure that the summary data about UAE and bleeding is accurate throughout. It is important to note that page 22/139 is summary of data from comparative effectiveness trials (not the data presented earlier for overall effects of UAE). Comparing UAE to myomectomy the changes in bleeding, subsumed in our summary statement about symptom relief, were similar. We have revised Page 112 of 139:
TEP Reviewer#3	General Comments	Typographical error: in both the introductory paragraph (page 11 of 139, line17-18, it should be woman, not women. Same error in the main report, (page 31 of 139, line 28).	Corrected, thanks.
TEP Reviewer#3	General Comments	Page 32 of 139, line 8-10, this mischaracterizes uterine embolization. The intent is to occlude the fibroid blood flow with the embolic material, which is carried by uterine artery flow to the fibroid. It is not similar to surgical occlusion, and a distinction should be made.	We disagree and do not see these approaches as fundamentally different. We group UAE, uterine artery ligation, and uterine artery coagulation as three approaches that all seek to achieve the same anatomic result which is occlusion of the uterine artery or more distal vessels to reduce blood flow to fibroids. The approaches are different and of varied invasiveness but all three are methods of occlusion of vessels.
TEP Reviewer#3	General Comments	Typographical error: page 65 of 139, line 8, should be bleeding not bleedings.	Corrected, thanks.
TEP Reviewer#3	General Comments	Page 68 of 139, line 7- UAE is an interventional radiology procedure, not a vascular radiology procedure.	According to the subspeciality definitions interventional radiology (IR) is also known as vascular and interventional radiology (VIR) or surgical radiology, and is a sub-specialty of radiology providing minimally invasive image-guided diagnosis and treatment of diseases in every organ system. The American Board or Radiology recognizes fellowship training in Vascular and Interventional Radiology for board certification. For brevity we have edited to interventional radiology.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#3	General Comments	Page 91 of 139, starting at line 14, the authors should note that this study had a fundamental flaw in its design that limited the ability to determine reproductive outcomes. According to the protocol, for any patient that underwent UAE and whose imaging 6 months later showed any fibroid remnants 5 cm or larger, they were then taken for myomectomy. A third of the UAE patients underwent a secondary procedure, not driven by symptoms but by the arbitrary protocol design. Therefore, the number of re-interventions was much higher than for UAE- which should be noted here- and the reproductive outcomes are completely muddled, because the poorer reproductive results for UAE could be from having the myomectomy and the UAE. In addition, since it was recommended not to get pregnant for 6 months after the intervention, UAE patients had much less time to get pregnant than the myomectomy patients. Because all the reproductive advantages ascribed to myomectomy over embolization rest on this study, it must be noted in the description of the study and it should be noted the strength of the conclusions is even further weakened from this already weak study.	Excellent point. This aspect is now specifically noted in the summary and rates are noted to be likely to have been influence by their protocol which is described.
TEP Reviewer#3	General Comments	Typographical error: page 98 of 139, line47-48- dimeter should be diameter.	Corrected, thanks.
TEP Reviewer#4	General Comments	The writing needs major copy editing. I know this is a draft, but it was a distraction and the sections were not written with a single "voice". This report's writing is problematic and detracts immensely from the presentation of the work. I know this is a draft, but it's still really sloppy. I started typing notes below but it became excessive. I have instead made notes on the text and will attach them as a pdf.	With apologies for errors, we have conducted close editing of the revised version. Edits include those items noted in the pdf markup. While we sought to unite the style/voice of the writing, the document is in fact prepared by many authors and content, for instance methods versus synthesis, at times calls for different styles.
TEP Reviewer#4	General Comments	The authors discuss that more trials have been conducted in Europe and more trials need to be conducted in the US. The authors do not address the willingness of US fibroid patients to allow their treatment to be randomized between very different treatments. US patients are different than European (or other patients) in that there is greater direct appeal by pharma, device companies and providers to the consumer (patient). Women may go in to consultation with their provider having already decided on a treatment course and they may not deviate from that course to participate in trials.	This interpretation is appealing but we are not aware of research that finds that there are national differences, for instance European versus United States, in the willingness of individuals to participate in clinical trials. In many conditions for where direct marketing to consumers is common (diabetes, back surgery, orthopedic procedures), the United States is a leader in the conduct of clinical trials. Strong precedent exists in surgical and procedure-based trials that US women are willing to be randomized to very different treatment modalities for instance physical therapy vs surgery for stress incontinence and bladder botox injection vs implanted nerve stimulator for overactive bladder. If there is a population-level difference in willingness it does not preclude the need for evidence generated in U.S. populations and practice settings.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#4	General Comments	PCOs have not been systematically identified. As stated, more treatment goals of women should be incorporated into studies. However, other than the UFS-QOL, PCOs have not been formally identified through processes such as focus groups and in-depth interviews. PCOs of interest may also vary by patient characteristics.	We concur. Nonetheless with the input of key informants and the technical expert panel for this report, we selected measures that are important to women and represent known areas of symptoms and bother as well as individual priorities that women report motivates them to seek care and by which they assess success of interventions. Such measures can be considered *de facto* to be patient-centered as they assess status for symptoms that are known to be of relevance, even if relative importance is not fully characterized for individuals with varying values and treatment goals. We note that patient-reported outcome measures are not fully developed and validated in this field and have noted this in the Future Research section of the report.
TEP Reviewer#4	General Comments	Time to event stat approaches are great but other methods could be helpful. Other stat methods such as some of the regression tree analyses may help tease apart the relative importance of factors. To the end that randomization could be limiting participation of women in studies resulting in fewer studies and less evidence, other analyses incorporating statistical techniques such as instrumental variables or propensity scores could greatly improve the quality of observational studies.	Noted this in Analysis Methods in the discussion. For this literature and scope of our review it did not have direct relevance to other areas of the text.
TEP Reviewer#4	General Comments	Finally, the authors focus on the role of fibroid characteristics as effect modifiers. These characteristics could be viewed differently for addressing Patient-Centered care. With all the RCTs publishing their eligibility requirements which can include fibroid characteristics, the results from the different arms of the different RCTs using the same eligibility criteria could be crudely compared across studies and not just within studies.	As noted, few studies reported effect modifier data to allow for such comparisons.
TEP Reviewer#4	General Comments	The word “data” is plural (problematic throughout)	The AMA Style Guide (and other style manuals) now indicate data may be used in both plural and collective noun (singular) form: "Many now consider acceptable the use of data as a singular. In this usage, data is thought of as a collective noun and, when considered as a unit rather than as the individual items of data that compose it, it takes the singular verb." We have endeavored to establish consistency: "data from this study are clear" versus "across the included studies the data is clear".
TEP Reviewer#4	General Comments	Define intermediate outcomes	Definition of intermediate outcomes provided as: "We did not include studies reporting only intermediate outcomes such as technical success, conversion to alternate procedure, estimated blood loss during procedure, wound healing status, length of stay, and readmission or reoperation." In some instances we leave summary statements such as "brief" or "limited" as the full report is available for detailed synthesis underpinning these summaries.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#4	General Comments	"greater or lesser association" ? Meaning stronger or weaker associations?	We revised this sentence to read: Lastly, we combed these papers for data about whether characteristics of women or the masses believed to be fibroids were associated with leiomyosarcoma presence or modified the likelihood that morcellation during a surgery for fibroids would be associated with harm.
TEP Reviewer#4	General Comments	Is "robotic" really an approach? Isn't it a laparoscopic approach with a robotic assist?	We revised this sentence to read: "Interventions include hysterectomy via abdominal, vaginal, and laparoscopic approaches and those with robotic assistance; myomectomy via laparotomy, laparoscopy, hysteroscopy, or with robotic assistance; uterine artery occlusion via embolization, ligation, or coagulation; ablative procedures...."
TEP Reviewer#4	General Comments	Table ES-A: The ranges of lengths of followup would be helpful.	The mean and standard deviation of followup times are provided. The full range emphases the extremes. Full details are available in the outcomes tables of the main report.
TEP Reviewer#4	General Comments	Page ES-5: Should "watchful waiting" and "expectant management" be in quotes?	Thank you for your comment, but we do not see a need to change this text.
TEP Reviewer#4	General Comments	Page ES-5, line 35: What does "meaningfully" mean? There isn't enough information presented to conclude that women with fibroids should not expect that bleeding patterns will worsen over the near term (defined as???).	Few measures were statistically significant (two out of 17 calculations reported) and ranged from no growth (at one year) and increases of 4 to 17cc at 3 and 6 months. We added this material: "One study reported a four percent reduction in size and those that reported volume measures documented an average increase in size of about 9cc, which is less than one-fifth the size of a golf ball. "
TEP Reviewer#4	General Comments	Page ES-5, lines 51-52: "which included seven with…" seven what?	We have edited the text to clarify:"...which included seven studies of "add-back" therapy (addition of a second agent to a GnRH agonist)."
TEP Reviewer#4	General Comments	Page ES-5, lines 56-57; ES-6 line 26: Were percentages provided? Absolute values are not that informative since the size of the original fibroid/fibroid volume is not known.	We prefer absolute change in size for reporting throughout the report. A decrease of 10% in volume a 100cc fibroid is 10cc (1/4th of a golf ball) while a decrease of 10% in a 30cc is 3cc (<1/10 if a golf ball). The absolute reduction in volume is more substantial in the former than the latter and we and the Key Informants believed it is a superior way of capturing physical change in fibroids. It was also the most directly available data in the literature as authors most often reported measures of volume before and after intervention.
TEP Reviewer#4	General Comments	Page ES-6, lines 6-7: Clinical or statistical significance?	We have reduced text here for brevity. Exact data and statistical testing is available in text and tables of full report.
TEP Reviewer#4	General Comments	ES-6, line 34: What are the percentages?	We have revised this text to include percentages.
TEP Reviewer#5	General Comments	Yes, this report is clinically meaningful and the audience is clearly defined. The key questions are also very appropriate and explicitly stated. See attached document for my specific comments.	Thank you for your comments.

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#5	General Comments	The present report evaluates and summarizes the evidence about management of uterine fibroids. The authors abstract data regarding the effectiveness of interventions, risks of harm, and whether individual or fibroid characteristics influence outcomes. The authors have done a phenomenal job synthesizing all the data from various studies in the field. This report is well-written and thorough. The authors have appropriately acknowledged the lack of adequate evidence in the field and that so much more research that remains to be done for women to make the most informed decisions about their care. I have only minor comments, which are listed below.	Thank you for your comments.
TEP Reviewer#5	General Comments	In first part of report (brief summary), why are fertility and pregnancy outcomes mentioned for some procedures and treatments but not others? Please try to make this consistent across the document.	We have added the following sentence added to Outcomes: "We sought to collect outcomes uniformly for all interventions, however if data were not available for a selected outcome within an intervention category the outcome is not listed."
TEP Reviewer#5	General Comments	Abstract-Results: “Subsequent intervention rates were lowest for initial medical management at two years of follow-up; higher for myomectomy and UAE especially among younger women. Also, I suggest moving the words “at two years of follow-up” to the beginning of this sentence for optimal clarity.	We have edited this text.
TEP Reviewer#5	General Comments	The word “data” is plural. Page 13, line 12. “as is data to determine” should read “as are data to determine.” Page 26, line 6: “There were insufficient data…”	The AMA Style Guide (and other style manuals) now indicate data may be used in both plural and collective noun (singular) form: "Many now consider acceptable the use of data as a singular. In this usage, data is thought of as a collective noun and, when considered as a unit rather than as the individual items of data that compose it, it takes the singular verb." We have endeavored to establish consistency: "data from this study are clear" versus "across the included studies the data is clear".

Commentator and Affiliation	Section	Comment	Response
TEP Reviewer#5	General Comments	At various points throughout the report, the authors use the term “rate” when referring to cumulative incidence (probability/percentage): For example, on page 17 line 46. “No women receiving UAE required transfusion; major complication rates during and following UAE ranged from 1.2 to 6.9 percent periprocedurally, up to about 5 percent at two years. The rate of major complications was high in two studies that reported longterm followup (21% at 5 years in the REST trial and 16.8% at 32 months in a second study) in part because they considered a subsequent procedure a complication.” Replacement with the term “probability” can rectify this problem: “No women receiving UAE required transfusion; the probability of major complications during and following UAE ranged from 1.2 to 6.9 percent periprocedurally, up to about 5 percent at two years. The probability of major complications was high in two studies that reported long-term follow-up (21% at 5 years in the REST trial and 16.8% at 32 months in a second study) in part because they considered a subsequent procedure a complication.”	We have searched for the word rate and corrected instances in which it does to refer to events per unit time.
TEP Reviewer#5	General Comments	“Compared to” should be replaced with “compared with” throughout.	We have replaced "compared to" to "compared with" throughout the document per AHRQ standard.
Peer reviewer #1	Clarity and Usability	Most certainly highlights the need for further extensive research into the management of uterine fibroids and women's health. I believe the findings will be highly relevant to policy and practice decisions. Unfortunately this literature anlaysis, although very important, does not provide enough definitive data to make such conclusions with confidence.	We concur. Many critical gaps in knowledge remain.
Peer Reviewer #2	Clarity and Usability	This MS simply has not been well edited and I hope that these deficiencies will be corrected in the finalization process. These errors are really too numerous to mention, but here are some examples: Lack of units or inconsistency of use of units in column headers on tables. Nearly all tables have this issue (e.g., Table ES-A, on pg ES-3, lines 51-52: "Mean followup duration" doesn't have units. Assume it is months, but it could easily be another unit of time.	We have revised the report extensively to improve clarity and correct errors.
Peer Reviewer #2	Clarity and Usability	Another example is on pg. 19, Table 6, where the column for "Uterine Size" has no units. To the left and right the units are cubic cms, but it needs to be specified for this column as well.	We have revised the table and improved the comparability of the headers throughout the report.
Peer Reviewer #2	Clarity and Usability	Table 13 on pg 35 is another example of the same sort of rather sloppy copy editing.)	We have revised the table and improved the comparability of the headers when the table content is similar.

Commentator and Affiliation	Section	Comment	Response
Peer Reviewer #2	Clarity and Usability	"Data" is plural. It should be "Data are . . . " rather than "data is . . . " (for example, pg ES-17, line 14, but you should do a search and replace because this error occurs multiple times).	The AMA Style Guide (and other style manuals) now indicate data may be used in both plural and collective noun (singular) form: "Many now consider acceptable the use of data as a singular. In this usage, data is thought of as a collective noun and, when considered as a unit rather than as the individual items of data that compose it, it takes the singular verb." We have endeavored to establish consistency: "data from this study are clear" versus "across the included studies the data is clear".
Peer Reviewer #2	Clarity and Usability	There are some awkward sentences. For example, The entire document does not read as if it has a single editorial voice. Very choppy from section to section and needs a strong editorial wash.	We have revised the report extensively to improve the flow of the text, to enhance clarity, and to correct typographical errors.
TEP Reviewer#5	Clarity and Usability	The report is very well-structured and organized. It was easy to read. The conclusions are indeed relevant to policy, though the research is still limited for many of the treatment options (not the fault of the authors). See attached document for additional specific comments.	Thank you for your comments.