A well designed, described and executed RCT. No obvious statistical flaws, only a query re 25% non-inferiority margin – seems large. Please justify.

We have given a justification for this is given in the Statistical Analysis section of the Methods. Statisticians from Birmingham Clinical Trials Unit have been involved in all stages of this project. ‘The pre-specified non-inferiority margin of 25% (relative reduction) was based on the assumptions that outpatient polypectomy would be more convenient for women and cheaper, permitting it to be considered the treatment of choice even if fewer women had alleviation of bleeding symptoms’. Based upon clinical judgment and the fact that in our pilot study many women wanted to have outpatient treatment once they knew it was available, the margin was thought to be appropriate. However, we have now added a comment in the Discussion, noting this as a limitation of the study on page 11 ‘The non-inferiority level of 25% might be considered large by some and hence a limitation of the study, however this was selected based upon the perceived advantages of outpatient treatment. These advantages include avoiding a general anaesthetic, the convenience of being treated immediately after diagnosis, not requiring more than half a day’s absence from work, not having to arrange for childcare and having fewer hospital appointments. Ultimately, the results of this study showed that outpatient polypectomy was at worst, only 18% less effective than inpatient treatment.’

• There are a lot of outcomes in Tables 2 and 3 - are they all secondary? The trial registry lists 4 secondary outcome measures as of 01/09/2011 and another 9 previous secondary outcomes. Has the new list replaced the original list or not? Please clarify this.

We believe all the patient completed outcomes listed in Table 3 were included in the registry. For Table 2, we admit that apart from the safety outcomes and operative success the other variables that we have analysed were not included in the registry listing. However, we feel it is
important to include these other variables in the publication as they describe the detail of the operative procedures used in the two randomised groups (use of speculum, hysteroscopic removal, length of treatment, etc.) and felt that this information aided interpretation of the trial results. In the ISRCTN registry, the revised secondary outcomes overlap with those originally listed, but not all are reported here e.g. the cost-effectiveness analysis.

**Please do not include the rejected CEA.**
We have not included the economic analysis.

**First and foremost, please revise your paper to respond to all of the comments by the reviewers. Their reports are available below.**
**Please also respond to the additional comments by the committee.**
We have addressed all reviewers’ comments below.

<table>
<thead>
<tr>
<th>Reviewer #1 - Jan Bosteels  (MD, PhD) CEBAM, the Belgian branch of the Dutch Cochrane centre, Leuven, Belgium</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Originality</strong></td>
</tr>
<tr>
<td>The present manuscript adds to the limited evidence in the published literature. In- and outpatient treatment of symptomatic uterine polyps are compared using a non-inferiority (one-sided) approach which is rational from the clinical viewpoint for this topic and sensible from the statistical viewpoint.</td>
</tr>
<tr>
<td><strong>Importance of work to general readers</strong></td>
</tr>
<tr>
<td>The findings of this manuscript are very relevant to everyday clinical practice. An evidence-based fundament is established for the widespread practice of removing symptomatic uterine polyps suggesting that the outpatient treatment is not inferior compared to inpatient approach but may be preferable for other reasons e.g. patient convenience to leave the hospital sooner and return to daily home and professional activities.</td>
</tr>
<tr>
<td><strong>Overall design of study</strong></td>
</tr>
<tr>
<td>The use of a multicentre pragmatic RCT to address the effectiveness of a surgical intervention is very adequate.</td>
</tr>
</tbody>
</table>
**Participants studied**  
Adequately described with clinically relevant in-and exclusion criteria.

**Abstract/summary/key messages/What this paper adds**  
This section reflects accurately what the paper says.

| Comments | * Scientific reliability  
Although the aims of the present study are clearly defined in the last paragraph of the introduction and in section 1.6 of the OPT trial protocol I would suggest that the authors clearly write out the full PICO(T) research question of this RCT. | We have added this to the introduction on page 3, however, we feel that the research question is clearly defined and the reviewer also comments on this. Therefore, if the editors feel that it does not add to the paper we are happy for it be removed again. |
| --- | --- | --- |
| *Methods  
The methods are adequately described and complying with the CONSORT statement and checklist. I would suggest to discard the use of the euphemism ‘masking’ in favor of ‘blinding’ as explained by David Sackett in his book ‘Clinical epidemiology’ on page 94: masks generally do not prevent blindness by allowing the wearer to see what is actually going on through the provided holes! There are no ethical concerns. The DMEC is independent and the relationship and actions to be taken dependent on the interim analyses by communication with the independent TSC are clearly described and concordant with good practice in clinical research conduct. | Masking has been changed to blinding |
| *Results  
The results answer the research question in a credible way. The interpretation of the clinical findings is adequately translated into clinical daily practice by the use of NNT (numbers needed to treat). Since the overall purpose in medicine is to ‘firstly do no harm’ I would suggest to replace ‘number-need-to-harm’ by the more appropriate use of NNTH (number needed to treat for a harm) as suggested by the Cochrane collaboration in its glossary of Cochrane terms. | Thank you, we have changed this on page 8. |
| *Interpretation and conclusions  
The interpretation is adequately based on the study findings without | We do comment on this on page 11, final |
Data dredging or post hoc analyses. The conclusions are exclusively clinical (implications for clinical practice); I likewise suggest to add a sensible future research agenda if appropriate.

<table>
<thead>
<tr>
<th>paragraph</th>
<th>Thus the OPT trial is timely, novel and relevant to contemporary clinical practice. Further randomised controlled trials will be needed to evaluate the merits of innovative outpatient interventions against conventional inpatient practices across all surgical specialties.</th>
</tr>
</thead>
</table>

**References**

The references are up to date and relevant. On page 3 in the last paragraph of the introduction there is a reference between brackets (cross reference the economic paper). I presume this is a minor omission and would ask the authors to insert this reference as an economic study was an integral part of this trial. Readers will be much interested in reading this complementary paper as well.

The economics paper has been submitted to the British Journal of Obstetrics & Gynaecology and will hopefully be accepted for publication. We will supply this reference as soon as it is available. If the economics paper is not accepted prior to publication we will reference the upcoming HTA report (in press- this will not be published before the journal papers arising from the OPT trial have been published).
Regarding the paper “A randomised trial of outpatient versus inpatient uterine polyp treatment for abnormal uterine bleeding”.

This is a very important paper, contributing with scientific evidence for the ongoing implementation of outpatients procedures. This is the first randomised trial to support that outpatient polypectomy procedures are not inferior to inpatient polypectomy procedures. Outcome measures for the present study are primarily the women's assessment of bleeding at six months, secondarily generic and disease specific quality of life, procedure feasibility and acceptability. **Unfortunately the important economic evaluation is not included in the present paper.**

The presented numbers of failure vs. success and acceptability demonstrates clearly the (expected) differences between the outpatient and inpatient procedures. Equal percentages of women found the procedure unacceptable. However the assessment of bleeding leads to several challenges. As stated in the text and in Table A7 concurrent treatment after polypectomy was necessary in about 19% in both the outpatient and inpatient polypectomy group. Uterine polyps are a frequent finding in asymptomatic women, and abnormal uterine bleeding is frequently not caused by the polyp in patients with both AUB and polyp, which is also demonstrated by the present numbers. In postmenopausal polypectomy does not change the time to re-bleeding, suggesting that the polyp is not causative of postmenopausal bleeding (Timmermans Obstet Gynecol 2008).

Two systematic reviews of studies of polypectomy are quoted in the introduction, a 75-100% alleviation of symptoms is found in the systematic reviews (contrasting the expected treatment success of 80-90% in the sample size calculation). The authors of reference no 8 concludes that concomitant treatment is necessary and refers to a single RCT comparing transcervical resection of polyp compared with no treatment finding no difference in pictorial blood assay chart, but improved subjective VAS and alleviation of “other gynaecological symptoms” (e.g. vaginal discharge) which is not evaluated in the present paper.

MMAS, bleeding duration VAS and bleeding amount VAS are all subjective measurements of menorrhagia/intermenstrual bleeding, the more objective and validated pictorial blood assay chart could have been more relevant in this large multicenter study. Several of the Tables (appendix) demonstrates no difference between the groups with regard to abnormal uterine bleeding, **an inclusion of the economic evaluation could have been more interesting to the reader.**

<p>| Scientific Comments | Page 7 l 34 It is noteworthy that 13% of the women randomised for the inpatient procedure chose the outpatient procedure and only 6% of the women randomised for the outpatient procedure chose the inpatient procedure. This could be commented on page 11. | This is interesting and may reflect the unwillingness of anaesthetic staff to give a general anaesthetic to some women when an outpatient option is available. It may also reflect patient preference. However, we can |</p>
<table>
<thead>
<tr>
<th>BMJ Reviewer comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>November 2014</td>
</tr>
<tr>
<td>Page 6 of 12</td>
</tr>
</tbody>
</table>

only speculate on this issue and therefore we do not feel that we can give a robust rationale within the discussion. A cohort study which ran alongside this RCT fully explores patients’ preferences for treatment with 81% choosing outpatient treatment. This paper will be published separately.

<table>
<thead>
<tr>
<th>At page 4 line 53, dilatation and curettage is mentioned as an option for polyp removal- was all specimens analysed by histology to confirm the removal of polyp, and was there a difference in success-rate according to the technique used (e.g. D&amp;C vs resectoscope)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histology was not followed up within the trial. The reviewer describes a potential prognostic study with successful removal as the outcome. Whilst this is an interesting question, we do not believe we can do justice to such an analysis within this manuscript and that prognostic indicators of success would make an interesting paper in its own right. Indeed, we plan to combine the data from this trial to data from a non-randomised group we have collected to give us a larger cohort for examination of ‘what predicts?’ type questions.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P 5 procedure: The skills of the surgeon (i.e. how many procedures per year) is relevant if obtainable. In the Table consultant/not consultant is reported. It is noteworthy that the serious adverse events only occurred in the inpatient group (p8 l.37), this could be commented in the discussion (or included in the abstract). It is possibly that a conscious woman would react with pain before false passage/perforation.</th>
</tr>
</thead>
</table>
| Surgical efficiency is a major confounder in surgical trials but uterine polypectomy is considered relatively simple and is conducted by consultant and non-consultant grades. We do not have the data regarding surgeon procedure numbers. We do comment on page 10 that ‘Despite
<table>
<thead>
<tr>
<th>P 10, pain scores on VAS, low VAS scores are reported, they could be commented using the results from Collins SL The visual analogue pain intensity scale: what is moderate pain in millimeters, Pain 1997; 72:92-7</th>
<th>Thank you, we have added detail on page 10 so that it now reads ‘Pain scores during the outpatient procedure were of moderate intensity but low postoperatively’ and our qualitative research (reference to be added) suggested that women felt that the discomfort of outpatient treatment was outweighed by convenience’ Please note that the qualitative paper has been submitted to the British Journal of Obstetrics and Gynaecology and if it is accepted before publication of our RCT paper we will reference the British Journal of Obstetrics and Gynaecology and if not we will reference the HTA report.</th>
</tr>
</thead>
<tbody>
<tr>
<td>More detail should be provided in the Statistical Analysis section to address this point</td>
<td>The important factors in outpatient procedures are acceptability and patient preferences for treatment, which we have fully explored both within the study and in accompanying</td>
</tr>
</tbody>
</table>

fewer serious complications and the convenience of avoiding hospital admission and general anaesthesia, the outpatient procedure was associated with more technical failures, increased post-operative pain and reduced patient acceptability. We agree with the observation made by this reviewer but as the study was not powered for evaluating the safety of the procedures we did not want to draw conclusions based upon small number of serious complications.
This is likely not a major issue with respect to the primary analysis. The expected effect of appropriately accounting for stratification in the interval estimate would be to increase precision, narrowing the confidence interval, and thus not changing the decision to conclude non-inferiority. However, if the nominal confidence level is to be achieved, the analysis should correspond to the randomization scheme.

As the reviewer points out the effect of the adjusted analysis would be to reduce the size of the confidence interval and make it more likely to declare non-inferiority. Hence, our unadjusted analysis could be viewed as conservative. We accept this point though and have already provided adjusted analysis for benefit of the reader in the appendix Table A1. There was only a minor change to the size of the confidence intervals.

This reference was cited due to errors with the referencing software. It is no longer cited within the manuscript.

The key to recruiting a large multicentre surgical study is recruiting across a large number of hospitals and clinics. All acknowledged clinicians took an active role in recruitment. We have acknowledged the study participants on page 13.

There was a problem with our referencing software and the references have been revised accordingly.
<table>
<thead>
<tr>
<th>Comments</th>
<th>Lacks a detailed description in the materials and methods of surgical technique for both inpatient and outpatient polypectomy. Lacks a description of anaesthesiologic technique.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>This was a pragmatic trial and under the title ‘Procedures’ on page 4 we have described the techniques. Table 2 also gives detailed information regarding the operative techniques.</td>
</tr>
<tr>
<td>Page 4 line 54: curettage?</td>
<td>Curettage is an accepted term. No revision / explanation required.</td>
</tr>
<tr>
<td>Page 5 line 1: clinicians were free to choose the technique. Patients should be randomized</td>
<td>This refers to technique of removal, not treatment setting. We have changed the sentence on page 5 so that it now reads ‘Clinicians were free to choose the technique for polypectomy post randomisation.’</td>
</tr>
<tr>
<td>Page 8 line 18: failed removal in 7% inpatient group. Line 25: reoperation 92%. Why? There is not uniformity of surgeons</td>
<td>We think that the reviewer has conflated the comparisons mistakenly. With reference to the operative detail on page 8, of the 25 failed outpatient procedures, 23 had repeat treatment as inpatients. We have added a denominator to the number of failed outpatient procedures to clarify this ‘25/46 (54%) of the failed outpatient removals were immediately scheduled for reoperation, usually as an inpatient (23/25, 92%).’</td>
</tr>
</tbody>
</table>
Reviewer #4  -  Ertan Saridogan University College London Hospitals

This manuscript describes a pragmatic noninferiority RCT comparing outpatient and general anaesthetic polypectomy for treatment of abnormal uterine bleeding (AUB). The study is clearly well-designed and conducted, primary and secondary outcome measures well-defined and analysed. The manuscript is well-written, the main text gives the key messages in a concise manner suitable for the BMJ style but further details can be found in the appendices.

I have a number of comments on the design which may need clarification:

<table>
<thead>
<tr>
<th>Comments</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>The sample calculation was made on the basis of ability to remove polyps under GA vs outpatients, but the primary outcome measure was successful treatment of AUB as assessed by the woman. In principle, it is appropriate to use the patient reported outcome as the primary measure and the results given both for this and also for successful removal of polyp. It would be useful to clarify this point in the discussion section.</td>
<td>The sample size was based upon patient reported outcomes. We have clarified this by changing the sentence to ‘The sample size was based on data suggesting that patient rated treatment success at six months would be 90% for inpatient polypectomy and 80% for outpatient polypectomy’</td>
</tr>
<tr>
<td>The study assumes that the polyps found are the cause of AUB the patient is complaining of, this may not be the case. The polyps are relatively common and may be incidental findings. There may be some differences in causality, depending on the type of AUB. For example postmenopausal or intermenstrual bleeding may be more likely to be due the polyp found, whereas the polyp may be unrelated to heavy menstrual bleeding. Hence, this may have some impact on the expected impact on the AUB outcome as assessed by the patient. The authors do give an impact analysis in Table A2 and there doesn't seem to be differences in these subgroups. It would be useful to add a few sentences to expand the discussion on this</td>
<td>We mention this in the first paragraph of the Discussion on page 10 ‘There was no evidence that successful resolution of symptoms varied by primary bleeding complaint, or polyp type and location.’</td>
</tr>
<tr>
<td>I agree with the authors that the pain score comparison needs to be interpreted with caution. The GA group patients would be given free access to oral and parenteral pain relief as opposed to the OPD patients, hence this is not a fair comparison. The</td>
<td>Thank you for this comment. We agree that patient experience is of major importance and we have evaluated this in a qualitative study which been submitted to the British Journal of</td>
</tr>
<tr>
<td>Requirement for analgesics was not compared, as far as I can see. Most OPD patients would normally leave the clinic in less than 60 minutes, in this study they must have been asked to stay to give their pain score. It is not clear from the Table 3 when they were discharged, I expect this would have been left to the discretion of the clinician. (Patient acceptability is probably a more valuable comparison for this purpose and this has been given in the manuscript)</td>
<td>Obstetrics and Gynaecology.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>How was the distinction between glandular and fibrous polyps made? Why was it thought to be significant?</td>
<td>Polyps are made up of varying amounts of stroma, glands and fibrous tissue. The proportion of these components determines whether they are glandular or fibrous. The hysteroscopic diagnosis was based upon clinical evaluation. A glandular polyp was defined as one that moved with the flow of the distension media and was soft when compressed. The reference for the hysteroscopic definition a uterine polyps has been given in the Methods ‘All women with abnormal uterine bleeding and a uterine polyp diagnosed at outpatient hysteroscopy(13) were eligible to be recruited into the OPT trial.’ This was thought to be important because glandular polyps are thought to be easier to remove through the narrow cervical os without the need for cervical dilatation in contrast to more fibrous polyps.</td>
</tr>
<tr>
<td>Was the impact of type of polypectomy (i.e under direct vision or blind avulsion) recorded and analysed? This is likely to be important for a couple of reasons; firstly blind avulsion would</td>
<td>The reviewer describes a potential prognostic study with successful removal and pain as outcomes. Whilst these are interesting</td>
</tr>
</tbody>
</table>
usually require some cervical dilatation and can potentially compromise acceptability/pain scores in the outpatients, and secondly it is more likely to lead to failed or incomplete removal which increases the risk of persistent symptoms and reintervention.

questions, we do not believe we can do justice to such an analysis within this manuscript and that prognostic indicators of success/pain would make an interesting paper in its own right. Indeed, we plan to combine the data from this trial to data from a non-randomised group we have collected to give us a larger cohort for examination of ‘what predicts?’ type questions.