National Institute for Health and Care Excellence

Final

Endometriosis: diagnosis and management

Full guideline

NICE guideline NG73

Methods, evidence and recommendations

September 2017

Final version

Developed by the National Guideline Alliance, hosted by the Royal College of Obstetricians and Gynaecologists

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Update information

November 2024: We have reviewed the evidence for, and updated and made new recommendations on, diagnosis of endometriosis. As part of this update some recommendations on signs and symptoms were also amended. These new and updated recommendations are marked **[2017, amended 2024]** or **[2024]** and can be seen in the main version of the guideline at http://www.nice.org.uk/guidance/NG73. Some of the evidence in this document has also been updated by the evidence review associated with these updated recommendations, which is linked in the main guideline.

Additionally, the diagnosis and referral section has been reviewed and reordered to better reflect current clinical practice and the care pathway. Unless otherwise indicated no changes have been made to recommendations that have been reordered.

We have also made an editorial update to a recommendation in the section on information and support to clarify what information should be provided and when. This recommendation is marked in the main guideline as **[2017, amended 2024].**

April 2024: We have reviewed the evidence and updated some recommendations, and made a recommendation for research, on treatment of endometriosis when fertility is a priority. These recommendations are marked **[2017, amended 2024]** or **[2024]** and can be seen in the main version of the guideline at http://www.nice.org.uk/guidance/NG73. Some of the evidence in this document has also been updated by the evidence review associated with these updated recommendations, which is linked in the main guideline.

ISBN 978-1-4731-2661-9

Contents

ISBN 978-1-4731-2661-9

ISBN 978-1-4731-2661-9

ISBN 978-1-4731-2661-9

ISBN 978-1-4731-2661-9

ISBN 978-1-4731-2661-9

ISBN 978-1-4731-2661-9

1 Introduction

Endometriosis is one of the most common gynaecological diseases needing treatment. It is defined as the growth of endometrial-like tissue (the womb lining) outside the uterus (womb). Endometriosis is mainly a disease of the reproductive years and, although its exact cause is unknown, it is hormone mediated and is associated with menstruation.

Endometriosis is typically associated with symptoms such as pelvic pain, painful periods and subfertility. Endometriosis is also associated with a lower quality of life. Women with endometriosis report frequent, or chronic, or severe pain, tiredness, more sick days, and a significant physical, sexual, psychological and social impact. Endometriosis is an important cause of subfertility and this can also have a significant effect on quality of life.

Women may also have endometriosis without symptoms, so it is difficult to know how common the disease is in the population. It is also unclear whether endometriosis is always progressive or can remain stable or improve with time.

Delayed diagnosis is a significant problem for women with endometriosis. Patient self-help groups emphasise that healthcare professionals often do not recognise the importance of symptoms or consider endometriosis as a possibility. In addition, women can delay seeking help because of a perception that pelvic pain is normal. Delays of 4 to 10 years can occur between first reporting symptoms and confirming the diagnosis. Many women report that the delay in diagnosis leads to increased personal suffering, prolonged ill health and a disease state that is more difficult to treat.

Diagnosis can only be made definitively by laparoscopic visualisation of the pelvis, but other, less invasive methods may be useful in assisting diagnosis, including ultrasound. Management options for endometriosis include pharmacological, non-pharmacological and surgical treatments. Endometriosis is an oestrogen-dependent condition. Most drug treatments for endometriosis work by suppressing ovarian function, and are contraceptive. Surgical treatment aims to remove or destroy endometriotic lesions. The choice of treatment depends on the woman's preferences and priorities in terms of pain management and/or fertility.

Endometriosis can be a chronic condition affecting women throughout their reproductive lives (and sometimes beyond). Women's priorities and preferences may change over time, and management strategies should change to reflect this.

Women with endometriosis typically present to community services (including GPs, sexual health services, practice nurses and school nurses) with pain, and may then be referred to gynaecology services for diagnosis and management. Some women may present to fertility services. Complex surgical treatment is carried out in specialist endometriosis services (endometriosis centres), which incorporate a multidisciplinary team.

This guideline makes recommendations for the diagnosis and management of endometriosis in community services, gynaecology services and specialist endometriosis services (endometriosis centres).

The guideline also covers the care of women with confirmed or suspected endometriosis, including recurrent endometriosis. It includes women who do not have symptoms but have endometriosis discovered incidentally. Special consideration was given to young women (aged 17 and under). The guideline does not cover the investigation of fertility problems related to endometriosis, care of women with endometriosis occurring outside the pelvis, nor postmenopausal women.

ISBN 978-1-4731-2661-9

2 Guideline summary

2.1 Committee membership, National Guideline Alliance (NGA) staff and acknowledgements

Table 1: Committee members

Name	Role
Rachel Brown	General Practitioner
Dominic Byrne*	Consultant Gynaecologist
Natasha Curran	Consultant in Anaesthsia and Pain Medicine
Alfred Cutner*	Consultant Gynaecologist
Cathy Dean	Endometriosis Clinical Nurse Specialist
Lynda Harrison	Lay member
Jed Hawe	Consultant Obstetrician and Gynaecologist
Lyndsey Hogg	Lay member
Andrew Horne	Professor of Gynaecology and Reproductive Science
Jane Hudson-Jones	Lay member
Wendy-Rae Mitchell	Endometriosis Clinical Nurse Specialist
Caroline Overton (Chair)	Consultant Gynaecologist
Carol Pearson	Clinical Commissioning Group Lay Member for Governance
Co-opted members	
Moji Balogun	Consultant Radiologist
Christian Becker	Associate Professor of Gynaecology and Reproductive Medicine
Mohammed Belal	Consultant Urological Surgeon
Sanjiv Manek	Consultant Gynaecological Pathologist
Natalie Lane	Clinical Psychologist (pelvic pain)
Tim Rockall	Consultant Colorectal Surgeon

**Job share position*

Table 2: NGA Staff

Name	Role
Alexander Bates	Senior Health Economist
Zosia Beckles	Information Scientist (until March 2016)
Shona Burman-Roy	Senior Systematic Reviewer (from May 2016)
Anne Carty	Project Manager (September to October 2016)
Melanie Davies	Clinical Adviser
Katharina Dworzynski	Guideline Lead (from March 2015)
Annabel Flint	Project Manager (until August 2016)
Maryam Gholitabar	Research Fellow (until January 2016)
Elise Hasler	Information Scientist (from April 2016)
Sadia Janjua	Systematic Reviewer (from November 2015 to March 2016)
Laura Kuznetsov	Systematic Reviewer (from April 2016)
Sabrina Naqvi	Project Manager (from October 2016)

ISBN 978-1-4731-2661-9

Name	Role
Amir Omidvari	Research Fellow (from January 2016 to March 2016)
Hugo Pedder	Statistician
Ferruccio Pelone	Assistant Systematic Reviewer (from August 2016)

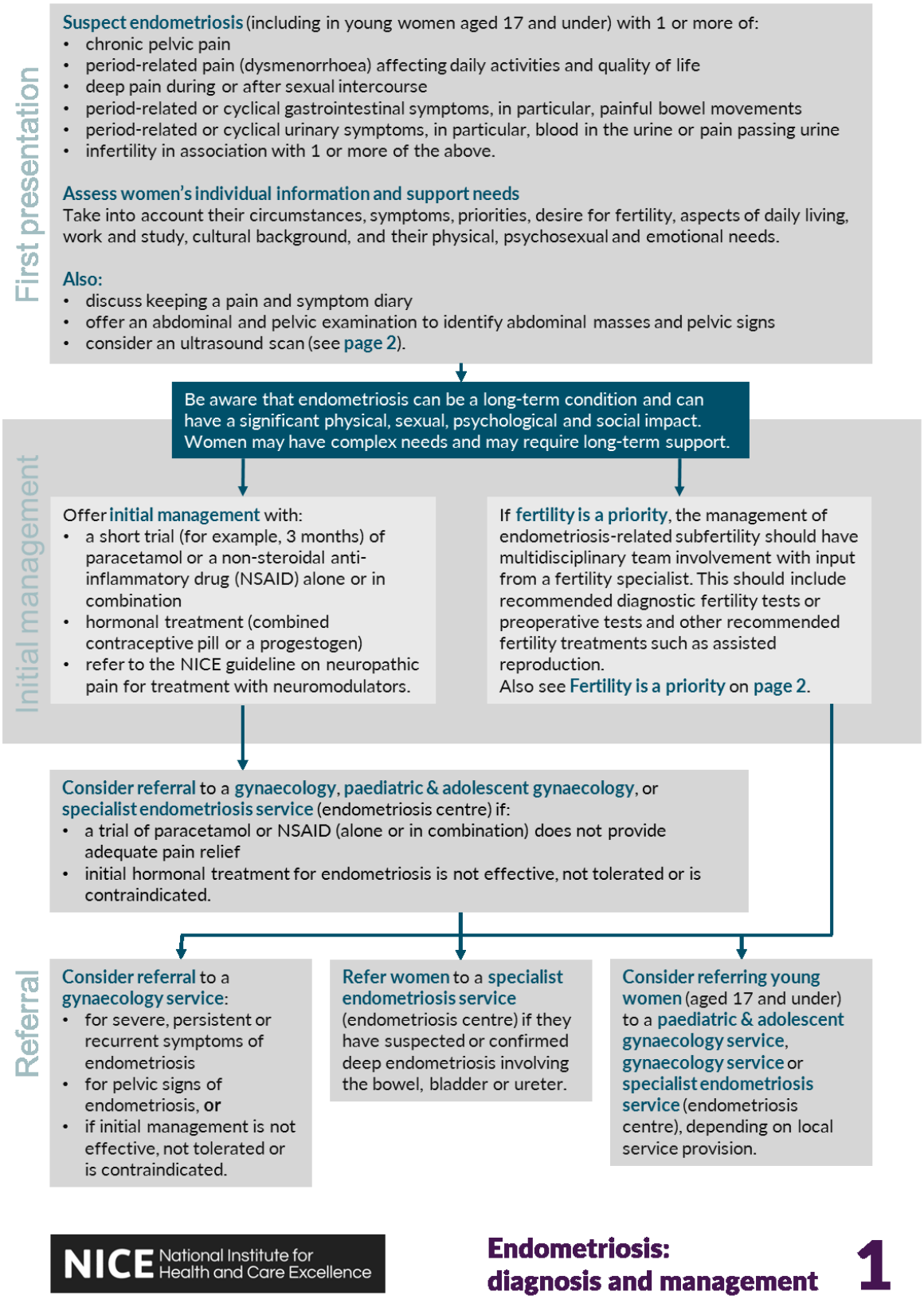
Acknowledgements

Additional support was received from Karen Head, Taryn Krause, Robin Pridy, Tim Reeves, Pauline Turner and Rachel Wheeler.

ISBN 978-1-4731-2661-9

2.2 Algorithm

Figure 1: Algorithm



ISBN 978-1-4731-2661-9

ISBN 978-1-4731-2661-9

2.3 Other versions of the guideline

The National Institute for Health and Care Excellence (NICE) produce a number of versions of this guideline:

- the 'short guideline' lists the recommendations, context and recommendations; and
- NICE Pathways brings together all connected NICE guidance.

2.4 Schedule for updating the guideline

For the most up-to-date information about guideline reviews, please see the latest version of the NICE guidelines manual available from the NICE website.

ISBN 978-1-4731-2661-9

3 Development of the guideline

3.1 What is a NICE clinical guideline?

NICE clinical guidelines are recommendations for the care of individuals in specific clinical conditions or circumstances within the NHS – from prevention and self-care through primary and secondary care to more specialised services. We base our clinical guidelines on the best available research evidence, with the aim of improving the quality of healthcare. We use predetermined and systematic methods to identify and evaluate the evidence relating to specific review questions.

NICE clinical guidelines can:

- provide recommendations for the treatment and care of people by healthcare professionals
- be used to develop standards to assess the clinical practice of individual healthcare professionals
- be used in the education and training of healthcare professionals
- help patients to make informed decisions
- improve communication between patients and healthcare professionals.

While guidelines assist the practice of healthcare professionals, they do not replace their knowledge and skills.

We produce our guidelines using the following steps:

- The guideline topic is chosen in consultation with NHS England, the Department of Health and Public Health England.
- Stakeholders register an interest in the guideline and are consulted throughout the development process.
- The scope is prepared by the NGA.
- The NGA establishes a Committee.
- A draft guideline is produced after the group assesses the available evidence and makes recommendations.
- There is a consultation on the draft guideline.
- The final guideline is produced.

The NGA and NICE produce a number of versions of this guideline:

- The ‘full guideline’ contains all the recommendations, together with details of the methods used and the underpinning evidence.
- The ‘short version’ lists the recommendations, context and recommendations for research.
- NICE Pathways brings together all connected NICE guidance.

3.2 Remit

NICE received the remit for this guideline from the Department of Health. It commissioned the NGA to produce the guideline.

The remit for this guideline is to develop a clinical guideline on the diagnosis and management of endometriosis.

The scope for this guideline is provided in Appendix A. Stakeholders were consulted on a draft of the scope (for a list of stakeholders see Appendix B).

ISBN 978-1-4731-2661-9

3.3 Who developed this guideline?

A multidisciplinary Committee comprising healthcare professionals and researchers as well as lay members developed this guideline (see the list of group members and acknowledgements).

NICE funds the NGA and thus supported the development of this guideline. The Committee was convened by the NGA and chaired by Dr Caroline Overton in accordance with guidance from NICE.

The group met every 4 to 6 weeks during the development of the guideline. At the start of the guideline development process all group members declared interests including consultancies, fee-paid work, shareholdings, fellowships and support from the healthcare industry. At all subsequent group meetings, members declared arising conflicts of interest.

Members were either required to withdraw completely or for part of the discussion if their declared interest necessitated it appropriate to do so. The details of declared interests and the actions taken are shown in Appendix C.

Staff from the NGA provided methodological support and guidance for the development process. The team working on the guideline included a guideline lead, a project manager, systematic reviewers, health economists, a statistician and information scientists. They undertook systematic searches of the literature, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate and drafted the guideline in collaboration with the group.

3.4 What this guideline covers

3.4.1 Groups that will be covered

This guideline covers the following groups:

- Women with confirmed or suspected endometriosis
- Women with recurrent symptoms of endometriosis
- Women with asymptomatic endometriosis discovered incidentally.

Young women (aged 17 and under) have been identified as a subgroup needing specific consideration.

3.4.2 Key clinical issues that will be covered

The following clinical issues that will be covered in this guideline:

- Symptoms and signs of endometriosis
- How and when to monitor and refer for complications and disease progression
- Use of diagnostic tests including imaging, biomarkers and surgical diagnosis
- Use of staging systems to guide treatment decisions
- Timing of interventions
- Pharmacological and surgical treatments including analgesics, hormonal medical treatments, neuromodulators, ablation, excision and hysterectomy with or without oophorectomy.
- Combining pharmacological and surgical treatments.
- Non-medical management specific to pain (for example acupuncture)
- Use of specialist services to deliver care
- Information and support for women with endometriosis.

ISBN 978-1-4731-2661-9

Note that guideline recommendations will normally fall within licensed indications. Exceptionally, and only if clearly supported by evidence, use outside a licensed indication may be recommended. This guideline will assume that prescribers will use a drug's summary of product characteristics to inform decisions made with individual patients.

3.5 What this guideline does not cover

3.5.1 Groups that will not be covered

This guideline does not cover:

- Women with endometriosis occurring outside the pelvis
- Postmenopausal women.

3.5.2 Clinical issues that will not be covered

This guideline does not cover:

- Investigation and assisted reproductive management of fertility problems related to endometriosis.
- Care during pregnancy for women with endometriosis
- Management of menopausal symptoms related to surgical treatment of endometriosis
- Treatment specific to adenomyosis in isolation.

3.6 Relationship between the guideline and other NICE guidance

3.6.1 Related NICE guidance

Menopause (2015) NICE guideline NG23.

3.6.1.1 NICE guidance that will be updated by this guideline

Fertility (2013) NICE guideline CG156. Recommendations 1.7.1.1–1.7.2.4

ISBN 978-1-4731-2661-9

4 Guideline development methodology

This chapter sets out in detail the methods used to review the evidence and to generate the recommendations that are presented in subsequent chapters. This guidance was developed in accordance with the methods outlined in the NICE guidelines manual 2014.

4.1 Developing the review questions and protocols

The 21 review questions developed for this guideline were based on the key areas identified in the guideline scope. They were drafted by the NGA and refined and validated by the Committee.

The review questions were based on the following frameworks:

- intervention reviews – using population, intervention, comparison and outcome (a PICO framework)
- reviews of diagnostic test accuracy – using population, diagnostic test (index tests), reference standard and target condition
- qualitative reviews – using population, area of interest and themes of interest
- prognostic reviews – using population, presence or absence of a risk factor, and outcome. This risk factor could be endometriosis itself as in the risk for cancer review (see chapter 7)

Full literature searches, critical appraisals and evidence reviews were completed for all review questions.

4.2 Searching for evidence

4.2.1 Clinical literature search

Systematic literature searches were undertaken to identify all published clinical evidence relevant to the review questions.

Databases were searched using relevant medical subject headings, free-text terms and study type filters where appropriate. Studies published in languages other than English were not reviewed. Where possible, searches were restricted to retrieve only articles published in English. All searches were conducted in MEDLINE, Embase and The Cochrane Library. All searches were updated in December 2016. Any studies added to the databases after this date (even those published prior to this date) were not included unless specifically stated in the text.

Search strategies were quality assured by cross-checking reference lists of highly relevant papers, analysing search strategies in other systematic reviews and asking the group members to highlight any additional studies. The questions, the study types applied, the databases searched and the years covered can be found in Appendix E.

The titles and abstracts of records retrieved by the searches were inspected for relevance, with potentially significant publications obtained in full text. These were assessed against the inclusion criteria.

During the scoping stage, a search was conducted for guidelines and reports on websites of organisations relevant to the topic. Searching for grey literature or unpublished literature was not undertaken. Searches for electronic, ahead-of-print publications were not routinely undertaken unless indicated by the Committee. All references suggested by stakeholders at the scoping consultation were initially considered.

ISBN 978-1-4731-2661-9

In terms of diagnostic test accuracy reviews (see chapter 8), 1 systematic literature search was carried out for all index tests listed in the review protocol. The resulting titles and abstracts were then sifted for all index tests generating:

- included studies for each index test; and
- a single excluded studies list for all studies that were not included in any of the diagnostic reviews.

4.3 Reviewing research evidence

4.3.1 Types of studies and inclusion and exclusion criteria

For most intervention reviews in this guideline, parallel randomised controlled trials (RCTs) were prioritised because they are considered the most robust type of study design that could produce an unbiased estimate of the intervention effects.

For diagnostic reviews, cross-sectional, retrospective or prospective observational studies were considered for inclusion. For prognostic reviews, prospective and retrospective cohort studies were included. Case-control studies were not considered for inclusion.

In the qualitative review, studies using focus groups, or structured or semi-structured interviews were considered for inclusion. Survey data or other types of questionnaires were only included if they provided analysis from open-ended questions, but not if they reported descriptive quantitative data only.

Where data from observational studies were included, the Committee decided that the results for each outcome should be presented separately for each study and meta-analysis was not conducted.

The evidence was reviewed following the steps shown schematically in Figure 2:

- Potentially relevant studies were identified for each review question from the relevant search results by reviewing titles and abstracts. Full papers were then obtained
- Full papers were reviewed against pre-specified inclusion and exclusion criteria to identify studies that addressed the review question in the appropriate population, as outlined in the review protocols (review protocols are included in Appendix D)
- Relevant studies were critically appraised using the appropriate checklist as specified in the NICE guidelines manual
- Key information was extracted on the study's methods, according to the factors specified in the protocols and results. These were presented in summary tables (in each review chapter) and evidence tables (in Appendix G)
- Summaries of evidence were generated by outcome (included in the relevant review chapters) and were presented in committee meetings (details of how the evidence was appraised is described in Section 4.5 below):
    - Randomised studies: meta-analysis was carried out where appropriate and results were reported in GRADE profiles (for intervention reviews)
    - Observational studies: data were presented as a range of values in GRADE profiles
    - Prognostic studies: data were presented as a range of values, usually in terms of the relative effect as reported by the authors
    - Diagnostic studies: data were presented as measures of diagnostic test accuracy (sensitivity and specificity) and were presented in modified GRADE profiles.

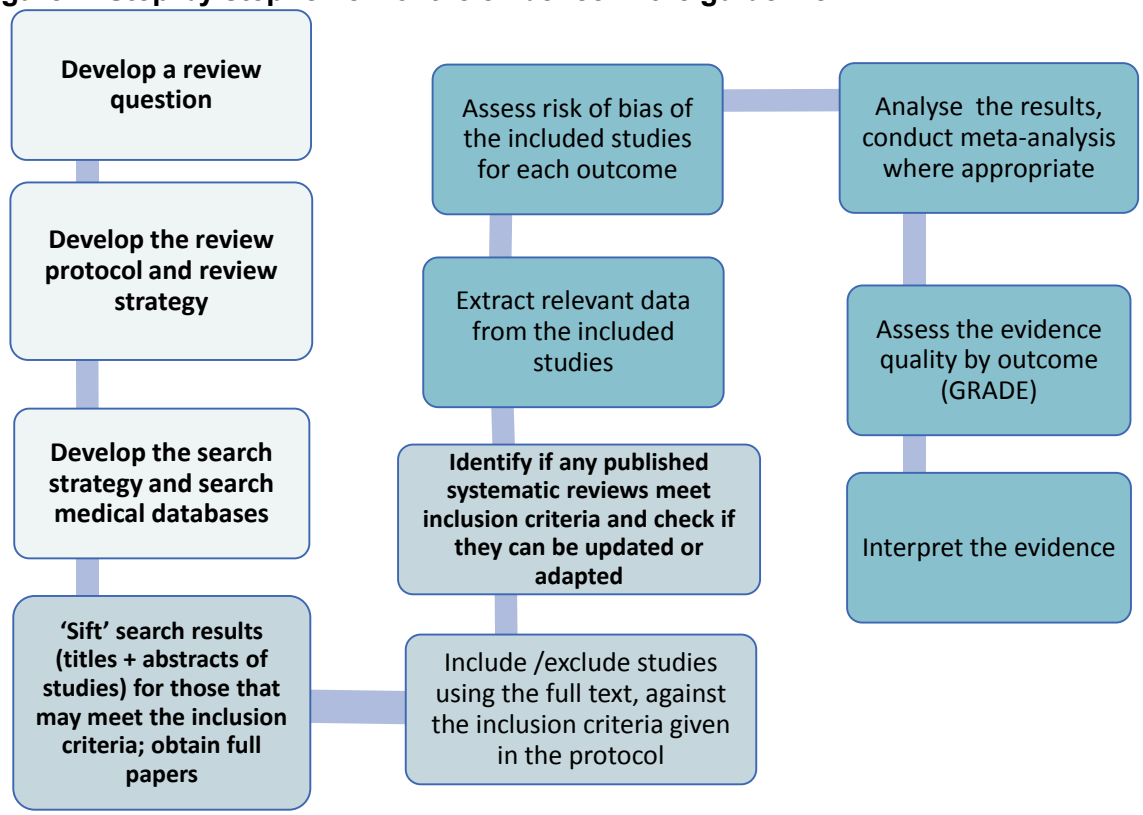
Qualitative studies: each study was summarised by theme and meta-synthesis was carried out where appropriate to identify an overarching framework of themes and subthemes. These were then presented in modified GRADE-CERQual (Lewin 2015) profile, where CERQual stands for Confidence in the Evidence from Reviews of Qualitative research.

ISBN 978-1-4731-2661-9

For quality assurance of study identification, either whole study selections or a sample of the study selection results were double checked by a second reviewer. This was carried out for 20% of all searches related to the Network Meta-Analysis and were double sifted.

A sample of all evidence tables was double extracted (20% of the Network Meta-Analysis). All drafts of reviews were checked by a second reviewer. Any discrepancies were resolved by discussion between the 2 reviewers.

Figure 2: Step-by-step review of the evidence in the guideline



4.3.1.1 Specific inclusions and exclusions

In chapter 11, where the impact of surgical or hormonal treatments on fertility are reviewed, the population was restricted to women with endometriosis who had been unsuccessfully trying to conceive and who did not have assisted reproductive treatment. The outcome that was then considered in the Network Meta-Analysis (for a description of the methods see section 4.4.1.1 and chapter 12) was spontaneous pregnancy (i.e. pregnancy that was not assisted by reproductive treatment).

Young women (aged 17 and under) are a specific subgroup highlighted in the scope. Endometriosis is particularly under recognised in the group of women. We therefore looked for evidence specific to this age group in each of our review question and reported this if the evidence was specifically reported in this way.

Adverse events were initially loosely, if at all, specified in the review protocols for hormonal treatments. After further discussion with the Committee it was agreed that 'withdrawal due to adverse events' would be the only outcome related to adverse events that should be extracted. There were several reasons for this:

ISBN 978-1-4731-2661-9

- Many of the adverse events for different classes of hormonal treatments are commonly known and recognised
- The Committee wanted to know whether the possible benefit from the treatment outweighed the adverse events, which could only be shown by whether or not women were more likely to persist taking one type of hormone over another.
- It makes the different hormonal treatments (with often very idiosyncratic adverse events) comparable.

These outcomes were therefore used in the Network Meta-Analysis of hormonal treatments (please see Chapter 11).

4.4 Method of combining clinical studies

When planning reviews (protocols), the following approaches for data synthesis were discussed and agreed with Committee.

4.4.1 Data synthesis for intervention reviews

It was planned to conduct meta-analyses where possible, to combine the results of studies for each review question using Cochrane Review Manager (RevMan5) software.

Fixed-effect (Mantel–Haenszel) techniques were used to calculate risk ratios (relative risk) for binary outcomes, such as rate of adverse events or rate of people with symptom improvements (Mantel–Haenszel 1959).

For continuous outcomes, measures of central tendency (mean) and variation (standard deviation) are required for meta-analysis. Data for continuous outcomes (such as level of pain on a visual analogue scale [VAS]) were analysed using an inverse variance method for pooling weighted mean differences. A generic inverse variance option in RevMan5 was used where any studies reported solely the summary statistics and 95% confidence interval (95% CI) or standard error; this included any hazard ratios reported. However, in cases where standard deviations were not reported per intervention group, the standard error (SE) for the mean difference is calculated from other reported statistics (p values or 95% CIs): meta-analysis was then undertaken for the mean difference and SE using the generic inverse variance method in RevMan5. When the only evidence was based on studies summarising results by presenting medians (and interquartile ranges) or only p values were given, this information was assessed in terms of the study's sample size and was included in the GRADE tables without calculating the relative or absolute effects. Consequently, aspects of quality assessment, such as imprecision of effect, could not be assessed for evidence of this type. However, the limited reporting of this outcome was classified as a risk of bias in study limitations.

Stratified analyses were predefined for some review questions at the protocol stage when the Committee identified that these strata are different in terms of biological and clinical characteristics and the interventions were expected to have a different effect.

Statistical heterogeneity was assessed by visually examining the forest plots (please see Appendix I) and by considering the chi-squared test for significance at $p<0.1$ or an I-squared inconsistency statistic (with an I-squared value of more than 50% indicating considerable heterogeneity). Where considerable heterogeneity was present, predefined subgroup analyses were performed.

Assessments of potential differences in effect between subgroups were based on the chi-squared tests for heterogeneity statistics between subgroups. If no sensitivity analysis was found to completely resolve statistical heterogeneity, then a random-effects (DerSimonian and Laird) model was employed to provide a more conservative estimate of the effect – (DerSimonian and Laird, 1986).

ISBN 978-1-4731-2661-9

4.4.1.1 Data synthesis for intervention reviews using Network Meta-Analysis (NMA)

As it is the case for ordinary pairwise meta-analysis, NMA may be conducted using either fixed or random effects models. A fixed effects model typically assumes that there is no variation in relative effects across trials for a particular pairwise comparison and any observed differences are solely due to chance. For a random effects model, it is assumed that the relative effects are different in each trial but that they are from a single common distribution. The variance reflecting heterogeneity is often assumed to be constant across trials.

For pain relief, a multivariate NMA was performed using the method of Achana (2014). This allows for results to be reported on a single scale (the VAS) that can easily be incorporated into a cost-effectiveness analysis. It also estimates treatment effects on all scales, even if they may only be reported on one scale in the original included study. The multivariate NMA used known correlations (Gerlinger 2012) between VAS, dysmenorrhoea (Biberoglu and Behrman 1981) and non-menstrual pelvic pain (Biberoglu and Behrman) to fully inform the network of treatments, with final results reported on the VAS. Dyspareunia was not included in the multivariate NMA as only 2 of 5 included studies reported data to calculate standard errors (SE) and therefore this outcome added very little information to the network.

For continuous outcomes, where SEs could not be calculated from the data, we imputed them from other studies that reported measures of uncertainty/variance, using the method of Stevens (2011). Though we did not directly allow for uncertainty in their imputation, we performed sensitivity analyses on the imputation by using the upper 95% credible interval (95% CrI) of the posterior of the imputed SEs.

For the VAS, any results reported on a scale ranging from 0-10 were converted to 0-100. Where medians and ranges or interquartile ranges were reported, we assumed the scale to be approximately normally distributed and converted them to means and SEs (Wan 2015).

In a Bayesian analysis, for each parameter the evidence distribution is weighted by a distribution of prior beliefs. Markov Chain Monte Carlo (MCMC) algorithm was used to generate a sequence of samples from a joint posterior distribution of 2 or more random variables and is particularly well adapted to sampling the treatment effects (known as posterior distribution) of a Bayesian network. A non-informative prior distribution was used to maximise the weighting given to the data and to generate the posterior distribution for each log odds ratio (OR) or mean difference (MD) of interest in the networks. We used the median of the distribution as our point estimate and the centiles provided the 95% Credible Intervals (CrI).

Non-informative priors were selected for discontinuation and VAS networks which were normally distributed with a mean of 0 and standard deviation of 100. However, for discontinuation, as there was sparse data on a number of treatments, we investigated whether the use of informative priors generated from empirical data would give a more stable between-study variance (Turner 2012; Appendix L). For the networks of Biberoglu and Behrman pain scales (dysmenorrhoea, dyspareunia and non-menstrual pelvic pain) we used truncated prior distributions that ensured estimates were kept between the 0-3 limits of the scale.

For the analyses, a series of 40,000 (100,000 for the multivariate NMA) burn-in simulations were run to allow the posterior distributions to convergence and then a further 100,000 simulations were run to produce the outputs. Convergence was assessed by examining the history, autocorrelation and Brooks-Gelman-Rubin plots.

Goodness-of-fit of the model was also estimated by using the posterior mean of the sum of the deviance contributions for each item by calculating the residual deviance and deviance information criteria (DIC). If the residual deviance was close to the number of unconstrained data points (the number of trial arms in the analysis) then the model was explaining the data

ISBN 978-1-4731-2661-9

at a satisfactory level. The choice of a fixed or random effects model can be made by comparing their goodness-of-fit to the data.

Incoherence in NMA between direct and indirect evidence can be assessed in closed treatment loops within the network. These closed treatment loops are regions within a network where direct evidence is available on at least 3 different treatments that form a closed "circuit" of treatment comparisons (for example, A vs B, B vs C, C vs A). If closed treatment loops existed then discrepancies between direct and indirect evidence was assessed for each loop using node-splitting (van Valkenhoef 2016).The outputs of the NMA were:

- Treatment specific log ORs and MDs with their 95% Credible Interval (CrI) were generated for every possible pairs of comparisons by combining direct and indirect evidence in each network
- The probability that each treatment is ranked within the best 3 or worst 3 treatments, based on the proportion of Markov chain iterations in which the treatment effect for an intervention is ranked best, 2nd best and so forth. This was calculated by taking the treatment effect of each drug compared to placebo and counting the proportion of simulations of the Markov chain in which each intervention had the highest treatment effect
- The ranking of treatments compared to placebo (presented as median rank and its 95% CrI).

One of the main advantages of the Bayesian approach is that the method leads to a decision framework that supports decision making. The Bayesian approach also allows the probability that each intervention is best for achieving a particular outcome, as well as its ranking, to be calculated.

We adapted a random effects model template for continuous and dichotomous data available from NICE Technical Support UNIT (TSU) technical support document number 2:

http://www.nicedsu.org.uk/Evidence-Synthesis-TSD-series (2391675).htm. This model accounts for the within-study correlation between treatment effects induced by multi-arm trials.

For further description of methods and the specific results of the NMA please see chapter 10.

4.4.2 Data synthesis for diagnostic test accuracy review

4.4.2.1 Data and outcomes

There are a number of diagnostic test accuracy measures. Sensitivity, specificity and the area under the curve were used as outcomes for diagnostic reviews in this guideline.

Sensitivity and specificity are measures of the ability of a test to correctly classify a person as having a disorder or not having a disorder. When Sensitivity is high, a Negative test result rules out the target disorder. When Specificity is high, a Positive test result rules in the target disorder – researchers have created the mnemonic SpPin/SnNout for this (Sackett 1992). An ideal test would be both highly sensitive and highly specific, but this is frequently not possible and typically there is a trade-off.

The area under the curve (AUC) of receiver operating characteristics (ROC) shows true positive rate (sensitivity) as a function of false positive rate (1 minus specificity).

ISBN 978-1-4731-2661-9

Data synthesis

Diagnostic paired sensitivity-specificity forest plots were produced for each diagnostic test using RevMan5. In order to do this, 2×2 tables (the number of true positives, false positives, true negatives and false negatives) were extracted.

If area under the ROC curve (AUC) data for continuous test results were given as AUC values with 95% confidence intervals, the Committee agreed on the following criteria:

- <0.50: the index test is worse than chance
- 0.50–0.60: very poor
- 0.61–0.70: poor
- 0.71–0.80: moderate
- 0.81–0.92: good
- 0.91–1.00: excellent or perfect test.

4.4.2.2 Diagnostic meta-analysis

When data from 3 or more studies were available, a diagnostic meta-analysis was carried out. To show the differences between study results, pairs of sensitivity and specificity were plotted for each study on one receiver operating characteristics (ROC) curve in RevMan5 (for plots please see Appendix I. Study results were pooled using the bivariate method for the direct estimation of summary sensitivity and specificity using a random effects approach (in WinBUGS® software). Using the output from WinBUGS®, we constructed and plotted confidence regions and, where appropriate ROC curves, using methods outlined by Novelli 2010. As it is a Bayesian analysis, the evidence distribution is weighted by a distribution of prior beliefs. Vague non-informative priors were used for all parameters. For each analysis, a series of 50,000 burn-in simulations were run to allow convergence and then a further 50,000 simulations were run to produce the outputs. Convergence was assessed by investigating density plots, auto correlation plots and history plots for parameters of interest. In cases where many cell counts were 0, 1 was added to each category (true positives, false positives, true negatives, false negatives) to ensure the model was able to run, while not significantly distorting the results.

The advantage of this approach is that it produces summary estimates of sensitivity and specificity that account for the correlation between the 2 measures (sensitivity and specificity). Other advantages of this method have been described elsewhere (Reitsma, 2005; Van Houwelingen, 1993; Van Houwelingen, 2002).

This model also assesses the variability by incorporating the precision by which sensitivity and specificity have been measured in each study. A confidence ellipse is shown in the graph that indicates the confidence region around the summary sensitivity / specificity point. A summary ROC curve is also presented. From the WinBUGS® output we report the summary estimate of sensitivity and specificity (plus their 95% confidence intervals) as well as between study variation measured as logit sensitivity and specificity as well as correlations between the 2 measures of variation.

4.4.3 Data synthesis for qualitative review

Where possible, a meta-synthesis was conducted to combine qualitative study results. The main aim of the synthesis of qualitative data was to produce a description of the topics that may influence the experience of a woman who has endometriosis, those people important to them and healthcare professionals involved in their care, rather than build new theories or reconceptualise the topic under review. Whenever studies identified a qualitative theme, this was extracted and the main characteristics were summarised. When all themes were extracted from studies, common concepts were categorised and tabulated. This included

ISBN 978-1-4731-2661-9

information on how many studies had contributed to an identified overarching theme. In qualitative synthesis, a theme being reported by different studies more often than other themes does not necessarily mean that it would be more important than those other themes. The aim of qualitative research is to identify new perspectives on a particular topic. Study type and population in qualitative research can differ widely, meaning that themes identified by just one or a few studies can provide important new information for a given topic. Therefore, for the purpose of the qualitative reviews in this guideline, we did not add further studies when they reported the same themes that had already been identified from the same perspectives (that is from women, their partners or families, or healthcare professionals) because the emphasis was on conceptual robustness rather than the quantitative completeness of evidence. This has implications for the types and numbers of studies that are included in the qualitative reviews. Study inclusion continued until no new relevant data could be found regarding a topic that would add to or refute it, a concept referred to in the literature as 'theoretical saturation' (Dixon-Woods 2005).

The most relevant evidence in this respect would originate from studies set in the target context of the UK NHS setting. Themes from individual studies were then integrated into a wider context and, when possible, overarching categories of themes with sub-themes were identified. Themes were derived from data presented in individual studies based directly on quotes from interviewees. When themes were extracted, theme names derived from the studies that provided it were used. The names of overarching themes, however, were named by the systematic reviewers (see section 7.4).

Emerging themes were then placed into a thematic map that presents the relationship between themes and subthemes. The purpose of the map was to show relationships between overarching themes and their subthemes. The mapping part of the review was drafted by a member of the technical team, but the final framework of themes was further shaped and, when necessary, re-classified through discussion with at least one other member of the technical team. The Committee could then draw conclusions from each theme in each setting or country and how they may help in forming recommendations.

4.4.4 Data synthesis for prognostic reviews

Signs and symptoms indicative of endometriosis (e.g. pain) could be construed as a characteristic that predicts a diagnosis of endometriosis. This would be classified as a prognostic/predictive factor. In this respect, odds ratios (ORs), risk ratios (RRs) or hazard ratios (HRs), with their 95% confidence intervals (95% CIs) for the effect of the pre-specified prognostic factors, were extracted from the papers when reported. Evidence came from observational studies because signs and symptoms that may indicate endometriosis are not factors that could be randomised. For this topic, we looked for studies that took into account possible key confounders as reported in multivariable analyses. The reported measures were therefore adjusted to take into account other characteristics less likely to be actual signs and symptoms of endometriosis. Studies did this in a pre-specified manner or used statistical methods that included variables that were likely to be signs and symptoms related to endometriosis and modelled them using statistical methods (such as multivariable logistic regressions). This would then indicate which characteristics are most likely to be independent prognostic factors rather than a factor only spuriously related to a diagnosis of endometriosis.

4.5 Appraising the quality of evidence

For intervention reviews, the evidence for outcomes from the included RCTs and observational studies were evaluated and presented using GRADE, which was developed by the international GRADE working group. Modified GRADE assessments were also carried out for accuracy measures in diagnostic reviews. For the appraisal of the quality of the evidence from qualitative reviews an adapted GRADE-CERQual (Lewin 2015) approach was

ISBN 978-1-4731-2661-9

used, where CERQual stands for Confidence in the Evidence from Reviews of Qualitative research.

The software developed by the GRADE working group (GRADEpro) was used to assess the quality of each outcome, taking into account individual study quality factors and the meta-analysis results. The clinical/economic evidence profile tables include details of the quality assessment and pooled outcome data, where appropriate, an absolute measure of intervention effect and the summary of quality of evidence for that outcome. In this table, the columns for intervention and control indicate summary measures of effect and measures of dispersion (such as mean and standard deviation or median and range) for continuous outcomes and frequency of events (n/N: the sum across studies of the number of patients with events divided by sum of the number of completers) for binary outcomes. Reporting or publication bias was only taken into consideration in the quality assessment and included in the clinical evidence profile tables if it was apparent.

The selection of outcomes for each review question was decided when each review protocol was discussed with the Committee. However, given the nature of most of the review questions included in this guideline (driven by short- or long-term outcomes), the categorisation of outcomes as critical and important did not follow the standard GRADE approach. The outcomes selected for a review question were critical for decision-making in a specific context.

The evidence for each outcome in interventional reviews was examined separately for the quality elements listed and defined in Table 3. Each element was graded using the quality levels listed in Table 4.

The main criteria considered in the rating of these elements are discussed below. Footnotes were used to describe reasons for grading a quality element as having serious or very serious limitations. The ratings for each component were summed to obtain an overall assessment for each outcome (Table 5).

The GRADE toolbox is designed only for RCTs and observational studies, but we adapted the quality assessment elements and outcome presentation for diagnostic accuracy and qualitative studies, subject to data availability. For example, for diagnostic accuracy studies, the GRADE tables were modified to include the most appropriate measures of diagnostic accuracy (sensitivity and specificity) whereas qualitative studies were presented in summary evidence tables around themes identified or direct participants' quotations. Quality of the evidence in the qualitative reviews was assessed per study level.

Table 3: Description of quality elements in GRADE (see details in sections 4.5.1.1 to 4.5.1.4)

Quality element	Description
Risk of bias (study limitations)	Limitations in the study design and implementation may bias the estimates of the treatment effect. High risk of bias for the majority of the evidence decreases confidence in the estimate of the effect.
Inconsistency	Inconsistency refers to an unexplained heterogeneity of results or findings.
Indirectness	Indirectness refers to differences in study population, intervention, comparator and outcomes between the available evidence and the review question, or recommendation made, such that the effect estimate is changed. This is also related to applicability or generalisability of findings.

ISBN 978-1-4731-2661-9

Quality element	Description
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect. Imprecision results if the confidence interval includes the clinically important threshold. For qualitative research this can relate to the sufficiency of data within each theme.
Publication bias	Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies.

Table 4: Levels of quality elements in GRADE level

Levels of quality elements in GRADE level	Description
None	There are no serious issues with the evidence.
Serious	The issues are serious enough to downgrade the outcome evidence by 1 level.
Very serious	The issues are serious enough to downgrade the outcome evidence by 2 levels.

Table 5: Overall quality of outcome evidence in GRADE level

Overall quality of outcome evidence in GRADE level	Description
High	Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low	Any estimate of effect is very uncertain.

4.5.1 Grading the quality of clinical evidence

After results were pooled, the overall quality of evidence for each outcome was considered. The following procedure was adopted when using the GRADE approach:

- A quality rating was assigned based on the study design. RCTs start as high, observational studies as moderate and uncontrolled case series as low or very low
- The rating was then downgraded for the specified criteria: risk of bias (study limitations); inconsistency; indirectness; imprecision; and publication bias. These criteria are detailed below. Evidence from observational studies (which had not previously been downgraded) was upgraded if there was a large magnitude of effect or a dose-response gradient, and if all plausible confounding would reduce a demonstrated effect, or suggest a spurious effect when results showed no effect.
- Each quality element considered to have 'serious' or 'very serious' issues was rated down by 1 or 2 points respectively. Value based judgements for relevant interpretation of the levels of quality elements were informed by discussion with the Committee for each review to balance consistency of approach across the guideline and clinical relevance within each review.

ISBN 978-1-4731-2661-9

- The downgraded/upgraded ratings were then summed and the overall quality rating was revised, taking into account the relative contributions from the individual studies within a meta-analyses, where performed. For example, RCTs start as high and the overall quality becomes moderate, low or very low if 1, 2 or 3 points are deducted respectively
- The reasons or criteria used for downgrading were specified in the footnotes.

The details of the criteria used for each of the main quality elements are discussed further in sections 4.5.1.1 to 4.5.1.4 below.

GRADE quality assessment was not performed for the reviews in Chapter 6 and 8 regarding monitoring and referral nor for the network meta-analysis. Quality statements were informed by assessment of risk of bias.

4.5.1.1 Risk of bias

Intervention studies

Bias can be defined as anything that causes a consistent deviation from the truth. Bias can be perceived as a systematic error.

The risk of bias for a given study and outcome is associated with the risk of over or underestimation of the true effect.

Sources of bias in randomised controlled trials are listed in Table 6.

A study with a poor methodological design does not automatically imply high risk of bias; the bias is considered individually for each outcome and it is assessed whether this poor design will impact on the estimation of the intervention effect.

Table 6: Sources of bias in randomised controlled trials

Risk of bias	Explanation
Allocation concealment	Those enrolling patients are aware of the group to which the next enrolled patient will be allocated (this is a major problem in 'pseudo' or 'quasi' randomised trials with allocation by, for example, day of week, birth date, chart number).
Lack of blinding	Patient, caregivers, those recording outcomes, those adjudicating outcomes or data analysts are aware of the arm to which patients are allocated.
Incomplete accounting of patients and outcome events	Missing data not accounted for and failure of the trialists to adhere to the intention to treat principle when indicated.
Selective outcome reporting	Reporting of some outcomes and not others on the basis of the results.
Other risks of bias	For example: • stopping early for benefit observed in randomised trials, in particular in the absence of adequate stopping rules • use of unvalidated patient-reported outcomes • recruitment bias in cluster randomised trials.

Diagnostic studies

For diagnostic accuracy studies, the Quality Assessment of Diagnostic Accuracy Studies version 2 (QUADAS- 2) checklist was used (http://www.bristol.ac.uk/social-community-

ISBN 978-1-4731-2661-9

medicine/projects/quadas/quadas-2/). Risk of bias and applicability in primary diagnostic accuracy studies in QUADAS- 2 consists of 4 domains:

- patient selection
- index test
- reference standard
- flow and timing.

Qualitative studies

For qualitative studies, quality was assessed using a checklist for qualitative studies (as suggested in Appendix H in the NICE guidelines manual 2014). This was based on the Critical Appraisal Skills Programme (CASP) checklist for qualitative studies. The quality rating for risk of bias (low, high and unclear) was derived by assessing the risk of bias across 6 domains. The evidence was then assessed by theme using GRADECerqual across studies as described above and labelled (no limitations, minor limitations, major limitations and unclear), see Table 7.

Table 7: Domains for quality assessment of qualitative studies

Risk of bias	Explanation
Aim and appropriateness of qualitative evidence.	This refers to an assessment of whether the aims and relevance of the study were clearly described and whether qualitative research methods were appropriate for investigating the research question.
Rigour in study design or validity of theoretical approach	This domain assesses whether the study approach has been clearly described and is based on a theoretical framework (for example, ethnography or grounded theory). This does not necessarily mean that the framework has to be explicitly stated, but that at least a detailed description is provided which makes it transparent and reproducible.
Sample selection	The background, the procedure and reasons for the chosen method of selecting participants should be stated. It should also be assessed whether there was a relationship between the researcher and the informant and if so, how this may have influenced the findings that were described.
Data collection	Consideration was given to how well the method of data collection (in-depth interviews, semi-structured interviews, focus groups or observations) was described, whether details were provided and how the data were collected (who conducted the interviews, how long did they last and where did they take place).
Data analysis	For this criterion it is assessed whether sufficient detail is provided about the analytical process and whether it is in accordance with the theoretical approach. For instance, if a thematic analysis was used, it is assessed whether there was a clear description of how the theme was arrived at. Data saturation is also part of this section. This refers to whether a theoretical point of theme saturation was achieved at which point no further citations or observations would provide more insight or suggest a different interpretation of this theme. This could be explicitly stated, or it may be clear from the citations presented that it may have been possible to find more themes.
Results	In relation to this section the reasoning about the results are important, for instance whether a theoretical proposal or framework is provided rather than being restricted to citations / presentation of data.

Prognostic studies

For prognostic studies, quality was assessed using the checklist for prognostic studies (Appendix H in the NICE guidelines manual 2014).

ISBN 978-1-4731-2661-9

This risk of bias for each risk factor across studies was derived by assessing the risk of bias across 6 domains for each study – selection bias, attrition bias, prognostic factor bias, outcome measurement bias, control for confounders and appropriate statistical analysis – with the last 4 domains being assessed for each outcome. A summary table on the quality of prognostic studies is presented at the beginning of each review to summarise the risk of bias across the 6 domains. More details about the quality assessment for prognostic studies are shown in Table 8:

Table 8: Sources of bias for prognostic factor studies

Risk of bias	Explanation
Patient selection	Selection bias would be suspected if the allocation to groups directly leads to differences in baseline characteristics. If only 1 risk factor is considered, risk of bias may be introduced when there was no attempt to achieve roughly comparable groups, and/or there is evidence of biased selection. If 2 or more risk factors are considered, the same may not apply for patient selection issues and then the study would have to have controlled for confounders.
Prognostic factor bias (or sign/symptom)	This refers to any biases that could directly be linked to the validity of the prognostic factor under investigation, such as how the signs or symptoms were assessed or measured.
Attrition bias	This is assessed by whether there are similar numbers of people who were followed up in groups who have or have not got the particular sign or symptom.
Outcome measurement bias	This usually refers to whether or not the outcome has been measured on a validated scale or was otherwise reliably assessed.
Control for confounders / statistical analysis	This domain is an assessment of whether confounders have been adequately accounted for. Confounders would be signs and symptoms that may be related to dying but that are not under direct investigation. For instance, age is related to dying, but we would not assess age in general as a sign or symptom of dying. We therefore wanted to assess whether signs and symptoms were independent predictors, regardless of other non-related factors.

4.5.1.2 Inconsistency / coherence of findings

Inconsistency refers to unexplained heterogeneity of results. When estimates of the treatment effect, prognostic risk factor or diagnostic accuracy measures vary widely across studies (that is, there is heterogeneity or variability in results), this suggests true differences in underlying effects.

Heterogeneity in meta-analyses was examined; if present, sensitivity and subgroup analyses were performed as pre-specified in the protocols (Appendix D).

When heterogeneity existed (chi-squared probability less than 0.1, I-squared inconsistency statistic of greater than 50%, or from visually examining forest plots), but no plausible explanation could be found (for example, duration of intervention or different follow-up periods), the quality of the evidence was downgraded in GRADE by 1 or 2 levels, depending on the extent of inconsistency in the results. When outcomes are derived from a single trial, inconsistency is not an issue for downgrading the quality of evidence. However, 'no inconsistency' is nevertheless used to describe this quality assessment in the GRADE profiles as this is the default option in the GRADEpro software used.

For diagnostic and prognostic evidence, this was assessed visually according to the differences in point estimates and overlap in confidence intervals on the sensitivity / specificity forest plots. In addition to the I-squared and chi-squared values and examination of forest plots, the decision for downgrading was dependent on factors such as whether the uncertainty about the magnitude of benefit (or harm) of the outcome showing heterogeneity would influence the overall judgment about net benefit or harm (across all outcomes).

ISBN 978-1-4731-2661-9

For qualitative research, a similar concept to inconsistency is coherence, which refers to the way findings within themes are described and whether they make sense. This concept was used in the quality assessment across studies for individual themes. This does not mean that contradictory data was downgraded automatically, but that it was highlighted and presented, and that reasoning was provided. As long as the themes, or components of themes, from individual studies fit into a theoretical framework, they do not necessarily have to have the same perspective. It should, however, be possible to explain these by differences in context (for example, the views of healthcare professionals might not be the same as those of family members, but they could contribute to the same overarching theme). Coherence was graded across studies with the following labels: coherent, incoherent or unclear.

4.5.1.3 Indirectness / applicability or relevance of findings

For quantitative reviews, directness refers to the extent to which the populations, intervention/risk factor/index test, comparisons and outcome measures are similar to those defined in the inclusion criteria for the reviews. Indirectness is important when these differences are expected to contribute to a difference in effect size, or may affect the balance of harms and benefits considered for an intervention.

Relevance of findings in qualitative research is the equivalent of indirectness for quantitative outcomes and refers to how closely the aims and context of the studies contributing to a theme reflect the objectives outlined in the review protocol of the guideline question.

4.5.1.4 Imprecision / theme saturation or sufficiency

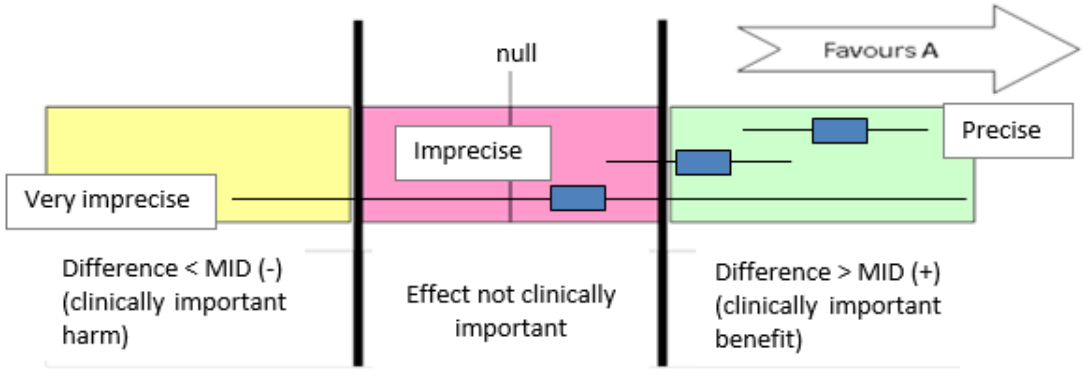
For quantitative reviews, imprecision in guidelines concerns whether the uncertainty (confidence interval) around the effect estimate means that it is not clear whether there is a clinically important difference between interventions or not (that is, whether the evidence would clearly support 1 recommendation or appear to be consistent with several different types of recommendations). Therefore, imprecision differs from the other aspects of evidence quality because it is not really concerned with whether the point estimate is accurate or correct (has internal or external validity); instead, it is concerned with the uncertainty about what the point estimate actually is. This uncertainty is reflected in the width of the confidence interval.

The 95% confidence interval (95% CI) is defined as the range of values that contain the population value with 95% probability. The larger the trial, the smaller the 95% CI and the more certain the effect estimate.

Imprecision in the evidence reviews was assessed by considering whether the width of the 95% CI of the effect estimate was relevant to decision-making, considering each outcome in isolation. This is explained in Figure 3, which considers a positive outcome for the comparison of treatment A versus treatment B. Three decision-making zones can be identified, bounded by the thresholds for clinical importance (minimal important difference, MID) for benefit and for harm. The MID for harm for a positive outcome means the threshold at which drug A is less effective than drug B by an amount that is clinically important to patients (favours B).

ISBN 978-1-4731-2661-9

Figure 3: Illustration of precise, imprecise and very imprecise evidence based on the confidence interval of outcomes in forest plots



When the confidence interval of the effect estimate is wholly contained in 1 of the 3 zones (for example, clinically important benefit), we are not uncertain about the size and direction of effect (whether there is a clinically important benefit, or the effect is not clinically important, or there is a clinically important harm), so there is no imprecision.

When a wide confidence interval lies partly in each of 2 zones, it is uncertain in which zone the true value of effect estimate lies and therefore there is uncertainty over which decision to make (based on this outcome alone). The confidence interval is consistent with 2 possible decisions and so this is considered to be imprecise in the GRADE analysis and the evidence is downgraded by 1 level ('serious imprecision').

If the confidence interval of the effect estimate crosses into 3 zones, this is considered to be very imprecise evidence because the confidence interval is consistent with 3 possible clinical decisions and there is therefore a considerable lack of confidence in the results. The evidence is therefore downgraded by 2 levels in the GRADE analysis ('very serious imprecision').

Implicitly, assessing whether the confidence interval is in, or partially in, a clinically important zone, requires the Committee to estimate an MID or to say whether they would make different decisions for the 2 confidence limits.

The literature was searched for established MIDs for the selected outcomes in the evidence reviews, such as symptom measurement tools. For the pain outcome, as measured on the visual analogue scale, a published MID was used (Gerlinger 2012) which was an interval of 1 cm on a 10 cm scale. In other words any differences larger than 1 cm were classed as clinically significant and then downgraded if the confidence interval crossed this line. For pain measured on other scales or all other outcomes (categorical or continuous) no further published MIDs were identified. In addition, the Committee was asked whether they were aware of any acceptable MIDs in the clinical community. Finally, the Committee considered whether it was clinically acceptable to use the GRADE default MID to assess imprecision: for binary outcomes a 25% relative risk increase and the related relative risk reduction was used, which corresponds to clinically important thresholds for a risk ratio of 0.8 and 1.25 respectively (due to the statistical characteristic of this measure this means that this is not a symmetrical interval). This default MID was used for all the binary outcomes in the interventions' evidence reviews as a starting point and decisions on clinical importance were then considered based on the absolute risk difference. For continuous outcomes default MIDs were also used. These use half of the median standard deviation of the control group.

The same principle was used for prognostic factors, for example, using the default MID as a starting point for the Committee discussion, to assess whether the size of the outcome effect would be large enough to be meaningful in clinical practice.

ISBN 978-1-4731-2661-9

In diagnostic accuracy measures, it was first considered whether sensitivity or specificity (or AUC for continuous variables) would be given more weight in the decision-making process. If one measure was given more importance than the other, then imprecision was rated on this statistical measure. In pooled estimates, the imprecision rating was based on the confidence region of the summary sensitivity and specificity point. A region that was reaching up to the line of chance (the 45 degree line of the ROC plot) was classed as imprecise and a region over the line of chance was classified as very precise.

Theme saturation or sufficiency refers to a similar concept in qualitative research. This refers to whether a theoretical point of theme saturation was achieved, at which point no further citations or observations would provide more insight or suggest a different interpretation of this theme. As already highlighted in a previous section on qualitative reviewing methods, it is not equivalent to the number of studies contributing to a theme, but rather to the depth of data and whether sufficient quotes or observations were provided that could underpin these findings.

4.5.2 Quality assessment of NMA

For the NMAs, quality was assessed by looking at risk of bias across the included evidence (using the standard GRADE approach for this domain), as well as heterogeneity and incoherence.

The following limits of the upper 95% CrI for between-study standard deviation were used to assess heterogeneity for NMAs in which a random effects model was used:

- less than 0.3 – low heterogeneity
- 0.3 to 0.6 – moderate heterogeneity
- 0.6 to 0.9 – high heterogeneity
- 0.9 to 1.2 – very high heterogeneity.

Where significant incoherence was found it was considered to be serious when the direction of effect for both direct and indirect estimates was the same (for example, an odds ratio of greater than 1 in both the direct and indirect estimates), and very serious when the direction of effect was different (for example, an odds ratio of greater than 1 for the direct estimate but less than 1 for the indirect estimate).

For fixed-effect NMAs that did not model heterogeneity, or for networks in which incoherence could not be assessed as no closed treatment loops existed, these criteria were not considered to impact the quality of evidence.

4.5.3 Assessing clinical significance (of intervention effects)

The Committee assessed the evidence by outcome in order to determine if there was, or potentially was, a clinically important benefit, a clinically important harm or no clinically important difference between interventions. To facilitate this, where possible, binary outcomes were converted into absolute risk differences (ARDs) using GRADEpro software: the median control group risk across studies was used to calculate the ARD and its 95% CI from the pooled risk ratio. For continuous outcomes, the mean difference between the intervention and control arm of the trail was calculated. This was then assessed in relation to the default MID (0.5 times the median control group standard deviation).

The assessment of clinical benefit or harm, or no benefit or harm, was not based on the default MID of the relative risk, which was only used as a starting point, but on the point estimate of the absolute effect, taking into consideration the precision around this estimate.

This assessment was carried out by the Committee for each critical outcome and an evidence summary table (used in the Committee meetings, but not presented in this guideline) was produced to compile the Committee's assessments of clinical importance per

ISBN 978-1-4731-2661-9

outcome, alongside the evidence quality and the uncertainty in the effect estimate (imprecision). In instances where the Committee's decision differed from the default assessment, decisions were captured in the 'Linking evidence to recommendations' sections.

4.5.4 Assessing clinical significance (of prognostic, diagnostic or qualitative findings)

Absolute risk differences were not calculated for prognostic findings in this guideline. The Committee considered the size of the relative effects and whether this was large enough to constitute a sign or symptom predicting the occurrence of the selected outcome.

In a similar manner, this was carried out for diagnostic accuracy statistics to interpret how likely the size of the effect reflects a clinically meaningful association between people having a positive test and the target condition.

For themes stemming from qualitative findings, clinical importance was decided upon by the Committee taking into account the generalisability of the context from which the theme was derived and whether it was convincing enough to support or warrant a change in current practice, as well as the evidence quality.

4.5.5 Evidence statements

Evidence statements are summary statements that are presented after the GRADE profiles, summarising the key features of the clinical evidence presented. The wording of the evidence statements reflects the certainty or uncertainty in the estimate of effect. The evidence statements are presented by outcome or theme and encompass the following key features of the evidence:

- the quality of the evidence (GRADE rating)
- the number of studies and the number of participants for a particular outcome
- a brief description of the participants
- an indication of the direction of effect (for example, if a treatment is clinically significant [beneficial or harmful] compared with another, or whether there is no difference between the tested treatments).

4.5.6 Evidence of cost effectiveness

The aims of the health economic input to the guideline were to inform the Committee of potential economic issues related to the diagnosis and management of endometriosis to ensure that recommendations represented a cost-effective use of healthcare resources. Health economic evaluations aim to integrate data on healthcare benefits (ideally in terms of quality-adjusted life-years (QALYs)) with the costs of different care options. In addition, the health economic input aimed to identify areas of high resource impact; recommendations which – while nevertheless cost-effect – might have a large impact on CCG or Trust finances and so need special attention.

The group prioritised a single economic model on interventions where it was thought that economic considerations would be particularly important in formulating recommendations and a review of the health economic literature was undertaken. This model covered multiple review questions, as a complete health economic analysis of the treatment pathway required consideration of all possible combinations of diagnostic strategy and treatment strategy together. For economic evaluations, no standard system of grading the quality of evidence exists and included papers were assessed using the economic evaluations checklist as specified in the NICE guidelines manual.

Health economic reviews were also undertaken for review questions relating to the timing of interventions and the configurations of services. In both of these cases it was thought that

ISBN 978-1-4731-2661-9

the Committee may wish to make recommendations that would lead to a high resource impact, although in practice this did not occur to a substantial degree.

No economic evaluation was undertaken for questions on information and support or signs and symptoms (of endometriosis) as it was agreed with the Committee that these reviews would focus primarily on the content and quality of information which is given to patients and clinicians respectively rather than whether the provision of such information represented a cost-effective use of NHS resources, which was thought to be clinically uncontroversial. Therefore these questions were not primarily about competing alternative uses for NHS resources and therefore were not considered suitable for economic analysis.

No economic analysis was undertake for a question on staging systems. While such an economic model might be valuable in deciding on the allocation of scarce NHS resources, no clinical evidence was uncovered which might populate an economic model which meant that no model could be constructed.

No economic analysis was undertaken for a question on monitoring and referral. This question was of a high health economic importance as the potential quality of life impact for misdiagnosing, for example, ovarian cancer is extremely high. However in order to perform a reasonable economic analysis on this question it would have been necessary to consider the cost-effectiveness of the treatment pathway for each possible reason to refer. Some of these pathways have existing NICE guidance but some do not, which would have required de novo modelling (taking away resources from the main health economic guideline). For this question it was agreed with the Committee that health economic input would be limited to resource impact and analysis, with a full health economic evaluation being left until all possible referral pathways had been costed in other NICE Guidelines.

4.6 Developing recommendations

Over the course of the guideline development process, the Committee was presented with:

- evidence tables of the clinical and economic evidence reviewed from the literature: all evidence tables are in Appendix H
- summary of clinical and economic evidence and quality assessment (as presented in Chapters 4 to 11)
- forest plots (Appendix J)
- a description of the methods and results of the cost-effectiveness analysis undertaken for the guideline (Appendix K).

Recommendations were drafted on the basis of the group's interpretation of the available evidence, taking into account the balance of benefits, harms and costs between different courses of action. This was either done formally, in an economic model, or informally. Firstly, the net benefit over harm (clinical effectiveness) was considered, focusing on the critical outcomes, although most of the reviews in the guideline were outcome driven. When this was done informally, the group took into account the clinical benefits and harms when one intervention was compared with another. The assessment of net benefit was moderated by the importance placed on the outcomes (the group's values and preferences) and the confidence the group had in the evidence (evidence quality). Secondly, the group assessed whether the net benefit justified any differences in costs.

When clinical and economic evidence was of poor quality, conflicting or absent, the group drafted recommendations based on their expert opinion. The considerations for making consensus-based recommendations include the balance between potential harms and benefits, the economic costs or implications compared with the economic benefits, current practices, recommendations made in other relevant guidelines, patient preferences and equality issues. The group also considered whether the uncertainty was sufficient to justify

ISBN 978-1-4731-2661-9

delaying making a recommendation to await further research, taking into account the potential harm of failing to make a clear recommendation.

The wording of recommendations was agreed by the group and focused on the following factors:

- the actions healthcare professionals need to take
- the information readers of the guideline need to know
- the strength of the recommendation (for example, the word 'offer' was used for strong recommendations and 'consider' for weak recommendations)
- the involvement of patients (and their support network if needed) in decisions about treatment and care
- consistency with NICE's standard advice on recommendations about drugs, waiting times and ineffective intervention.

The main considerations specific to each recommendation are outlined in the 'Recommendations and link to evidence' sections within each chapter.

4.6.1 Research recommendations

When areas were identified for which good evidence was lacking, the group considered making recommendations for future research in accordance with the NICE Research Recommendations Process and methods guide (2011), available from the NICE website.
Validation process

This guidance is subject to a 6-week public consultation and feedback as part of the quality assurance and peer review of the document. All comments received from registered stakeholders are responded to in turn and posted on the NICE website at publication.

4.6.2 Updating the guideline

Following publication, and in accordance with the NICE guidelines manual, NICE will undertake a review of whether the evidence base has progressed significantly to alter the guideline recommendations and warrant an update.

4.6.3 Disclaimer

Healthcare providers need to use clinical judgement, knowledge and expertise when deciding whether it is appropriate to apply guidelines. The recommendations cited here are a guide and may not be appropriate for use in all situations. The decision to adopt any of the recommendations cited here must be made by practitioners in light of individual patient circumstances, the wishes of the patient, clinical expertise and resources.

ISBN 978-1-4731-2661-9

5 Organisation of care

5.1 Specialist services

Review question: What is the clinical and cost effectiveness of specialist endometriosis services?

5.1.1 Introduction

Women who suffer from endometriosis of all levels of severity will present with a wide variety of symptoms to clinicians in different settings. The symptoms do not always correlate well with the severity of endometriosis. It is important that women with endometriosis are triaged to receive treatment in the setting that best suits their needs, symptoms and preferences. The expertise and the opportunity for management of these women will differ in each setting, but for women with severe endometriosis that may involve lesions affecting the bowel, urinary tract or other sites beyond her reproductive organs, it is generally thought that a specialist multidisciplinary team would be required.

There is currently variation in the time taken for referral to specialist services and how these services are configured to best meet women's needs. For instance, the skill mix in the multidisciplinary team and the access to pain clinics or diagnostic tests varies across specialist endometriosis services (endometriosis centres). Specialist endometriosis services may not need to comprise all relevant specialists working in the same place as long as there is access to additional expertise or specialist training in the management of endometriosis.

How care for women with endometriosis is best organised to meet their needs is the topic of the current chapter.

For full details, see review protocol in Appendix D, the study selection flow chart in Appendix F and study exclusion list in Appendix H.

5.1.2 Description of clinical evidence

No clinical evidence was identified for this review. For full details of excluded studies, see Appendix H.

5.1.3 Summary of included studies

No clinical evidence was identified for this review.

5.1.4 Clinical evidence profile

No evidence was identified.

5.1.5 Economic evidence

No health economic studies were found contrasting specialist services to conventional gynaecology services. Consequently a de novo model was constructed to support Committee recommendations.

ISBN 978-1-4731-2661-9

5.1.5.1 Economic model description

5.1.5.1.1 Introduction

Specialist endometriosis services (endometriosis centres) are medical units designed with endometriosis patients as their primary users and with healthcare professionals who have expertise and training in the management of endometriosis. Consequently there is good reason to think that patients with endometriosis will receive better care at these units compared to less specialised services. However, because of these units' specialist nature they are likely (on average) to cost more than conventional gynaecology services.

As some women have a complex form of endometriosis that may not be optimally managed in conventional care, there is a belief that some of these women might be more cost-effectively treated in specialist endometriosis services, as it is assumed the higher quality of care will lead to reduced complications, reduced reoperation and a higher quality of life.

As no clinical or economic study was identified considering the expected cost-effectiveness of specialist endometriosis services, a de novo costing model was constructed. As there was no evidence on the clinical effectiveness of gynaecology or specialist endometriosis services, the model was designed with a 'cost-minimisation' approach. This meant that the model was designed to identify what percentage of women could be treated in specialist endometriosis services without exceeding the current budget in the UK. The Committee then used this information to draw conclusions on how these services might best be configured.

5.1.5.1.2 Review of the literature

Rather than studies considering the cost-effectiveness of gynaecology or specialist endometriosis services, 2 studies were identified that could inform a de novo model. These examine the distribution of costs arising from women with endometriosis.

Simoens 2012

Simoens conducted a costing study on 909 women across 10 countries as part of the EndoCost Consortium. This included UK women and so was considered suitable for inclusion despite not directly representing a UK population.

The perspective of the study was not suitable for NICE analysis as the main outcome measures included productivity loss rather than health related quality of life. However, the study disaggregated the outcome measures, which meant it was possible to use its figures for total cost and health-related quality of life. The design of the study was questionnaire based, using the EndoCost questionnaires and with a response rate of 28%. In general, costs were calculated using national repayment tariffs, but UK costs in particular were taken directly from actual resource use. Where costs could not be calculated, a conservative value of €0 was used, indicating that the figures published are probably slight underestimates.

Results were given as mean, standard deviation, minimum cost and maximum cost in a variety of fields. On average, it cost €3113 (£2651) to treat women with endometriosis with a standard deviation of €13,244 (£11,279). The most significant items of this cost were surgery, monitoring costs, hospitalisation and physician visits. As this was a costing study, no specific hypothesis about the data was to be tested – therefore no comment can be made on the statistical significance or otherwise of this data. On regression analysis the study found that the treatment of UK-based patients costs around half as much as the treatment of patients from other countries, but this finding was probably better explained by chance (p=0.815) and so was not used to inform the model.

ISBN 978-1-4731-2661-9

Prast 2013

Prast conducted a costing study of 73 Austrian women with endometriosis over a time period of 1 year. The small numbers of patients and non-UK setting would typically make such a study a weak source of evidence for a NICE costing analysis.

The perspective was, again, on productivity losses and direct healthcare costs, which was not suitable for NICE analysis. However, the study disaggregated these costs, which allowed it to be included. No quality-of-life information was collected. The study was a direct cost analysis design with a questionnaire method to elicit expected subsequent costs.

Results were given as mean and standard deviation. On average, it costs €3,466 (€3,712) in surgical costs and a further €117 (€294) in medical costs to treat Austrian women with endometriosis, which is equivalent to £2953 (£3161) in surgical costs and £100 (£250) for medical. This is comparable with Simoens' results, but with significantly less variation in costs; it is unclear whether this is because Simoens includes the typically high-cost US system in his analysis or because the small number of patients means the Prast is less likely to find extreme outliers.

5.1.5.1.3 Methods

Basic model structure

The model is based on a threshold analysis, where the estimated costs of a referral into specialist endometriosis services are contrasted against the distribution of costs of women with endometriosis, and the crossover point (i.e. the marginal woman) is identified with sensitivity analysis tables.

Some additional complexity is added by considering various probabilistic factors such as the ability of the healthcare system to accurately discriminate between high- and low-need cases on referral to gynaecology services.

Time horizon

The time horizon of the model is 1 year. This is a limitation on the model imposed by the data sources used to construct it.

Discount rate

As the time horizon is 1 year or less, no discount rate was applied.

Interventions and comparisons

The intervention is referral to specialist endometriosis services, which are defined as centres specialising in the treatment of endometriosis, with the following clinicians available for the treatment of endometriosis. It is assumed the vast majority of this treatment will be surgical in specialist endometriosis services (endometriosis centres):

- gynaecologists with expertise in diagnosing and managing endometriosis, including advanced laparoscopic surgical skills
- a colorectal surgeon with an interest in endometriosis
- a urologist with an interest in endometriosis
- an endometriosis specialist nurse
- a multidisciplinary pain management service with experience in pelvic pain
- a healthcare professional with specialist expertise in gynaecological imaging of endometriosis
- advanced diagnostic facilities (for example, radiology and histopathology)

ISBN 978-1-4731-2661-9

- fertility services.

The comparison is conventional treatment, which is defined in Simoens and Prast, and approximately translates to treatment in gynaecology services.

Outcome modelling assumptions

Effectiveness of specialist endometriosis services

The model assumes that most of the variation in the cost of women being treated for endometriosis is related to errors or complications in the treatment of their endometriosis. This assumption might not be true; it is not clear from the evidence what might happen to a woman who is diagnosed with endometriosis but who then incidentally has a heart attack while in hospital (i.e. where her costs would no longer be related to the treatment of endometriosis). However, it is likely (based on similar papers) that there would be no way the authors would be able to exclude this woman from the study. The more important assumption is that specialist endometriosis services have negligible variation in costs, apart from known variation in the complexity of operation.

This assumption implies that women referred to specialist endometriosis services are treated correctly the first time (and do not require multiple rounds of retreatment), do not have unexpected complications during an operation owing to surgical error and have comprehensive aftercare, meaning they do not have unexpectedly long post-surgical recoveries. While it is probably true that specialist endometriosis services reduce such errors, it is a strong assumption that they disappear completely. However, based on Committee experience, it was assumed that major surgical error would be rare in skilled specialist endometriosis surgeons. Therefore the assumption that these errors are negligible is supportable from their clinical experience.

Accuracy of stratification

The model relies on clinicians who would otherwise refer to gynaecology services instead referring to specialist endometriosis services, or alternatively, gynaecologists recognising when they are faced with an especially complex case and referring from there. It was assumed that healthcare professionals in the NHS would be unlikely to identify the most costly cases perfectly, but on the whole their stratification would be considered to be reasonably good. For an illustrative example of what this means in practice, if faced with 10 patients of varying complexity and expense to treat, the Committee might be able to identify the 3 most expensive patients, given 4 attempts to select them – the most expensive 2 patients are easy to select, but the difference between the third and fourth most expensive patient might be slight.

To reflect this potential for inaccuracy, an estimate of 75% was used for an ‘accuracy of stratification’ parameter, meaning that 75% of patients who are sent to specialist services will – in hindsight – have been correctly sent there.

Prevalence vs incidence

The two data sources used for this model both give prevalence figures (based on the EndoCost consortium), but the model makes more sense if it assumes that these are incidence figures. This is because the assumption is that treatment for endometriosis (particularly surgical treatment) is functionally a ‘one off’ and does not need repeating. This assumption is incorrect for e.g. drug prescribing, but since postoperative drugs such as hormonal contraception will be prescribed by both specialist and non-specialist services this is not thought to represent an opportunity cost. The Committee agree that the figures appear sensible, based on their experience, and so published literature sources were preferred.

ISBN 978-1-4731-2661-9

In order to correct for any effects of prevalence vs incidence, the key output from the model was in percentage format. This means that if the number of women seeking treatment for endometriosis decreases dramatically once the 'stock' of women who currently have badly treated endometriosis dwindles (due to treatment at specialist endometriosis services, perhaps), it should be simple to calculate the number of patients it is likely to be cost-effective to treat from these figures.

Costs

The costs of conventional care are given in Simoens and Prast, and are based on distributions calculated from their figures. As the two papers report only summary figures and do not appear to have appendices with associated data, assumptions over the correct distribution must be made. Committee opinion is that the majority of women will have middling costs of around £1000-£5000 to treat, while some women will have very large costs associated with their treatment, suggesting that standard distributions such as the normal distribution will significantly underestimate the costs of most women. Consequently it was decided to use a 'fat tailed' distribution such as the Weibull or log-normal distribution. More complex distributions like the gamma were considered, but did not appear to add much to the fit of the model. As there was no guidance on fitting distributions in the NICE Reference Case, the log-normal was chosen for the base case since its statistical properties would be easier to explain to the Committee and was a more 'natural' choice since normal distributions were used elsewhere in the model. Since Simoens both reported more data and reported data for more patients, his figures were used in the 'base case' of the model, fit to a log-normal distribution for reasons described above. Both of these assumptions were tested in sensitivity analysis.

The costs of specialist endometriosis services are modelled de novo and are assumed to be mostly related to staff wages. These wages are assumed to be related to those in the Personal Social Services Research Unit (PSSRU) Cost of Health and Social Care:

Table 9: Annual cost to NHS of specialties involved in specialist treatment of endometriosis (wages plus additional oncosts)

Role	Wage	Wages plus oncosts
Gynaecological specialist	£87,449	£195,684
Endometriosis specialist nurse	£38,550	£91,469
Non-specialist nurse	£25,902	£40,502
Colorectal surgeon	£87,449	£195,684
Urologist	£87,449	£195,684
Pain management specialist	£87,449	£195,684
Radiologist	£87,449	£195,684
Fertility specialist	£87,449	£195,684

(a) All values taken from PSSRU Unit Cost of Health and Social Care, 2016 (http://www.pssru.ac.uk/project-pages/unit-costs/2016/index.php)

Each of these specialities is required for a varying amount of time depending on the complexity of the procedures. These timings are based on Committee consensus, which is discussed in greater detail in Appendix K.

Table 10: Estimated time required per operation by complexity

Role	Hours per superficial operation	Hours per adnexal operation	Hours per deep operation	Hours per complex deep operation
Gynaecological specialist[a]	0.50	1.50	2.50	4.00

ISBN 978-1-4731-2661-9

Role	Hours per superficial operation	Hours per adnexal operation	Hours per deep operation	Hours per complex deep operation
Endometriosis specialist nurse[b]	0.67	2	2	3.33
Non-specialist nurse[b]	1.33	4	4	6.67
Colorectal surgeon[c]	0.00	0.00	1.92	3.08
Urologist[c]	0.00	0.00	1.69	2.70
Pain management specialist[d]	0.10	0.87	1.45	2.32
Radiologist[d]	0.50	1.00	1.00	1.50
Fertility specialist[d]	0.50	0.50	0.50	0.50

(a) Based on Committee consensus, see Appendix K.
(b) Based on care provided by 1 specialist and 2 non-specialist nurses for the duration of hospitalisation following operation, which is also given by Committee consensus in Appendix K. Assumes 6 patients per ward.
(c) Based on The British Society for Gynaecological Endoscopy (BSGE) staffing figures, available from http://bsge.org.uk/centre/ retrieved 28/10/16.
(d) Based on assumption informed by the Committee.

The second major cost is the cost of complications, which are also calculated in Appendix K. As the model is not probabilistic, these complications are averaged over each operation.

Table 11: Expected cost of complications by operation complexity

Operation	Expected cost of complications
Superficial operation	£0.00
Adnexal operation	£7.56
Deep operation	£68.06
Complex deep operation	£544.44

Finally, a cost of £24.50 per recovery hour is added, based on the cost of an excess elective inpatient bed day for 'Non-Malignant Gynaecological Disorders with Interventions, with CC Score 0-2' divided by 24, and a cost of £1766.95 added per operating room hour based on the difference between the staff and recovery costs and the NHS Reference Costs for a day case 'Intermediate Female Pelvic Peritoneum Adhesion Procedures (https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016) to account for the opportunity cost of using an operating theatre for an endometriosis excision rather than another operation.

From these tables it is possible to calculate the expected cost per operation, which is given in Table 12. There is good agreement with the method of cost calculation in Appendix K, but this method of costing is more appropriate for a service delivery question as it allows the Committee to test assumptions such as varying the number of specialists involved in the operation or see 'what if' for instance discharge planning could be sped up.

Table 12: Expected cost of operations of various complexity.

Operation	Expected cost (using NHS Reference Cost uprating from Appendix K)	Calculated cost using 'bottom up' model for service delivery
Superficial operation	£1,494.89	£1364.14
Adnexal operation	£4,201.06	£4042.10
Deep operation	£6,614.77	£6398.66

ISBN 978-1-4731-2661-9

Operation	Expected cost (using NHS Reference Cost uprating from Appendix K)	Calculated cost using 'bottom up' model for service delivery
Complex deep operation	£10,622.33	£10,710.60

There are no figures on the estimated split of condition of patients who are referred to specialist services (not least because Committee recommendations could seek to alter this balance). However, figures from the units of Committee members who operate in a specialist environment suggest that around 25% superficial and adnexal endometriosis, 30% bowel infiltrating and 20% complex bowel infiltrating is probably the right order of magnitude, which would give the average operating costs on a typical patient referred to specialist endometriosis services as £6940.

5.1.5.2 Health-related quality of life

There was no comparative evidence available on the quality of life of women treated in specialist endometriosis services (endometriosis centres) compared to conventional care. The possibility of improved quality of life serving as the argument for more referrals into specialist endometriosis services is considered in sensitivity analysis.

5.1.5.3 Results

Analysis shows that – depending on the assumptions chosen – somewhere between 2.6% and 3.2% of women with endometriosis could be treated in specialist endometriosis services for less than they would cost to treat in gynaecology services. This is estimated to be somewhere between 7,800 and 9,300 women, depending on estimates of the population of England and Wales with symptomatic endometriosis.

The model is designed to be cost-minimising, meaning that the estimate of between 2.6% - 3.2% of women may not be the most cost-effective number of women to treat (but it is likely to be the cheapest, and highly likely to use fewer resources on net than currently). Therefore the Committee may wish to consider recommending a service which services more or fewer women depending on clinical considerations not included in this model.

The estimated saving of the most cost-minimising choice of specialist endometriosis services design is on the order of magnitude of £25m, but since the NHS does already provide some specialist endometriosis services to high-risk women, the actual saving is likely an order of magnitude lower.

However, the results strongly imply that there should be a large transfer of resources from gynaecology to specialist endometriosis services; likely well above the threshold for a 'high' resource impact. The Committee therefore considered their recommendations in light of this. In particular, the Committee was careful to allow for a variety of possible implementation strategies – provided a minimum clinical competence threshold was reached – to try and limit the extent of resource transfer where possible.

The model assumes the transfer of women to specialist endometriosis services does not improve their health (although this assumption is varied in sensitivity analysis). Instead it finds an economic case for the recommendations by identifying that a small fraction of women with very complex endometriosis are unlikely to have their condition properly addressed outside of highly specialisted services, causing the potential for reoperation or side effects of treatment

ISBN 978-1-4731-2661-9

5.1.5.3.1 Sensitivity analysis

Choice of distribution

Table 13 demonstrates the results of the economic model for a variety of possible distributions and underlying data selections. The results are not notably sensitive to the choice of distribution, provided that distribution is 'fat tailed'. However, attempts to use statistically simpler but less well-fitting distributions such as gamma do not produce good agreement with the evidence and therefore create very unusual results.

Table 13: Results for different distribution profiles

Distribution	Percentage of patients that could be referred to cost-minimise	Estimated number of patients in specialist services	Cost of marginal patient
Simoens, log normal	3.1%	9,300	£1,666
Simoens, Weibull	2.6%	7,800	£1,617
Prast, log normal	2.8%	8,400	£2,073
Prast, Weibull	2.8%	8,400	£2,069
Simoens, gamma (poor fit)	0.7%	2,100	£1,253

The results are well clustered around the central estimate, indicating that the choice of distribution is not important provided it is well parameterised for the data.

Accuracy of risk stratification

The Committee determined that the accuracy of stratification was likely high in reality; they discussed that based on clinical experience it would usually be clear when lesions of a similar size would provoke complexity owing for example, to site of the lesion. Consequently, a lower-bound estimate of 75% was used in the base case, as discussed above. Varying this 'accuracy' parameter produces estimates for the number of patients who should be referred in scenarios of high and low accuracy, and is demonstrated in Table 14.

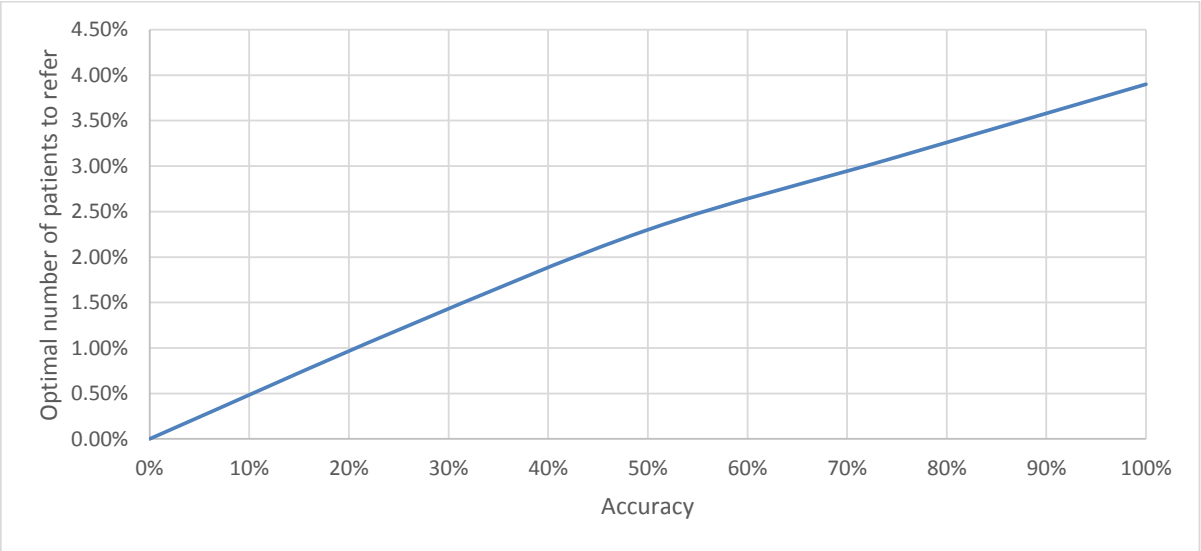
Table 14: Results for different risk stratifications (Simoens, log normal)

Accuracy	Percentage of patients that minimises cost to NHS
100%	3.90%
75% (base case)	3.10%
50%	2.30%
25%	1.20%
0%	0.00%

The relationship between accuracy and percentage of patients who should optimally be referred is roughly linear, as demonstrated by Figure 4, indicating that it would be valuable to become more accurate in this their assessment because– better stratification would have a moderate and direct effect on the cost effectiveness of specialist endometriosis services.

ISBN 978-1-4731-2661-9

Figure 4: Relationship between stratification accuracy and optimal number of patients to refer

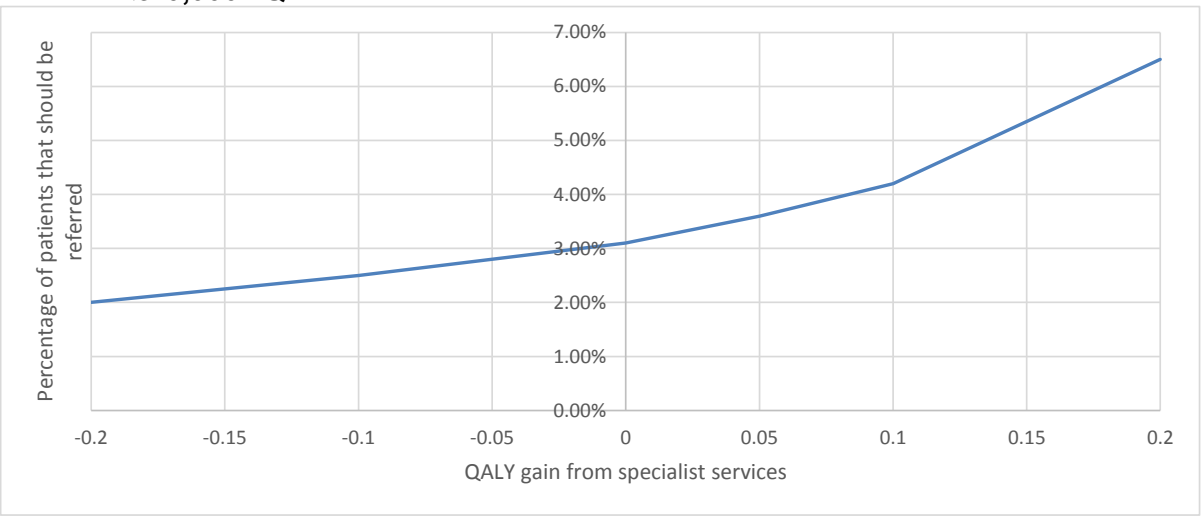


Source: 'Specialist Endometriosis Services' economic model

Quality-of-life impact of specialist endometriosis services

In the base case, specialist endometriosis services do not improve quality of life. The Committee strongly disagreed with this assumption, and asked for this parameter to be considered in sensitivity analysis. Unpublished data from The British Society for Gynaecological Endoscopy (BSGE) suggests that the maximum possible gain from specialist endometriosis services is 0.20 quality adjusted life years (QALY) sustained over a period of some years, so sensitivity analysis will consider QALY gain from -0.2 to +0.2 QALY (Figure 5).

Figure 5: Relationship between QALY gain from specialist endometriosis services and percentage of patients to refer for optimal cost / effects trade-off at £20,000 / QALY



Source: 'Specialist Endometriosis Services' economic model

Over a plausible range of QALY values, the percentage of patients who should be referred to specialist endometriosis services varies from around 2.00% to around 6.50%. Although the

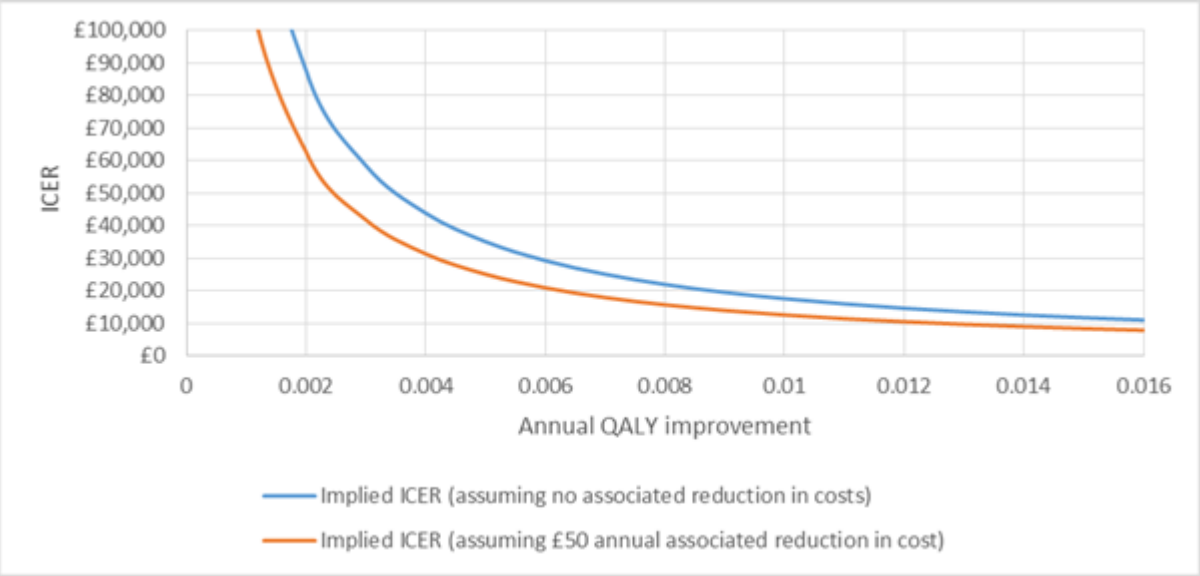
ISBN 978-1-4731-2661-9

extreme values here are quite large, in general the effect is small for more plausible effect sizes (0.05 QALY, for example).

It should be noted that at around 0.3 QALYs added from specialist endometriosis services (which is an extremely unlikely value for the quality of life gain from specialist endometriosis services), almost 100% of patients are recommended into specialist endometriosis services; this is where the average value of the QALY gain is higher than the average cost of treatment in specialist endometriosis services. The model therefore cannot be relied on for accurate values given extreme parameters for QALY gain.

Figure 6 shows the implied ICER for a pain clinic costing £1500 (Committee estimate based on ten group sessions and two consultant appointments) and giving a variable number of QALYs per year for ten years, discounted at 3.5%. Two results were presented; one where the clinic did not reduce overall spending for the NHS and one where the clinic reduced NHS spending by £50 / year, for a discounted overall total reduction of £428 over the ten years of the estimate.

Figure 6: Estimated ICER of pain management programme given possible QALY improvement



Source: Economic model

5.1.6 Clinical evidence statements

No clinical evidence was identified.

5.1.7 Evidence to recommendations

5.1.7.1 Relative value placed on outcomes considered

The Committee agreed that improvement in pain, better quality of life and improved participation in activities of daily living would be critical outcomes for this review. However, no evidence was identified to address these or any other outcomes. Outcomes related to costs were also considered to be important, such as length of hospital stay, further treatments, and readmission to hospital.

ISBN 978-1-4731-2661-9

No clinical evidence was identified.

5.1.7.2 Consideration of clinical benefits and harms

The Committee acknowledged that although the review did not bring any clinical evidence to light, based on their experience, the Committee agreed that having specialist endometriosis services provided better outcomes. However, this data was not specific to endometriosis. The Committee emphasised that the specialist endometriosis service should be provided by professionals who have expertise and training in the management of endometriosis and follow good practice to provide a high standard of care.

The Committee further noted that referral to these services may take time, but that the benefits of the care provided by these would outweigh the harms of having to wait for this to happen. It was noted that a delay in referral was sometimes related to a misdiagnosis for women with deep endometriosis involving the bowel, bladder or ureter. The Committee acknowledged that some symptoms do overlap with irritable bowel syndrome or painful bladder syndrome. However, the hormone dependent cyclical pattern should distinguish these symptoms and should make appropriate referral or at least the suspicion of endometriosis possible.

5.1.7.3 Consideration of economic benefits and harms

As no evidence was found for specialist endometriosis services, recommendations were based on expertise and discussion of the Committee and information from the Health Economic model.

The model found that if specialist endometriosis services improved quality of life outcomes following operation, a very large proportion of women should be treated in these services. The Committee argued that it was reasonable to assume that most women were being treated well in gynaecological services (those with endoemtriosis that is not involving the bowel, bladder or ureter and those with uncomplicated endometriomas or endometriosis that responed well to medical treatment) and so the possibility of a large quality of life increase was unlikely in these women, but that there was a potentially large improvement for women with highly complex endometriosis currently being treated in gynaecological services. Nevertheless the Committee agreed that no comparative data existed comparing gynaecological services to specialist endometriosis services when controlling for casemix so it was reasonable for the model to attempt to assume a zero quality of life increase and draw conclusions based on cost alone. Unpublished indicative data suggests that this assumption might be too conservative.

The model finds that some women have endometriosis which is so complex it is being poorly managed in gynaecological services. The data upon which the model are based do not go into details on what is causing these women to accrue large treatment costs, but the Committee agreed it was reasonable to assume it would include reoperation following an unsuccessful operation, complications caused by errors in surgery and increased recuperation time due to inexpert or non-specialist post-operative nursing care. The model therefore tries to identify the fraction of women for who it would be cost-saving to treat in specialist endometriosis services.

Summary of model findings that were discussed

The Committee understood that the analysis showed that – depending on the distribution and primary source for variance chosen – somewhere between 2.6% and 3.2% of women with endometriosis could be treated in specialist endometriosis services (endometriosis centres) for less than they would cost to treat in gynaecology services. This is estimated to be somewhere between 7,800 and 9,300 women depending on estimates of the population of England and Wales with symptomatic endometriosis.

ISBN 978-1-4731-2661-9

Furthermore the Committee were reassured that the planned sensitivity analyses (change in distribution, accuracy of risk stratification and the impact on the quality of life of specialist endometriosis services) did not change the conclusion of the model.

Discussion of pain management services

The Committee believed that a pelvic pain management service would likely be helpful for women with endometriosis, but were unsure if it would be cost-effective. The Committee described how a pelvic pain management service required a somewhat high initial outlay of resources for consultations and group discussions, but then was expected to offer benefits for a very long time, potentially as long as the pain lasted. These benefits included direct cost savings to the NHS through medicines optimisation and reduced GP appointments, and QALY benefits due to helping women with their pain.

Although the Committee considered that the £50 / year estimate would be unrealistically low - as it equated to only one marginal GP appointment not made per year –the figures that related to these results were otherwise accepted as illustrative. The Committee discussed how the simple cost estimate showed that pain management was likely to be cost-effective at £20,000 / QALY if it added around 0.009 QALY annually in the case where it led to no cost savings and around 0.006 QALY in the case where it led to very minor cost savings. The Committee believed that the true QALY improvement was likely to be greater than 0.03 QALY based on their clinical experience and evidence from fields other than endometriosis, which is substantially higher than either of the threshold analyses seen by the Committee.

Consequently the Committee recommended that women should have access to a pain management programme if it was thought it would help with pain, on the grounds that their clinical experience suggested that women would benefit and the costing analysis discussed above strongly implies that this recommendation is cost-effective.

Other economic considerations

The Committee considered that the NHS is already commissioning such services (https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2014/04/e10-comp-gynae-endom-0414.pdf) and it was therefore agreed that the economic impact would not include a significant amount of implementation costs, although as described above it may involve a large transfer of resources across sectors.

5.1.7.4 Quality of evidence

No clinical evidence was identified.

5.1.7.5 Other considerations

Current practice and issues

In current practice, gynaecologists with expertise in advanced laparoscopic surgery for endometriosis are restricted to specialist endometriosis services. However, there are not many gynaecologists who act as specialist leads, which is further reflected in gynaecology services where there are no such specialist gynaecologists in service provision for women who have mild or moderate endometriosis. As a consequence, women with mild to moderate endometriosis are also referred to specialist endometriosis services (endometriosis centres) for further treatment, so women with any level of severity of endometriosis are currently been seen in specialist endometriosis services (endometriosis centres). The Committee felt that an endometriosis service would benefit from having 3 separate levels of care so that only women with severe endometriosis are referred to specialist endometriosis services (endometriosis centres) for complex treatment such as surgery, whereas women with mild to moderate endometriosis are referred to gynaecology services for surgical and non-surgical

ISBN 978-1-4731-2661-9

treatments. Women who come to gynaecology services who are diagnosed with severe endometriosis could then be referred further to specialist endometriosis services (endometriosis centres), so that the correct women are being referred to the appropriate service. It has also been further noted that there should be awareness regarding young women (17 years age and under) with symptoms of pelvic pain or endometriosis who are referred to specialist endometriosis services (endometriosis centres) for surgery even though they may not require surgery but may benefit from non-surgical diagnosis (imaging such as magnetic resonance imaging [MRI]).

Components of the multidisciplinary team (MDT)

In order to develop a service for endometriosis that would cover 3 levels of care (community, gynaecology and specialist endometriosis services (endometriosis centres)), the Committee considered the configuration of the service that would be most appropriate and cost-effective. The Committee considered the model of managed clinical networks to provide better access to women and therefore earlier diagnosis of the condition. Since there was no clinical evidence that was identified, the Committee suggested that the de novo economic model could provide evidence on whether the gynaecology and the specialist endometriosis MDT would be cost-effective in the model.

At gynaecology service level, the Committee considered that the team should include: a gynaecologist with interest in pelvic pain and expertise in diagnosing and managing endometriosis, including training and skills in laparoscopic surgery, and a gynaecology nurse with an interest in endometriosis. There should be access to a specialist pain management team and to fertility services. Diagnostic services would require a radiologist with an interest in gynaecological imaging who would identify cases in the gynaecology service to refer further to specialist endometriosis services (endometriosis centres).

At specialist endometriosis services (endometriosis centres) level, it was considered that women with severe endometriosis should have access to a full multidisciplinary team, including a gynaecologist with expertise in endometriosis, including advanced laparoscopic surgery, a colorectal surgeon, urologist, endometriosis specialist nurse, multidisciplinary pelvic pain management service and advanced diagnostic services (i.e., radiology and histopathology); there should be access to fertility services. Because of the issues of delayed or missed diagnosis, the Committee noted the importance of the surgeon's expertise to identify endometriosis in the range of places it can occur and of good visual documentation during laparoscopy and written description which can save further operations. The Committee therefore considered that gynaecologists should be able to perform techniques to skill levels described by professional body technical standards and in specialist endometriosis services have advanced training (e.g. as provided by RCOG). The Committee agreed that, even though more costly, these specialists would treat women with the most severe type of endometriosis who are a small proportion of all women with endometriosis. Therefore if triaged this would be a cost-effective service.

Of particular concern to the Committee was that all women with endometriosis should have access to appropriate support from specialist nurses. The nurse acts as a vital link between the women and their management pathway, being available to communicate with and support them when required. As it is likely that most nurses recruited to this role will have other roles within gynaecology (as part of their working week) the Committee emphasised that the change in practice implied here would be dependent on nurses receiving robust training, preferably with an accredited course. This should equip the nurse to provide knowledge and acquire expert skills enabling them to meet the woman’s physical, psychological/psychosexual and social needs and expectations and provide support to them and their families through their pathway. The Committee discussed how a job specification and accreditation pathway might be developed from the existing specification for the Endometriosis Specialist Nurse, and suggested that professional organisations such as the Royal College of Nursing might be well positioned to develop such a specification. The

ISBN 978-1-4731-2661-9

Committee added that they expected the role will be constantly developing, as will the management of women with endometriosis. The Committee considered that the specialist services would be expensive and would require time in terms of implementation. There are downstream costs such as time for regular MDT meeting, planning of surgery or other treatment strategies. The Committee was aware that there is experience from the cancer MDT formation that such a configuration does work and is feasible. In addition, to make this a clinical and cost-effective service it would have to have healthcare professionals with the appropriate expertise, and also would have to treat a sufficient number of women to make it viable (based on the Committee's experience and expert opinion, the minimum requirement of cases was 12 per year per gynaecologist).

Linking of the endometriosis services network

The Committee discussed how the network from community services to gynaecology services could be linked in terms of identifying suspected cases from the community to be referred to gynaecology services. At the community services level, it was discussed whether suspected cases of endometriosis could be identified and triaged by a GP, since there are only a few who are specialised in gynaecological conditions and practice is varied across the UK. Currently endometriosis is not always suspected even if a women presents with symptoms and signs (see chapter 4).

Mental health/psychological issues

The Committee highlighted that mental health issues, including depression and anxiety, need to be addressed in the services (see chapter 7 which highlights those support needs). These may present to the GP, nurse or gynaecologist and may arise at any stage; for example, for women who have delayed diagnosis resulting in loss of fertility. Access to psychological services can be provided by GP by direct referral to counselling services rather than by referral to specialist endometriosis services (endometriosis centres).

Key conclusions

Due to the lack of applicable clinical evidence, the Committee based the recommendations on the health economic model as well as on their experience and expertise. They considered that it would be possible to stratify women with endometriosis involving the bowel, bladder or ureter to specialist endometriosis services (endometriosis centres) and that this is therefore a targeted smaller group of women that would receive this service. Access to these services would be improved through managed clinical networks. Since these services already exist it will not require a significant cost in setting up these services and therefore strengthen the cost effectiveness of these services.

5.1.8 Recommendations

1. **Set up a managed clinical network for women with suspected or confirmed endometriosis, consisting of community services (including GPs, practice nurses, school nurses and sexual health services), gynaecology services and specialist endometriosis services (endometriosis centres).**

Gynaecology services for women with suspected or confirmed endometriosis

2. **Gynaecology services for women with suspected or confirmed endometriosis should have access to:**
    - a gynaecologist with expertise in diagnosing and managing endometriosis, including training and skills in laparoscopic surgery
    - a gynaecology specialist nurse with expertise in endometriosis

ISBN 978-1-4731-2661-9

- a multidisciplinary pain management service
- a healthcare professional with an interest in gynaecological imaging
- fertility services.

Specialist endometriosis services (endometriosis centres)

3. Specialist endometriosis services (endometriosis centres) should have access to:

- gynaecologists with expertise in diagnosing and managing endometriosis, including advanced laparoscopic surgical skills
- a colorectal surgeon with an interest in endometriosis
- a urologist with an interest in endometriosis
- an endometriosis specialist nurse
- a multidisciplinary pain management service with expertise in pelvic pain
- a healthcare professional with specialist expertise in gynaecological imaging of endometriosis
- advanced diagnostic facilities (for example, radiology and histopathology)
- fertility services.

5.2 Timing: association between duration of symptoms before laparoscopy and treatment outcomes

Review question: Is there an association between duration of symptoms before laparoscopy and /or treatment and treatment outcomes?

5.2.1 Introduction

This section will assess whether there is an inverse association between the length of time that a women had symptoms before laparoscopy and the effectiveness of the treatment.

Endometriosis patients present with a range of symptoms; which can vary from very mild to severely debilitating. Often women with endometriosis have experienced symptoms for a long time before they are diagnosed or treated. This delay may alter the stage of the disease and result in a need to adopt different treatment options. It can be argued that any delay in treatment will prolong the women's suffering and have a negative impact on quality of life, including social and work interactions. A delay in treatment may accrue costs for the NHS because treatment options may become more complex and costly due to the progression of the condition, or could potentially be less effective.

For full details, see review protocol in Appendix D, the study selection flow chart in Appendix F and study exclusion list in Appendix H.

5.2.2 Description of clinical evidence

No clinical evidence was identified for this review.

5.2.3 Summary of included studies

No clinical evidence was identified for this review.

ISBN 978-1-4731-2661-9

5.2.4 Clinical evidence profile

No clinical evidence was identified for this review.

5.2.5 Description of economic evidence

The issue of the timing of interventions was of very great importance to stakeholders and members of the Committee, and might carry large health economic consequences. A literature search was undertaken of the health economics literature and no studies were found comparing early to late interventions.

Consequently this question was prioritised for de novo health economic modelling, the details of which are described in Appendix K and a summary is provided below.

5.2.5.1 Summary of relevant section of the health economic model

A summary table from the model is reproduced in Table 15. This shows that for all reasonable cost / quality adjusted life year (QALY) thresholds the NHS might consider, there would have to be extremely strong reasons to delay treatment for women with pain and/or infertility as a main symptom. This is not true for a group of women with asymptomatic endometriosis (which is discovered incidentally); these women would only be cost-effective to treat at £179,943 / QALY. This is most likely an artefact of the model due to not simulating enough women to completely eliminate random variation, as there is no biological reason why these women should benefit from treatment.

Table 15: Summary table of health economic results by subgroup

Subgroup	Cost 1 year faster diagnosis	QALY gain 1 year faster diagnosis	ICER of 1 year faster diagnosis	Probability 1 year faster diagnosis cost effective at £20,000 / QALY
Pain only	£806	0.20	£4,075	93.7%
Infertility only	£1,907	0.19	£10,000	82.9%
Both	£1,068	0.21	£5,093	84.6%
Asymptomatic	£1,584	0.01	£179,943	N/A

(e) ICER: Incremental cost-effectiveness ratio; QALY: Quality adjusted life years

5.2.6 Clinical evidence statements

No clinical evidence was identified for this review.

5.2.7 Evidence to recommendations

5.2.7.1 Relative value placed on the outcomes considered

The aim of this review was to identify whether it would be both clinically and cost-effective to treat symptoms as early as possible (early with regard to presentation or how long the symptoms have been present rather than early as in the age of the women reporting the symptoms). The Committee prioritised relief of endometriosis-related pain, health-related quality of life and adherence to the treatment programme as critical outcomes when considering recommendations. The remaining outcomes of improvement in fertility rates (spontaneous, i.e. unassisted, pregnancy rates), reduction in the size and extent of endometriotic cysts, improvement of endometriosis-related symptoms apart from pain (e.g. fatigue), adverse effects resulting from the intervention, rates of reoccurrence and activities of daily living were considered to be important. However, no evidence was identified.

ISBN 978-1-4731-2661-9

5.2.7.2 Consideration of clinical benefits and harms

As clinical evidence was not identified in the review, the Committee suggested that timing of interventions could be addressed by cost-effectiveness in the de novo health economic model. The Committee noted that it was important that women diagnosed with endometriosis were treated early as this would be a cost-effective approach, as a delay in referral would result in endometriosis becoming more severe and therefore may be more harmful for women. The Committee suggested that a recommendation related to the organisation of services could be made since women who are not treated early may develop more severe symptoms of endometriosis. The Committee wanted to make a strong recommendation for early referral, diagnosis and treatment. There was also discussion about persistent symptoms (when and how long is a symptom considered to be persistent) and prompt referral (what is meant by 'prompt').

The Committee discussed the fact that no individual healthcare professional intentionally delays the diagnosis of endometriosis, but that there was nevertheless concern among patients that delays in diagnosis may be being introduced by clinicians not suspecting endometriosis until some time after initial presentation (for example because some symptoms or signs could be misinterpreted as another condition). The Committee agreed that clinicians should suspect endometriosis as soon as symptoms and signs are reported at the time of first presentation. It was agreed that the guideline should promote the awareness of this condition and therefore speed up the recognition of endometriosis in future.

5.2.7.3 Consideration of economic benefits and harms

The Committee agreed that the cost-effectiveness model showed that a delay in treatment was extremely unlikely to be cost-effective for the NHS.

It was noted that this does not consider the costs of actually implementing services to reduce the delay of diagnosis and treatment; if, for example, it was discovered that the main reason for the delay was that women did not recognise the symptoms then it could be a cost-effective solution to raise awareness of this condition (which this Guideline would promote).

The Committee identified this need and pointed out that there were many reasons for a delay in diagnosis and treatment, and indeed delay was introduced at many different stages.

It is unclear what effect – if any – these recommendations will have on NHS resources, as the resource impact is entirely to do with how strongly these recommendations can be implemented. For example, each year faster endometriosis is diagnosed costs approximately £806, which means if approximately 1250 women are diagnosed a year faster each year, the resource impact will be high under NICE definitions. It should additionally be noted that the resource impact of these recommendations are – to a certain extent – out of the NHS' hands; patients can reduce the delay in diagnosis by asking doctors to consider treatment for endometriosis, meaning that regardless of the recommendations made the resource impact may go up or down depending on changes to patient understanding of the disease.

5.2.7.4 Quality of evidence

No clinical evidence was identified.

5.2.7.5 Other considerations

The Committee highlighted that, although no studies had been carried out to address timing of interventions, research should continue because they felt that the lack of evidence did not reflect on the efficacy of carrying out such research; however, it was also acknowledged that evidence from this area of research would be difficult to identify as it was unclear which study design would be appropriate to identify such data.

ISBN 978-1-4731-2661-9

They decided not to prioritise this as a research recommendation because actively delaying treatment would not be ethical and that retrospective research would suffer from a number of biases. To be robust it would have to be a very large study to account for confounders. For instance, it would most likely be the case that those diagnosed early were those women who had more severe symptoms and that this group would therefore be over-represented. Differences in treatment regimes, analgesic regime and other factors may also bias results.

5.2.7.6 Key conclusions

The Committee agreed with the conclusions from the de novo health economic model which showed that in all patient populations with endometriosis, a delay in diagnosis and treatment was not beneficial to the NHS given their typical willingness to trade resources for health at around £20,000. The model demonstrated that delays in treatment led to an overall cost saving despite the increased cost of treating more progressed endometriosis, but found that this saving was outweighed by the harm to the quality of life of the women with endometriosis that a delay caused. In the absence of clinical evidence the conclusion from the de novo economic model is consistent with clinical expert consensus.

5.2.8 Recommendations

4. Community, gynaecology and specialist endometriosis services (endometriosis centres) should:

- provide coordinated care for women with suspected or confirmed endometriosis
- have processes in place for prompt diagnosis and treatment of endometriosis, because delays can affect quality of life and result in disease progression.

ISBN 978-1-4731-2661-9

6 Signs and symptoms of endometriosis (monitoring and referral)

Review question 1: What are the symptoms and signs of endometriosis?

Review question 2: How and when should women with endometriosis be monitored and referred for the following symptoms or condition progression and complications:

- pelvic pain disrupting daily activities
- cyclical bowel pain
- cyclical voiding pain?

6.1 Introduction

In the UK the average time from symptom onset to diagnosis of endometriosis is 8 years. The key to earlier diagnosis, avoiding unnecessary pain, distress and possible disease progression, is awareness and knowledge of endometriosis among health professionals. Women often find health professionals normalise their symptoms and have limited knowledge of endometriosis. These can contribute to a delay in diagnosis and increase the risk of misdiagnosis. Women present to health professionals with a variety of symptoms that may suggest endometriosis, including pelvic pain, painful periods, painful sex, infertility, gastrointestinal and urological problems. Symptoms of endometriosis are non specific and overlap with other diseases, for example, irritable bowel syndrome (IBS) and pelvic inflammatory disease (PID). Symptoms are usually cyclical but can occur at any time throughout the month. Symptoms experienced by women may depend on the location of the disease but do not always correlate with the severity of the disease and some women with endometriosis are asymptomatic. Signs suggestive of endometriosis may be found during physical examination of the pelvis and include tenderness, tethering of pelvic organs, palpable nodules of endometriosis and visible vaginal endometriosis lesions. However, signs may be subtle and a normal examination does not exclude endometriosis.

The objective of this systematic review is to identify what symptoms and signs (or combinations of them) are predictive of endometriosis and, once identified, when women with these signs should be monitored and referred.

For full details, see review protocol in Appendix D, the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I and study evidence tables in Appendix G.

6.2 Description of clinical evidence

Three studies (Calhaz-Jorge 2004, Peterson 2013, Whitehill 2012) were included in this review. Two were prospective cohorts (Calhaz-Jorge 2004, Peterson 2013) and 1 was a retrospective cohort (Whitehill 2012).

All of the studies used a questionnaire to collate information about the patients' symptoms and all are subject to recall bias. The subjective rating of pain varied among the studies:

- broad categories with no clear definition (absent, mild, moderate, severe);
- a descriptive definition of dysmenorrhoea, for example; mild pain, being mild discomfort with no use of analgesic medication, or
- use of pain scale from 0 (none) to 10 (severe).

ISBN 978-1-4731-2661-9

Two studies (Calhaz-Jorge 2004, Peterson 2013) also reported results stage III/IV endometriosis as defined by the American Fertility Society (AFS).

None of the studies reported the following symptoms: bowel (rectal bleeding, bloating, constipation and diarrhoea), bladder (bladder irritability, blood in the urine), referred pain (leg, thigh and hip), fatigue, psychological effects (isolation, depression, anxiety, low self-esteem, low mood, poor body image, loss of libido) and signs: vaginal (visible endometriosis, severe vaginismus) or renal (loin tenderness, palpable mass). The provided evidence relates to individual symptoms and signs rather than combinations of them.

All studies used a combination of visualisation at laparotomy/laparoscopy or biopsy histological confirmation to confirm the diagnosis of endometriosis (Calhaz-Jorge 2004, Peterson 2013, Whitehill 2012).

All studies adjusted for age in the multivariable analyses, however, only 1 study also adjusted for the use of oral contraceptives (Calhaz-Jorge 2004). Other risk factors were also used in the multivariable analysis (see Table 16).

The main reason that studies were excluded from this review was due to them not performing multivariable analyses.

See also the study selection flow chart in Appendix F, study evidence tables in Appendix E and the exclusion list in Appendix H.

6.3 Summary of included studies

A summary of the studies that were included in this review are presented in Table 16.

Table 16: Summary of included studies

Study	Risk factors and their method of measurement	Outcome ascertainment measure	Analysis and outcomes (aOR and 95%CI)	Critical confounders	Comments
Calhaz-Jorge 2004 Prospective cohort Portugal N=1079 (488 with endometriosis, 591 without endometriosis)	Measured by: interview and questionnaire • pelvic symptoms (chronic pelvic pain) • uterus: pain (dysmenorrhoea), abnormal bleeding (prolonged and heavy) • vaginal pain (dyspareunia)	Endometriosis: laparoscopy (direct visualisation) or biopsy of lesions	Multivariable analysis • aOR (95%CI) endo AFS any type • mild dysmenorrhoea: 0.62 (0.46 to0.83) • irregular cycle: 0.60 (0.43 to 0.84) • aOR (95%CI) endo AFS grade III/IV: • dysmenorrhoea (any type): 2.5 (1.2 to 5.2) • moderate dysmenorrho	• age • OC use Other covariates in MVA: • ethnicity • BMI • smoking status • previous pregnancy • ever use of OC • dysmenorrhoea any type • mild dysmenorrhoea • moderate dysmenorrhoea • severe dysmenorrhoea	• subfertile population • descriptive pain definition rather than scale used • moderate risk of bias

ISBN 978-1-4731-2661-9

Study	Risk factors and their method of measurement	Outcome ascertainment measure	Analysis and outcomes (aOR and 95%CI)	Critical confounders	Comments
			ea: 1.7 (1.1 to 2.7) • severe dysmenorrhoea: 2.8 (1.5 to 5.1) • recently intensified dysmenorrhoea: 2.4 (1.3 to 4.5) • chronic pelvic pain: 2.0 (1.2 to 3.4) • irregular cycle: 0.29 (0.15 to 0.54)	• recently intensified dysmenorrhoea • primary dysmenorrhoea • dysmenorrhoea day 1–2 • chronic pelvic pain • generally regular menstrual cycle • irregular cycle	
Peterson 2013 Prospective cohort (part of the ENDO study) USA n=495 operative cohort; n=131 (a population cohort who did not have suspected endometriosis – results for these are therefore not reported in this review)	Measured by computer-assisted interview (telephone or in person), approx. 2 months prior to surgery: • pelvic symptoms (surgical indication pelvic pain vs. other) • uterus: pain (dysmenorrhoea) • infertility	visualisation at laparotomy/ laparoscopy	Multivariable analysis • aOR (95%CI) for stage III/IV endometriosis (n=473): • infertility history: 2.43 (1.57 to 3.76) • dysmenorrhoea: 2.46 (1.28 to 4.72) • pelvic pain: 1.39 (0.95 to 2.04)	• age Oother covariates in MVA: • clinical site • socioeconomic status • education • BMI • gravid • parous • infertility history • age at first consenting sex • surgical indication for laparoscopy (pelvic pain vs. other) • menstruation (past 12 months) • age at menarche • mean no. of periods • mean cycle length • mean length shortest cycle	• unclear how pain was measured • moderate risk of bias

ISBN 978-1-4731-2661-9

Study	Risk factors and their method of measurement	Outcome ascertainment measure	Analysis and outcomes (aOR and 95%CI)	Critical confounders	Comments
				• mean length longest cycle • dysmenorrhoea • pelvic pain	
Whitehill 2012 Retrospective cohort Canada N=429 (168 with endometriosis, 261 without endometriosis)	Standard questionnaire: • pelvic symptoms (chronic pelvic pain) • uterus (dysmenorrhoea) • vaginal pain (dyspareunia) • infertility (type and duration of) • pelvic signs (uterosacral/ cul-de-sac tenderness and nodularity)	Laparoscopy visualised or by biopsy	Multivariable analysis • aOR (95%CI): • primary infertility: 1.98 (1.29 to 3.04) • degree of dysmenorrhoea: 1.34 (1.1 to 1.65) • pelvic signs: • 3.81 (1.64 to 8.83)	• age Other covariates in the MVA: • primary infertility • duration of infertility • dysmenorrhoea (none, mild, moderate, severe) • deep dyspareunia • chronic pelvic pain • uterosacral/cul-de-sac tenderness • uterosacral/cul-de-sac nodularity • intrauterine filling effect • polypoid endometrium • endometriosis-focused practice	• no clear definition of the levels of pain • moderate risk of bias

AFS: American Fertility Society; aOR: adjusted odds ratio; BMI: body mass index; CI: confidence interval; MVA: multivariable analysis; OC: oral contraceptive pill

6.4 Economic evidence

No health economic studies were found relevant to this question, and therefore no health economic modelling was conducted for this question.

6.5 Clinical evidence statements

6.5.1 Risk of endometriosis

6.5.1.1 Pelvic pain

Evidence from 1 study (n=1079, moderate risk of bias) showed there was a significantly increased risk of stage III/IV endometriosis in women who had symptoms of chronic pelvic pain.

ISBN 978-1-4731-2661-9

Evidence from 1 study (n=495, moderate risk of bias) showed there was no increased risk of endometriosis in women who had pelvic pain.

6.5.1.2 Dysmenorrhoea

Evidence from 1 study (n=1079, moderate risk of bias) showed there was no increased risk of endometriosis in women who had symptoms of mild dysmenorrhoea; however, moderate quality evidence from 1 study (n=429) showed a significantly increased risk of endometriosis in women with increasing severity of dysmenorrhoea.

Evidence from 2 studies (moderate risk of bias) showed that there was a significantly increased risk of stage III/IV endometriosis in women who had dysmenorrhoea of any type (n=495 and n=1079) as well as moderate, severe or recently intensified dysmenorrhoea (n=1079).

6.5.1.3 Irregular cycle

Evidence from 1 study (n=1079, moderate risk of bias) showed there was no increased risk of any type or stage III/IV endometriosis in women who had an irregular cycle.

6.5.1.4 Infertility history

Evidence from 2 studies (n=495 and n=429, moderate risk of bias) showed a significantly increased risk of endometriosis or stage III/IV endometriosis in women who had a history of (primary) infertility.

6.5.1.5 Pelvic signs (uterosacral/cul-de-sac tenderness and nodularity)

Evidence from 1 study (n=429, moderate risk of bias) showed that there was a significantly increased risk of endometriosis in women with uterosacral/cul-de-sac tenderness and nodularity.

6.6 Evidence to recommendations

6.6.1 Relative value placed on the outcomes considered

The following outcomes were considered to be important for decision-making by the Committee:

- later diagnosis of endometriosis at follow-up
- severity of endometriosis
- referral to diagnostic services

The Committee also considered which symptoms and signs, once identified as a risk factor (particularly pain, bowel and bladder or ureter symptoms) should lead to the following courses of action:

- monitoring
- referral.

6.6.2 Consideration of clinical benefits and harms

The Committee discussed the impact a diagnosis of endometriosis has on women. The Committee agreed that GPs do not always suspect endometriosis and that earlier diagnostic investigation of symptoms would be of benefit to women.

ISBN 978-1-4731-2661-9

They agreed that confirmation of a diagnosis generally improves quality of life and emotional wellbeing of women who have had long-term symptoms in terms of recognition and explanation of their symptoms, and because it provides a gateway for accessing further information and support. They commented that no confirmation of a diagnosis following investigation can be difficult for women who have had symptoms.

The Committee also considered the need to distinguish pain symptoms that were associated specifically with endometriosis. For example, dysmenorrhoea is commonly experienced and can be managed successfully with analgesia, whereas in endometriosis, dysmenorrhoea would typically be more severe, perhaps requiring women to take time off work despite analgesia. They concluded that recommendations should be based on severity, frequency and persistency of symptoms to distinguish physiological from pathological pain associated with endometriosis in order to help GPs decide which women required further investigation.

6.6.3 Consideration of economic benefits and harms

The identification of signs and symptoms indicative of endometriosis might carry a very small direct cost as some signs and symptoms require examination by a medical professional. However, the main costs of this area are indirect; labelling signs that are – in actual fact – not indicative of endometriosis as being useful indicators will likely result in women without endometriosis being sent for detailed diagnosis and evaluation. Alternatively, ignoring signs and symptoms that are helpful in indicating a problem will cause women with a potentially treatable condition to go without examination, which is likely to have a quality-of-life impact and may have a direct cost if the disease progresses untreated.

Many women presenting with chronic pelvic pain or dysmenorrhoea may be treated in a similar way to those with endometriosis and hence the health economic impact of these conditions would be similar. However, for women whose primary symptom is infertility, a sign that can differentiate endometriosis from, for example, partner-related infertility is likely to have a stronger economic impact. The Committee recognised the importance of this issue, and explicitly reflected this in the recommendation made.

There is a direct cost of repeat visits to a healthcare provider such as a GP and it is well understood that failure to identify endometriosis from a description of signs and symptoms causes women to present multiple times. Consequently any recommendations that improve the recognition of signs and symptoms of endometriosis are very unlikely to carry a significant resource impact to the NHS and have a good probability of being resource saving.

6.6.4 Quality of evidence

There were only 3 studies available that provided evidence to inform this review. All included studies were assessed as having moderate risk of bias according to the NICE prognostic study checklist. The Committee broadly agreed with the evidence that mild dysmenorrhoea was not significantly associated with a diagnosis of endometriosis but that more severe dysmenorrhoea would be associated with endometriosis and that dysmenorrhoea, pelvic pain and a history of infertility would be significantly associated with more severe endometriosis.

Despite the lack of evidence, the Committee agreed that dyspareunia is one of the most common symptoms of endometriosis and that pain can be a symptom that occurs during or after sexual intercourse. It was suggested that there may be underreporting of this symptom as women may be less likely to admit experiencing dyspareunia if asked in the presence of their partner. The Committee also considered that understanding of dyspareunia might be subjective, influencing responses, for example, pain occurring during sexual intercourse or afterwards.

ISBN 978-1-4731-2661-9

The Committee found the lack of evidence for digestive symptoms (such as cyclical painful bowel movements, constipation, diarrhoea and nausea) and urinary symptoms surprising as these are quite common in women with endometriosis. They summarised that this might be because these questions were not asked but also noted that these signs can often be misdiagnosed, e.g. as IBS.

The Committee agreed that it is important to ask the right questions about symptoms and in sufficient depth to ascertain whether there was underlying pathology and an accurate differential diagnosis of this, for example, digestive or urinary signs or symptoms associated with endometriosis, would tend to be cyclical.

6.6.5 Other considerations

The Committee agreed that the pelvic examination could identify several signs that could be felt by palpation, such as reduced organ mobility and tender nodularity in the posterior vaginal fornix. The Committee commented that pain associated with pressure on the ovaries or uterine ligaments elicited during palpation can also be an indicator of endometriosis but did not include this observation in the recommendations. However, they highlighted that other signs, such as endometriotic vaginal lesions may need to be visualised by examination with a speculum. They agreed that it would therefore be important to add to the recommendation which signs may only be visualised rather than identified by touch. They also agreed that a negative abdominal or pelvic examination does not exclude endometriosis and that the persistence of symptoms may indicate that a referral may be needed for further investigations. This applies also to other diagnostic tests (e.g. ultrasound and MRI) and it was therefore decided to add this as a 'general principle' at the beginning of the recommendations on diagnostic tests (please see section 9.2.8).

The Committee also discussed the symptoms and signs that may require further monitoring or referral.

The Committee discussed particular conditions that indicate that further monitoring or referral may be required. They agreed that referral should be considered based on the severity, persistence and recurrence of symptoms. If a clinical examination indicates pelvic signs of endometriosis, this should also lead to referral.

Those women with signs suggestive of deep endometriosis involving bowel, bladder or ureter would require further investigations, surgery or both and would need to be referred to specialist services. The Committee did not want to be too prescriptive about what these signs suggestive of deep endometriosis involving bowel, bladder or ureter were because these could vary on a case by case basis. However they discussed signs such as the presence of bilateral endometriomas on ultrasound scan, presence of dyschezia, particularly cyclical dyschezia, urinary symptoms, presence of palpable or visible lesions in the posterior cul-de-sac, a fixed uterus and others. They also discussed that there are some women that may require referral to a specialist endometriosis service even though not suspected of having deep endometriosis. These could be for example women with significant neurological symptoms suggestive of sciatic or pudendal endometriosis whose scans may be negative for deep endometriosis. The Committee agreed that these specific cases requiring referral to a specialist endometriosis service could be difficult to define and that there is always room for clinical judgment in decisions about referral. Women's preferences were then discussed and it was highlighted that some women may not choose to have surgery. The Committee agreed that these women should be considered for further monitoring because their symptoms would, most likely, persist and there may also be disease progression.Equalities considerations also featured in the discussion of the evidence. One of the groups identified to be in need of specific considerations were young women (aged 17 and under). For young women (aged 17 and under) suspected of having endometriosis, referral to a paediatric and adolescent gynaecology service was seen to be more appropriate and a recommendation stating this was agreed.

ISBN 978-1-4731-2661-9

The Committee recognised the value of further research into the origins of endometriosis and its pathophysiology but a research recommendation was not made as the research question would be broader than the protocol of this review.

6.6.6 Key conclusions

The Committee agreed that the guideline should raise awareness of signs and symptoms that could indicate endometriosis and provide guidance for GPs on thresholds for further investigation and diagnosis as well as monitoring and referral. They noted that diagnostic investigation might not be performed by the GP and referral might be necessary.

The Committee agreed that almost all women with symptomatic endometriosis have severe dysmenorrhoea and chronic pelvic pain, but that other symptoms may be more variable. Chronic pelvic pain was defined as a minimum of 6 months of cyclical or continuous pain. They also considered that there should be a distinction between superficial and deep dyspareunia as the latter is more likely to be associated with endometriosis.

The Committee agreed that it is important to ask questions about symptoms and have a full discussion with women, considering the diagnosis of endometriosis when a positive history is given.

The Committee concluded that recommendations should reflect the available evidence and that severe dysmenorrhoea, chronic pelvic pain and a history of infertility (where relevant) were key symptoms associated with endometriosis. Other symptoms were agreed by consensus and the strength of the recommendations for further intervention should reflect this.

6.7 Recommendations

5. **Suspect endometriosis in women (including young women aged 17 and under) presenting with 1 or more of the following symptoms or signs:**
    - chronic pelvic pain
    - period-related pain (dysmenorrhoea) affecting daily activities and quality of life
    - deep pain during or after sexual intercourse
    - period-related or cyclical gastrointestinal symptoms, in particular, painful bowel movements
    - period-related or cyclical urinary symptoms, in particular, blood in the urine or pain passing urine
    - infertility in association with 1 or more of the above

6. **Inform women with suspected or confirmed endometriosis that keeping a pain and symptom diary can aid discussions.**

7. **Offer an abdominal and pelvic examination to women with suspected endometriosis to identify abdominal masses and pelvic signs, such as reduced organ mobility and enlargement, tender nodularity in the posterior vaginal fornix, and visible vaginal endometriotic lesions.**

8. **If a pelvic examination is not appropriate, offer an abdominal examination to exclude abdominal masses.**

ISBN 978-1-4731-2661-9

Referral for women with suspected or confirmed endometriosis

9. **Consider referring women to a gynaecology service for an ultrasound or gynaecology opinion if:**
    - they have severe, persistent or recurrent symptoms of endometriosis
    - they have pelvic signs of endometriosis **or**
    - initial management is not effective, not tolerated or is contraindicated.

10. **Refer women to a specialist endometriosis service (endometriosis centre) if they have suspected or confirmed deep endometriosis involving the bowel, bladder or ureter.**

11. **Consider referring young women (aged 17 and under) with suspected or confirmed endometriosis to a paediatric and adolescent gynaecology service, gynaecology service or specialist endometriosis service (endometriosis centre), depending on local service provision.**

Monitoring for women with confirmed endometriosis

12. **Consider outpatient follow-up (with or without examination and pelvic imaging) for women with confirmed endometriosis, particularly women who choose not to have surgery, if they have:**
    - deep endometriosis involving the bowel, bladder or ureter **or**
    - 1 or more endometrioma that is larger than 3 cm.

ISBN 978-1-4731-2661-9

7 Information and support

Review question: What information and support do women with endometriosis and their families find helpful and what are the barriers and facilitators in the provision of these information and support needs?

7.1 Introduction

The reported average delay of 8 years to a diagnosis of endometriosis means that many women with endometriosis have been told their pain, bleeding, painful sex, fatigue and other symptoms are normal. This can lead to isolation, stress, depression and exhaustion through coping with symptoms without information and support. At the point of diagnosis it has been reported that many women express relief at finally knowing what is wrong.

Accurate, evidence-based, up-to-date and easily accessible information is crucial to support women to understand and self-manage the condition. General information on symptoms and management is of particular importance.

In the clinical setting, specialist nursing staff are a key source of information and support. It is important that the woman understands the consequences of her choices and is able to make an informed decision. The challenge for healthcare professionals is to tailor information to the individual needs, preferences and circumstances of each woman while also allowing for flexibility because information needs may also change with time or if new symptoms develop.

7.2 Description of clinical evidence

The aim of this review was to identify information and support that makes a positive difference to women and their families when diagnosed with endometriosis. The objectives of the review are:

- To test the effectiveness of interventions or package of care to provide additional information and support needs compared to usual care.
- To explore areas of information and support that women and their families find helpful.
- To identify how women would like to receive this information or support.

Qualitative and quantitative studies were selected for inclusion for this review. We looked for studies that collected data using qualitative methods (such as semi-structured interviews, focus groups and surveys with open-ended questions) and analysed data qualitatively (including thematic analysis, framework thematic analysis, content analysis etc.). Survey studies that reported descriptive data that had been analysed quantitatively were excluded. For quantitative studies, we looked for effectiveness of interventions resulting from randomised controlled trials (RCTs), or comparative cohort studies.

For full details, see review protocol in Appendix D, the study selection flow chart in Appendix F, study exclusion list in Appendix H and study evidence tables in Appendix G.

No quantitative studies (RCTs or comparative cohorts) were identified for effectiveness of the following interventions compared with no treatment or usual care:

- support groups
- volunteer groups
- methods of information provision (verbal, written, online, apps, in groups, 1:1 advocacy support
- online health forums.

A total of 17 qualitative studies were identified for inclusion in this review. Of them:

ISBN 978-1-4731-2661-9

12 studies focused on the perspective of women with endometriosis (Ballard 2006, Cox 2003a, Cox 2003b, Denny 2004, 2007, 2009, Gilmour 2008, Jones 2004, Markovic 2008, Seear 2009, Treloar 2007, Whelan 2007). Two studies interviewed both women with endometriosis as well as their partners (Butt 2007, Culley 2013) and 1 study interviewed partners of women with endometriosis (Fernandez 2006). One study was based on blogs from women with endometriosis (Neal and McKenzie 2011). One focused on the perspective of women with endometriosis who use endometriosis online support groups (Shoebotham 2016).

The majority of included studies collected data by semi-structured interviews or focus groups. One study collected data by open ended questions. The most common data analysis method employed across studies was thematic analysis. With regard to the setting of studies:

- Seven studies were conducted in the UK (Ballard 2006; Culley 2013; Denny 2004, 2007, 2009, Jones 2004).
- Five studies were conducted in Australia (Cox 2003, Fernandez 2006, Markovic 2008, Seear 2009, Treloar 2007).
- One study was conducted in the USA (Butt 2007).
- Two studies were conducted in Canada (Neal and McKenzie 2011, Whelan 2007).
- One study was conducted in New Zealand (Gilmour 2008).
- One study was conducted in the UK and the USA (Shoebotham 2016).

Assessment of risk of bias was completed using the Critical Appraisal Skills Programme (CASP) checklist for qualitative studies (see chapter 4). The risk of bias in the included studies ranged from low to high (1 study with low risk of bias; 9 studies with moderate risk of bias; 7 studies with high risk of bias).

Evidence on all themes was considered important by the Committee and was searched for. A number of further themes emerged from the studies and were incorporated in the review. Three systematic reviews were also identified and the majority of individual studies in the reviews were also covered by the search for this review.

A brief description of the studies is provided in Table 17. See also the study selection flow chart in Appendix F, study evidence tables in Appendix G and the exclusion list in Appendix H. For presentation of findings, a theme map was generated according to the themes emerged from studies (Figure 7). Due to the nature of these studies, evidence is summarised in GRADECerqual (Table 18 to Table 23).

7.3 Summary of included studies

A summary of the studies that were included in this review are presented in Table 17.

Table 17: Summary of included studies

Study	Study design/ methods	Participants /respondent	Aim of the study	Comments
Interviews/focus-groups				
Ballard 2006 UK	Semi-structured interviews	N=32 women (including 28 women with subsequent diagnosis of endometriosis)	To obtain women's experience of being diagnosed with endometriosis, delays in diagnosis, treatments	• data collection and analysis clearly reported • researchers' role and potential influences in the analytical process not critically reviewed

ISBN 978-1-4731-2661-9

Study	Study design/ methods	Participants /respondent	Aim of the study	Comments
			available after failure of therapeutic interventions, benefits from diagnosis, diagnosis as an access to social support, absence from work and social obligations	
Butt 2007 USA	Interviews	N=13 comprising women with endometriosis and their partners	To investigate the relationships of couples who are living with chronic pelvic pain from endometriosis	• analysis clearly reported • recruitment of patients was through public and private treatment providers and clinics, as well as endometriosis support and informational groups
Cox 2003a Australia	Focus group	N=61 women contributed to 5 focus groups led by researcher	To determine needs for information related to day surgery for endometriosis-related problems	• 65% response rate from survey contributed to focus group • data collection and analysis clearly reported • researchers' role in the analytical process not critically reviewed
Cox 2003b Australia	Focus groups	N=61 women contributed to 5 focus groups	To determine needs for information related to day surgery for endometriosis-related problems	• 3 of the focus groups were face-to-face and the other 2 were telephone discussions • information from the focus group reported was that of use of complementary therapies
Culley 2013 UK	Face-to-face, semi-structured, in-depth interviews	N=44, comprising 22 women with endometriosis and their partners	To explore the impact of endometriosis on couples	• data collection and analysis clearly reported • researchers' role and potential influences in the analytical process not critically reviewed • self-selected sample
Denny 2003 UK	Semi-structured interviews	N=15 women diagnosed with endometriosis following laparoscopy	To explore women's experiences of living with endometriosis	• participants were approached via a message board on a self-help website, the gynaecological department of a local hospital, or by snowballing • interviews took place either in women's homes or mutually convenient

ISBN 978-1-4731-2661-9

Study	Study design/ methods	Participants /respondent	Aim of the study	Comments
				locations, or over the telephone • thematic and content analysis were carried out using identified key areas and themes were elicited from initial analysis of interview transcripts
Denny 2007 UK	Semi-structured interviews	N=30 women with laparoscopy-confirmed endometriosis	To understand the impact of dyspareunia on women's lives	• women attended endometriosis outpatient clinic • data saturation reported: recruitment to the study was stopped when no new themes emerged from additional data collected
Denny 2009 UK	Semi-structured interviews at baseline and 1 year later	N=30 women with endometriosis outside of the uterus	To explore women's experience of living with endometriosis	• participants were recruited from an endometriosis clinic • data saturation reported: recruitment to the study was stopped when no new themes emerged from additional data • 27/30 women were interviewed after 1 year • storytelling approach was used for collection of data, narrative analysis was considered most appropriate
Fernandez 2006 Australia	Survey /interviews	N=16 male partners of women with endometriosis	To explore experiences of partners of women with endometriosis	• recruitment was achieved via female partner's participation in a previous questionnaire-based study conducted by the authors • response rate was low (32%) • saturation of data collection not reported • not clear which participants were interviewed over the telephone
Gilmour 2008 New Zealand	Unstructured interviews, interactive format	N=18 women recruited through endometriosis support group meeting	To explore women's perceptions of living with endometriosis, its effect on their lives and strategies used to	• data was analysed through a thematic approach • women were aged 16 to 45 years • many of the women were educated at tertiary level and all except 1

ISBN 978-1-4731-2661-9

Study	Study design/ methods	Participants /respondent	Aim of the study	Comments
			manage their disease	participant (16 years age) were or had been in paid employment
Jones 2004 UK	Face-to-face, individual, in-depth interviews	N=24 women with endometriosis diagnosed by laparoscopy	To explore and describe the impact of endometriosis on quality of life	• women attended a gynaecology outpatient clinic • saturation of data reported • interviewer bias was checked by a research nurse who went through the same transcripts as the interviewer
Markovic 2008 Australia	In-depth interviews	N=30 women diagnosed with endometriosis; 6/30 women were menopausal	To understand the relationship between the patients socio-demographic background and health-related phenomena by identifying distinct differences among women's narratives	• women were invited by awareness through community newspapers and noticeboards, snowballing • saturation of data reported
Neal and McKenzie 2011 Canada	Discourse analysis	N=11 blogs authored by women with endometriosis	To understand how bloggers present information sources and make cases for and against the authority of those sources	• saturation of data reported
Seear 2009 Australia	Semi-structured interviews	N=20 women diagnosed with endometriosis	To explore the experiences of women living with endometriosis	• women were recruited by snowballing and also by advertisement of the study being placed in a newsletter of an Australian support group for sufferers, inviting them to contact the author if interested in participating in the study • saturation of data collection reported
Shoebotham 2016 UK and USA	Web-based survey with open-ended questions	N=69 women who were using endometriosis online support groups UK=45 USA=15	To explore the therapeutic affordances of online support group use in women with endometriosis	• only 66 out of the overall study sample (n=69) had a confirmed diagnosis of endometriosis (95%) • data collection and analysis clearly reported • researchers' role in the analytical process critically reviewed

ISBN 978-1-4731-2661-9

Study	Study design/ methods	Participants /respondent	Aim of the study	Comments
Treloar 2007 Australia	Semi-structured interviews	N=21 women with endometriosis recruited from the large GBE study	To investigate motivations and reflections of participants who had provided epidemiological information, blood samples and access to clinical records and data in a large genetic epidemiological study of endometriosis	• participants were contacted individually • women were asked about their motivation to take part in the GBE study • themes were identified from the data according to the direction of questions asked • the researcher took an open-ended approach to interviews. • saturation of data collection not reported
Whelan 2007 Canada	Focus group sessions	N=6 women with endometriosis	To understand how women gather, evaluate and use information about a medical treatment as a specific element of the endometriosis experience	• women were recruited from an endometriosis support group • focus groups involved face-to face group conversations and accounts of endometriosis • the focus of the sessions was GnRH agonists

N: number of participants in study; GBE: Genes Behind Endometriosis

7.4 Clinical evidence

7.4.1 Evidence summary

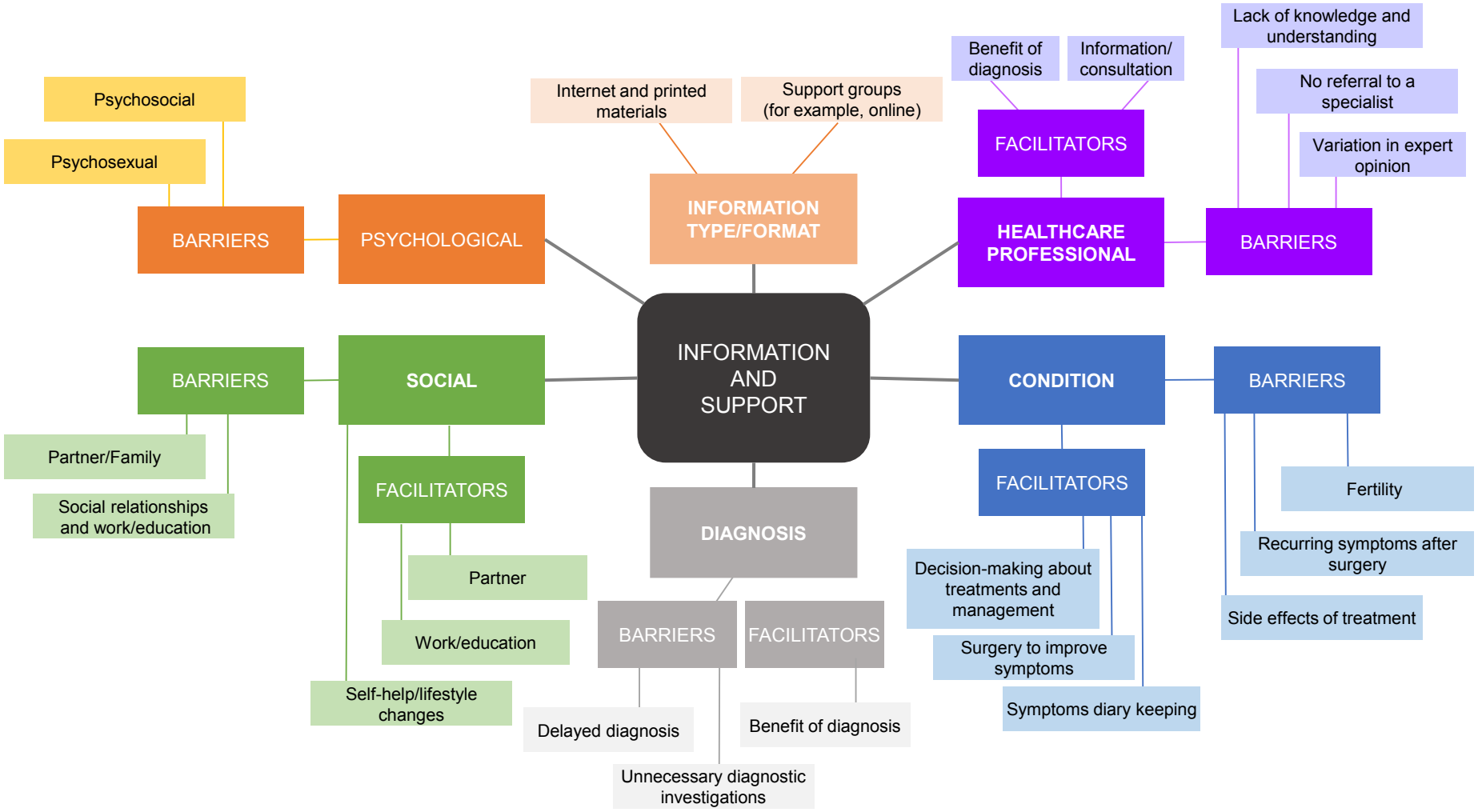
Figure 7 provides a theme map for the qualitative evidence found. At the centre of the map is the main theme that is overarching and which was mentioned as part of most of the other themes and subthemes. Six main overarching themes emerged from interviews or focus groups of women with endometriosis. Themes included facilitators and barriers surrounding diagnosis of endometriosis, issues around interaction with healthcare professionals, how partners were coping with their partners having endometriosis and symptoms affecting their lives and how endometriosis was having psychological effects (psychosocial and psychosexual). Studies also identified that women with endometriosis were accessing different information formats to find out about their condition and that support groups were helpful to understand the condition and have an impact on their decision-making on how to manage treatment. Women also had concerns about how endometriosis would affect fertility and their chances of having children.

Table 18 to Table 23 provide further details on the themes and subthemes found.

ISBN 978-1-4731-2661-9

7.4.2 Clinical evidence profile

Figure 7: Theme map – for description of themes, see Table 18 to Table 23



ISBN 978-1-4731-2661-9

Table 18: Summary of evidence: Theme 1– Information type/content that was perceived helpful to women with endometriosis

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Sub-theme 1:Support groups					
3 (Gilmour 2008, Seear 2009, Whelan 2007)	1 unstructured interview; 1 structured interview; 1 focus group	3 studies conducted in different settings (New Zealand, Australia and Canada) among women with endometriosis reported that support groups were key resources for self-management and exchange of information.	Limitation of evidence	Major limitations	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Sub-theme 1: Online support groups					
1 (Shoebotham 2016)	Web-based survey with open-ended questions	1 multicentre study conducted in 2 countries (the UK and the USA) among women with endometriosis reported 4 therapeutic affordances related to online support group: 1) the ability to connect in order to support each other, exchange advice and to try to overcome feelings of loneliness; 2) the ability to look for information, learn and bolster their knowledge; 3) ability to share their experiences, as well as read about the experiences of others; and 4) the ability to manage how they present themselves online.	Limitation of evidence	Major limitations	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Sub-theme 2: Internet and printed materials					
6 (Gilmour 2008, Markovic 2008, Neal and McKenzie 2011, Seear 2009, Treloar 2007, Whelan 2008)	1 unstructured interview; 1 in-depth interview; 1 discourse analysis; 2 semi-structured interviews; 1 focus group	6 studies conducted in different settings (in New Zealand, Australia and Canada) among women with endometriosis reported that various forms of information (for example, internet and printed materials) were important resources in understanding their condition, treatment options (pros and cons) to help with decision-making.	Limitation of evidence	Major limitations	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	

ISBN 978-1-4731-2661-9

Table 19: Summary of evidence: Theme 2: Psychological barriers encountered by women with endometriosis

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
Sub-theme 1: Psychosocial: not coping with pain, fear of pain starting in public, not socialising, low mood (barrier)					
4 (Culley 2013, Denny 2004, 2007, Jones 2004)	2 semi-structured interviews; 1 face-to-face interview; 1 face-to-face, in-depth interview	4 studies conducted in the UK found that women with endometriosis felt 'worried about the pain starting in public because if the pain occurred most of the women "wanted to be by themselves" and not surrounded by others' or the pain made them 'tired and lacking energy' 'Some women did not want to appear to others that they were not coping' 'Most women described feeling hormonal or had premenstrual tension all the time. They spoke about feeling moody and having short tempers that were often taken out on their friends, family or children' "There's been times in the past where basically she hasn't been up to going out, and I've said 'right well I'm going out anyway because it's the weekend' … I need that time and that space, she knows that. I'm quite a social person." (male partner)	Limitation of evidence	Minor limitations	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Sub-theme 2: Psychosexual: dyspareunia (barrier)					
3 (Denny 2004, 2007, Jones 2004, Culley 2013)	3 semi-structured interviews; 1 face-t - face interview	4 studies conducted in the UK found that women with endometriosis encountered painful intercourse, and would put off due to pain. Some women who went to the GP were told: "'…it's perfectly normal' and [they] suggested that it might be a psychological problem, and I might just be anxious." (female participant). The severity and frequency of dyspareunia and its impact on sex	Limitation of evidence	Minor limitations	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

ISBN 978-1-4731-2661-9

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
		varied, for example for some women this was not perceived as a major problem, however for other women the pain was significant: "When we did get down to it, it was just, it was kind of, it was an ordeal really." (female participant)			
Sub-theme 2: Psychosexual: worries about partners leaving (barrier)					
2 (Culley 2013, Denny 2007)	1 semi-structured interview; 1 face-to-face, in-depth interview	2 studies conducted in the UK among women with endometriosis found that they were worried that their partners would leave due to lack of sexual activity or arguments and tensions in their relationship with their partner: "I do get worried that he's going to go off and meet someone who can give him a lot more than I can." (female participant) "It causes arguments obviously... he doesn't understand that I get frustrated as well but I'd rather just forget about it than go through with the pain I suppose." (female participant) "Coming to terms with not having children of our own and the whole process of IVF, going through it, is really traumatic and for me that's been the most painful element of the whole process." (male partner)	Limitation of evidence Coherence of findings Applicability of evidence Sufficiency or saturation	Minor limitation Coherence Applicable Unclear	Moderate

Table 20: Summary of evidence: Theme 3 – Social facilitators and barriers encountered by women with endometriosis

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
Sub-theme 1: Relationship with partner (facilitator)					
			Limitation of evidence	Major limitations	Low

ISBN 978-1-4731-2661-9

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
3 (Culley 2013, Denny 2004, 2007)	2 semi-structured interviews; 1 face-to-face, in-depth interview	3 studies conducted in the UK among women with endometriosis found that their partners were supportive and 'tended to feel that they were lucky to have such support'.	Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 (Culley 2013)	Face-to-face, in-depth interview	1 study conducted in the UK among women with endometriosis and their partners found that men supported their female partners by providing support in relation to healthcare and treatment (e.g. attending consultations, providing care after surgery), by helping with managing everyday life (e.g. looking after children) or by provided emotional support (by, for example, 'being there', 'listening and understanding').	Limitation of evidence	Major limitations	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Sub-theme 2: Workplace (facilitator)					
1(Denny 2004)	Semi-structured interview	1 study conducted in the UK among women with endometriosis found that their employers were supportive and was 'sympathetic to their needs and made adjustments to their work', stating: "work has been brilliant.'"	Limitation of evidence	Major limitations	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Sub-theme 2: Work/education – school (facilitator)					
1 (Markovic 2008)	In-depth interview	1 study conducted in Australia among women with endometriosis found that some support from teachers ('referred to sick room, given pain killers or hot water bottles') was helpful when they had period pain, although teachers were not experienced in discussion around the severity of the disease.	Limitation of evidence	Minor limitations	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

ISBN 978-1-4731-2661-9

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Sub-theme 3: Self-help/lifestyle changes (facilitator)					
7 (Butt 2007, Cox 2003a, 2003b, Denny 2007, Gilmour 2008, Markovic 2008, Seear 2009)	1 interview; 2 focus groups; 2 semi-structured interviews; 1 unstructured interview; 1 in-depth interview	7 studies conducted in different settings (in Australia, New Zealand, the UK and the USA) among women with endometriosis found that self-help and lifestyle changes (diet and exercise, spiritual healing and positive thinking) helped to 'manage life' (and pain) and be drug free.	Limitation of evidence	Major limitations	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Sub-theme 3: Self-help/lifestyle changes (facilitator)					
7 (Butt 2007, Cox 2003a, 2003b, Denny 2007, Gilmour 2008, Markovic 2008, Seear 2009)	1 interview; 2 focus groups; 2 semi-structured interviews; 1 unstructured interview; 1 in-depth interview	7 studies conducted in different settings (in Australia, New Zealand, the UK and the USA) among women with endometriosis found that self-help and lifestyle changes (diet and exercise, spiritual healing and positive thinking) helped to 'manage life' (and pain) and be drug free.	Limitation of evidence	Major limitations	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Sub-theme 4: Relationship with partner (barrier)					
6 (Butt 2007, Cox 2003a, Culley 2013, Denny 2004, 2007, Fernandez 2006)	1 interview; 1 focus group; 2 semi-structured interviews; 1 survey; 1 face-to-face, in-depth interview	6 studies conducted in Australia, the UK and the USA found that there was some strain on women's relationships with partners and even break-up of some, as men tried and failed to cope with the illness: "I wasn't able to go out and enjoy myself, and I was in a serious relationship for 3 years...and at that point endometriosis made my relationship break up." (female participant)	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Sub-theme 4: Relationship with partner: partner's perspective (barrier)					
			Limitation of evidence	Major limitation	Low

ISBN 978-1-4731-2661-9

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
2 (Culley 2013, Fernandez 2006)	1 survey;1 face-to-face, in-depth interview	2 studies conducted in the UK and Australia among women with endometriosis and their partners found that partners felt 'alarmed and concerned when told about endometriosis and felt shock and disbelief due to the nature and suddenness of surgery'. 'Partners also felt powerless as they saw their partner in pain and did not know what to do to help. Male partners indicated the feeling that they had very limited control over decision-making related to the management of endometriosis, which seemed to induce a sense of powerlessness.' (male partner) "I worry about how she will feel at night and feel helpless to see her in pain, not being able to do anything about it." "Seeing the physical and emotional trauma that my wife had to endure through her years of surgery and procedures was particularly hard to take." (male partner) "I can't feel the pain, I don't even know what a period feels like, whether it's a particularly heavy one or whether it's bad or the period pain beforehand. I don't know what any of that feels like. I can try and put myself in her shoes as best as possible but I will still never understand." (male partner)	Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Sub-theme 5: Social relationships and work/education (barrier)					
			Limitation of evidence	Major limitation	Moderate

ISBN 978-1-4731-2661-9

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
5 (Cox 2003a, Culley 2013, Denny 2004, Gilmour 2008, Jones 2004)	1 unstructured interview; 1 face-to-face interview; 2 semi structured interviews; 1 face-to-face, in-depth interview	5 studies conducted in Australia, New Zealand and the UK among women with endometriosis found that symptoms of endometriosis had caused disruption to education, social relationships and full-time work. Women felt that their social life had 'diminished and the illness forced them to cancel social events'. In the workplace, living with severe pain led to taking sick leave and being unable to perform the job adequately: "I was really worried about having more work, and I had a warning that my sickness rate was unacceptable. My boss was pushing for me to be sent to occupational health because he didn't believe that I was ill." (female participant) "Unfortunately there's a lot of employers out there that just aren't understanding. It's hard to find a good employer … in the end you end up being forced out. It's as simple as that." (female participant)	Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	

Table 21: Summary of evidence: Theme 5 - Healthcare professional

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
Sub-theme 2: Information (facilitator)					
1 (Whelan 2007)	Focus group	1 study conducted in Canada among women with endometriosis considered doctors to be useful	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

ISBN 978-1-4731-2661-9

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
		starting places or authorities on particular subjects: “Surgery, I’d have to say my main source would be my doctor. I read a lot of books and I heard from a lot of people, and I heard all the wrong things. So I got the truth from my doctor.” (female participant)	Sufficiency or saturation	Unclear	
Sub-theme 2: Consultations (facilitator)					
1 (Culley 2013)	Face-to-face, in-depth interview	1 study conducted in the UK among women with endometriosis and their partners recommended (based on their findings) that consultations should be inclusive of the impact of endometriosis on quality of life, and on women, partners and the couple relationship. Healthcare practitioners can improve women’s and couple’s experiences by referring them to specialist services (e.g. pain clinics, psychosexual counselling, etc.). Following diagnosis of endometriosis, healthcare practitioners should raise the topic of planning for and having children, and open up a discussion that allows women and couples to explore this important issue and to receive evidence-based information (also balancing the potential risks of infertility created by the treatments). The same study also suggested the production of a NICE guideline on the management of endometriosis and chronic pelvic pain, incorporation of endometriosis-related information into the training and development of	Limitation of evidence	Major limitations	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

ISBN 978-1-4731-2661-9

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
		health professionals, and adoption of a more couple-focused approach.			
Sub-theme 3: Lack of knowledge and understanding (barrier)					
8 (Ballard 2006, Cox 2003a, Culley 2013, Denny 2004, 2009, Jones 2004, Markovic 2008, Whelan 2007)	3 semi-structured interviews; 1 unstructured interview; 2 face-to-face, in depth interviews; 1 in-depth interview; 1 focus group	8 studies conducted in different settings (in Australia, Canada, New Zealand and the UK) among women with endometriosis found that healthcare professionals were dismissive of women's symptoms, and that pain was due to periods and was 'normal'. "Some physicians would not take endometriosis seriously and knew little about the disease". (female participant) Such issues resulted in women going to see many doctors before they found one who would take them seriously. "This is in a nutshell what is so frustrating about my disease, all the conflicting messages I am receiving, and trying to seek the best possible treatment and dealing with various GPs all the time, just to make me feel like I am always going back to square one. Why can't I go straight to a designated specialist or walk-in clinic?	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	

ISBN 978-1-4731-2661-9

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
		I have a chronic disease that GPs are clearly not knowledgeable about. I am just so frustrated that I do not have access to someone who is able to treat all the aspects of the disease." (female participant)			
Sub-theme 4: Refusal of referral to specialist (barrier)					
1 (Cox 2003a)	Focus group	1 study conducted in Australia among women with endometriosis found that 'doctors refused to give referrals because they simply did not believe in them'. (female participant) One woman in the study suggested to her GP that she may have endometriosis because her mother had the condition, but her GP told her that it was highly unlikely and was reluctant to refer her on to a specialist.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Sub-theme 5: Variation in expert opinion (barrier)					
1 (Whelan 2007)	Focus group	1 study conducted in Canada among women with endometriosis found that there was variation in expert opinion regarding endometriosis	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicability	
			Sufficiency or saturation	Sufficiency	

Table 22: Summary of evidence: Theme 7 – Condition: facilitors and barriers encountered by women with endometriosis

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
Sub-theme 1: Decision-making about treatments and management (facilitator)					
1 (Markovic 2008)	In-depth interview	1 study conducted in Australia among women with endometriosis found that	Limitation of evidence	Minor limitation	Low
			Coherence of findings	Coherent	

ISBN 978-1-4731-2661-9

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
		reclaiming control of women's own health resulted in women becoming their primary decision-maker	Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Sub-theme 2: Surgery to improve symptoms (facilitator)					
2 (Denny 2004, 2009)	2 semi-structured interview	2 studies conducted in the UK found that women who had multiple operations still felt positive about the surgery than medical treatments and did not seem to have the same anxiety about long-term effects of surgery. Symptoms for some women had reduced and did not encounter setbacks such as return of pain	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Sub-theme 3: Symptoms diary keeping (facilitator)					
4 (Cox 2003a, 2003b, Denny 2009, Markovic 2008)	2 focus groups; 1 semi-structured interview; 1 in-depth interview	4 studies conducted in different settings (in Australia and the UK) among women with endometriosis helped to record their symptoms and to 'work out for themselves what was happening and to have a record to show the doctor'	Limitation of evidence	Major limitations	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Sub-theme 4: Worries about fertility (barrier)					
5 (Butt 2007, Culley 2013, Denny 2009, Jones 2004, Markovic 2008)	1 interview; 1 semi-structured	5 studies conducted in different settings (in Australia, the UK and the USA) among women with	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

ISBN 978-1-4731-2661-9

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
	interview; 1 face to face interview; 2 in-depth interviews	endometriosis found that women were concerned about their 'uncertainty about fertility as a result of being diagnosed with endometriosis, and difficulties in deciding which pathways to parenthood would be most appropriate'. (female participant) "It's a complete nightmare to realise that you're not able to have children and you still have to keep trying. There is this pressure on you to keep trying, you kind of feel that it might not work. It's heart-breaking, it's been very hard. So yes, we've had some very low points. It's just yes, very, very stressful." (female participant) Data from 1 study conducted in the UK strongly suggested that either actual or anticipated infertility was a significant issue for the vast majority of the study participants: "The biggest concern for me is will I be able to have children? So I'm very emotional about my period and the pain every month. So it's kind of slipped into another dimension now … I accept it, I'm ok that I've got endometriosis, but now I'm worried about the impact it's going to have." (female participant)	Sufficiency or saturation	Unclear	
Sub-theme 5: Recurring symptoms after surgery (barrier)					
1 (Denny 2004)	Semi Structured interview	1 study conducted in the UK found that majority of women with endometriosis were worried about symptoms returning: "I'm scared that it will come back…I know there's a huge chance of it	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

ISBN 978-1-4731-2661-9

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
		coming back at some point and that really does scare me. I could almost say that I don't think that I could ever go through that again. I'm not sure I could cope with that a second time round."			
Sub-theme 3: Worry about side effects of medical treatment (barrier)					
1 (Whelan 2007)	Focus group	1 study conducted in Canada among women with endometriosis found that women argued that 'endometriosis treatments –specifically GnRH agonists –may cause depression, irritability, confusion, anxiety and memory loss'	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Table 23: Summary of evidence: Theme 9: Diagnosis: facilitators and barriers encountered by women with endometriosis

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
Sub-theme 1: Benefit of diagnosis (facilitator)					
5 (Ballard 2006, Cox 2003a, 2003b, Culley 2013, Denny 2009)	2 semi-structured interviews; 2 focus groups; 1 face-to-face, in-depth interview	5 studies conducted in different settings (in Australia and the UK) among women with endometriosis found that after considerable length of time, an eventual diagnosis was described as a 'relief' and a 'burden lifted from the minds of the women'.	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Sub-theme 2: Delayed diagnosis of endometriosis (barrier)					
8 (Ballard 2006, Cox 2003a, 2003b, Culley 2013, Denny 2004, 2009, Markovic 2008, Whelan 2007)	3 semi-structured interviews; 3 focus groups; 1 in-depth interview; 1	8 studies conducted in different settings (in Australia, Canada and the UK) found that women experienced delay in diagnosis and women experiencing symptoms of pain in their teen years were told by their GPs that their symptoms were normal and that	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	

ISBN 978-1-4731-2661-9

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
	face-to-face, in-depth interview	they were 'unlucky' that they suffer from 'bad periods'.			
Sub-theme 3: Unnecessary diagnostic investigations (barrier)					
4 (Ballard 2006, Cox 2003b, Culley 2013, Markovic 2008)	1 semi-structured interview; 1 focus group; 1 in-depth interview; 1 face-to-face, in-depth interview	4 studies conducted in different settings (in Australia and the UK) found that women experienced unnecessary investigations before they were diagnosed with endometriosis. 'It was awful going for internals all the time and being told there's nothing there. To keep going backwards and forwards and having it, and then there's nothing showing up.' (female participant)	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

ISBN 978-1-4731-2661-9

7.5 Economic evidence

This question focused on the content and quality of information that is routinely provided rather than whether the provision of information represented – in itself – a cost-effective use of resources. This question is not primarily about competing alternative uses of resources with different opportunity cost and therefore was not considered suitable for health economic review. A global health economic search was undertaken which did not find any evidence relating to information and support.

7.6 Clinical evidence statements

A number of themes emerged from the semi-structured interviews, interviews, focus groups and support groups of women with endometriosis and also their partners. The central theme of information content with subthemes of information type, social, healthcare professional, diagnosis, condition and psychological information are interlinked and have been perceived as important and helpful or as barriers by women with endometriosis and their partners and families.

7.6.1 Information type

Moderate quality evidence from 3 studies, carried out among women with endometriosis using interview or focus group design, showed that women found support groups to be key resources for self-management of their condition and also exchange of information with other endometriosis sufferers.

Low quality evidence from 1 study, carried out among women with endometriosis using a web-based survey with open-ended questions, showed that endometriosis online support groups provide 4 therapeutic opportunities to connect in order to support each other, exchange advice and to try to overcome feelings of loneliness; the ability to look for information; the ability to share their experiences, as well as read about the experiences of others.

Moderate quality evidence from 6 studies, carried out among women with endometriosis using interview, focus group or discourse analysis study design, showed that women found various forms of information including books, Internet, internet chat rooms, blogs, newspapers, guest speakers, recorded material, medical publications, leaflets, drug reference manual and being part of a research study to be important resources to understand their condition and treatment options (benefits and harms) to help them with decision making.

7.6.2 Psychological barriers

Moderate quality evidence from 4 studies showed that women with endometriosis did not want to appear that they were not coping with symptoms in front of others. Some women were concerned about their symptoms (mainly pain) starting in public as they would rather be by themselves and not surrounded by others. Symptoms of pain also made them tired and lacking energy. Some women also spoke about their mood and that they had short tempers that were often taken out on family, friends or children.

Moderate quality evidence from 4 studies showed that women with endometriosis encountered dyspareunia, which was disruptive to their wellbeing. Moderate quality evidence showed that women were concerned that their partners would leave them due to dyspareunia.

ISBN 978-1-4731-2661-9

7.6.3 Social facilitators

Moderate to low quality evidence from 4 studies carried out among women with endometriosis using semi-structured interview and in-depth interview study design independently showed that partner, workplace and teachers (at school) were supportive and helpful for women managing their condition.

Low quality evidence from 1 study carried out among women with endometriosis and their partners using semi-structured interviews found that men supported their female partners by providing support in relation to healthcare and treatment (e.g. attending consultations, providing care after surgery), by helping with managing everyday life (e.g. looking after children) or by provided emotional support (as for example 'being there', 'listening and understanding').

Moderate-quality evidence from 7 studies carried out among women with endometriosis using semi- or unstructured interview and focus group study design also showed that self-help and making lifestyle changes (e.g., diet/nutrition, exercise, spiritual healing and positive thinking) helped to manage life and pain and be drug free.

7.6.4 Social barriers

Moderate quality evidence from 6 studies showed that partners of women with endometriosis found it difficult to cope with the condition and it put strain on their relationships, resulting in relationships breaking up.

Low quality evidence from 2 studies from a partner's perspective, showed that they were concerned when told about endometriosis and that they felt powerless as they did not know what to do to help. In addition, women encountered disruption to relationships with family (one low quality study, focus group) especially as women were convinced by their mothers, aunts, teachers or others that symptoms of period pain were normal.

Moderate quality evidence from 5 studies (using semi- or unstructured interview and face-to-face interview study design) showed that women encountered disruption in their education, social relationships and full-time work as employers did not believe their illness.

7.6.5 Healthcare professionals as facilitators

Moderate quality evidence from 1 study carried out among women with endometriosis using focus group study design, indicated that the healthcare professional was a starting point for women requiring information about their condition and treatments.

Low quality evidence from 1 study carried out among women with endometriosis and their partners found that consultations should be inclusive of the impact of endometriosis on quality of life and on women, partners and the couple's relationship. Moreover, healthcare professionals should raise the topic of planning for and having children and discuss this important issue openly.

7.6.6 Healthcare professionals as barriers

Moderate quality evidence from 8 studies carried out among women with endometriosis with semi- and unstructured interview, face-to-face interview and focus group study designs, found that women's symptoms were trivialised by the doctor and felt that the doctors knew little about the disease.

Low quality evidence from 1 study carried out among women with endometriosis and their partners found that healthcare professionals can improve women's and couple's experiences by referring them to specialist services (e.g. pain clinics, psychosexual counselling, etc.).

ISBN 978-1-4731-2661-9

Moderate quality evidence from 1 study showed that women encountered variation in expert opinion of doctors about endometriosis and led to women going to a number of doctors before being seen by one who would take their symptoms seriously. Low quality evidence from 1 study showed that doctors refused to refer women to specialists because they did not believe them.

Low quality evidence from 1 study carried out among women with endometriosis and their partners found that healthcare professionals should raise the topic of planning for and having children and open up a discussion that allows women and couples to explore this important issue and to receive evidence-based information (also balancing the potential risks of infertility created by the treatments).

7.6.7 Condition facilitators

Low quality evidence from 1 study carried out among women with endometriosis using interview study design found that acquiring knowledge about endometriosis allowed them to reclaim control of their own health, resulting in them becoming their primary decision-maker for further treatments.

Low quality evidence from 2 studies carried out among women with endometriosis using semi-structured interview study design, indicated that symptoms for some women had reduced after surgery and those women who had multiple operations were still positive about the surgical treatment reducing symptoms compared with medical treatment and did not have the same anxiety about long-term effects of surgery.

Moderate quality evidence from 4 studies, carried out among women with endometriosis using focus group and structured interview study design, showed that recording symptoms helped to understand and manage symptoms better and to have a record to show to their GP.

7.6.8 Condition barriers

Moderate quality evidence from 5 studies carried out among women with endometriosis found that women were concerned about their chances of conception and the uncertainty of actual or anticipated infertility. They also encountered difficulties in deciding which pathways to parenthood would be appropriate for them.

Low quality evidence from 1 study carried out among women with endometriosis found that women were concerned about recurring symptoms of pain.

Moderate quality evidence from 1 study also showed that women with endometriosis were concerned about side effects of medical treatments such as GnRH agonists.

7.6.9 Diagnosis facilitators

Moderate quality evidence from 5 studies carried out among women with endometriosis using semi-structured interview and focus group study design, found that a diagnosis of endometriosis was beneficial to women as it allowed them to have discussion about their condition with their doctor.

7.6.10 Diagnosis barriers

Moderate quality evidence from 8 studies carried out among women with endometriosis with semi-structured interview and focus group study design found that women experienced delay in diagnosis for a number of years that was related to doctors not taking their symptoms seriously or doctors were dismissive of women's symptoms of pain. In addition, moderate

ISBN 978-1-4731-2661-9

quality evidence from 4 of these studies found that women experienced unnecessary diagnostic investigations before they were diagnosed with endometriosis.

7.7 Evidence to recommendations

7.7.1 Relative value placed on the outcomes considered

The Committee agreed that the frustrations related to delays in diagnosis was a critical theme and that social support and the psychological impact of endometriosis were also important themes identified by the review. Themes relating to the perspective and involvement of partners of women with endometriosis were also considered important in drafting the recommendations (such as emotional support needs and participation in decision-making). The Committee also acknowledged that the principles set out in the Patient Experience Guideline (CG138) regarding the presentation of information in a personalised manner was important for women with endometriosis.

7.7.2 Consideration of clinical benefits and harms

The evidence included in this review showed that women with endometriosis found information and support provided through all forms to be helpful, for example, support groups, written or online, face to face, and this information enabled them to be actively involved in decision-making for the management and treatment of endometriosis. However, the evidence also identified barriers that women and their support network faced in their endometriosis pathway. The Committee made recommendations on general information and support, as well as specific guidance on essential information for a woman with suspected or confirmed endometriosis.

7.7.3 Consideration of economic benefits and harms

Providing information and advice is part of routine clinical practice. It typically involves a small opportunity cost in terms of staff time or consumables (e.g. patient information leaflets). There is significant potential gain if a better understanding of symptoms and treatment options leads women to pursue fewer healthcare provider contacts, request more appropriate treatment and are able to function better in activities of daily living (for example, work and social interaction).

Some recommendations for increased information and support could carry a direct economic cost; patient information leaflets might have to be expanded and reprinted as booklets, and clinical time would have to be found to discuss the issues arising. However, in actuality, the Committee did not think that their recommendations were likely to cause this cost (and the effect would be very marginal even if the recommendations did), so it is very unlikely that these recommendations will have a significant resource impact on the NHS as some information is already disseminated and the recommendations are limited to ensuring this information is useful.

7.7.4 Quality of evidence

Moderate to low quality evidence was presented in the review. The main limitations of the evidence base were:

- Lack of saturation in the data analysis and data collection.
- Lack of critical review of the researcher's role in sample recruitment, data collection or data analysis process. Very few studies clearly reported the relationship between researchers, interviewers and the respondents, whether the researchers had a pre-understanding about the topic or the possible influence of that in data collection and analytical process.

ISBN 978-1-4731-2661-9

- Lack of verification of findings: very few studies verified their findings with participants or external sources. They also did not report the reason why verification was not necessary or applicable. Some studies did not report in detail how findings/themes were derived or emerged from the data in their research.
- A number of facilitating themes were consistently reported by many women but, due to the limitation in studies, the evidence should be interpreted with caution.

7.7.5 Other considerations

The Committee suggested that reference could be made to the Patient Experience Guideline (CG138) as it covered appropriate and timely manner of information provision, but information must be specific to the women's age in terms of:

- signs and symptoms
- fertility
- information on support groups
- treatment options such as self-management
- medical services that were available.

In terms of service provision, the Committee considered that endometriosis should be acknowledged as a long term chronic condition in order to ensure long-term support as necessary for women according to the severity of their condition and the complexity of their needs. The Committee also stressed the importance of acknowledging the cultural background of women affected by it. The Committee was aware that there was a report by Denny (2010) highlighting that services for minority ethnic women with endometriosis could be improved. It was therefore agreed that this was an important point to highlight in the recommendation.

The Committee believed that assessment of information needs should include women as well as their 'support network'. Women may rely on their friends or on support groups for women who have endometriosis rather than worry their family or partner. Endometriosis could impact many aspects of women's lives in terms of pain, fertility and activities of daily living. It may also have a significant impact on the emotional, psychosocial and psychosexual wellbeing of women and their partners and families, therefore the Committee agreed that this should be addressed by health professionals to support women according to their specific needs. There is also potential for a significant impact on finances as the workplace may not be sympathetic towards women with the condition, owing to lack of awareness.

It was acknowledged by the Committee that there is a significant delay in diagnosis and thus training and education for healthcare professionals should be identified. For most women, the healthcare professional is the first point of contact regarding information about their condition and the Committee highlighted that this is often insufficiently provided, e.g. information about local support groups. The Committee suggested that GPs should be willing to discuss with women and/or their support network, the sensitive aspects of endometriosis, specific to the need of the patient. In addition, it was raised that women and their partners and family may also want their healthcare professional to know that delayed diagnosis can have physical, emotional, social and psychosexual impact and also impact on fertility. The Committee were aware of organisations that provide diary templates for recording symptoms and noted that keeping a diary could help diagnose the condition sooner. It was recommended that this could consist of a list of the symptoms (severity, type and location) that they experience, which can then be shared with their healthcare professional. The Committee also noted that specialist nurses would be well placed to have discussions with and provide support to women and their families through their pathway.

The Committee decided no additional recommendations were necessary for adolescents, but chose to make a research recommendation.

ISBN 978-1-4731-2661-9

7.7.6 Key conclusions

Although some of the themes/findings that emerged from the review are covered by the Patient Experience Guideline (CG138), the Committee considered specific recommendations to address the information and support needs of women with endometriosis and their partners and families. These were based on the themes identified in the current review.

7.8 Recommendations

13. **Be aware that endometriosis can be a long-term condition, and can have a significant physical, sexual, psychological and social impact. Women may have complex needs and require long-term support.**

14. **Assess the individual information and support needs of women with suspected or confirmed endometriosis, taking into account their circumstances, symptoms, priorities, desire for fertility, aspects of daily living, work and study, cultural background, and their physical, psychosexual and emotional needs.**

15. **Provide information and support for women with suspected or confirmed endometriosis, which should include:**
    - what endometriosis is
    - endometriosis symptoms and signs
    - how endometriosis is diagnosed
    - treatment options
    - local support groups, online forums and national charities, and how to access them.

16. **If women agree, involve their partner (and/or other family members or people important to them) and include them in discussions. For more guidance on providing information to people and involving family members and carers, see the NICE guideline on patient experience in adult NHS services.**

7.9 Research recommendations

1. **What information and support interventions are effective to help women with endometriosis deal with their symptoms and improve their quality of lives?**

Why this is important?

This guideline has identified that women with endometriosis and their partners feel that information and support is not always provided in the way that best meet their needs. However, the direct effectiveness of different types or formats of information and support interventions on measurable outcomes such as health-related quality of life and level of function (for example, activities of daily living) have not been tested. Good practice in this area in non-specialist and specialist settings can improve satisfaction with the care provided. It may also improve quality of life and positively affect relationships between healthcare professionals and the woman with endometriosis, as well as the woman's personal family relationships.

ISBN 978-1-4731-2661-9

Table 24: Research recommendation rationale

Research question	What information and support interventions are effective to help women with endometriosis deal with their symptoms and improve their quality of lives?
Importance to 'patients' or the population	Diagnosis and early treatment of the disease is of prime importance to the health of women. There is a potential that the consequences from delayed diagnoses may affect daily activities of living (work/relationships/sexual function/fertility) impacting on their mental health.
Relevance to NICE guidance	High Priority: Minimising known risk factors such as delayed treatment affecting fertility
Relevance to the NHS	Very large, lack of an effective treatment pathway and crucial support through all agencies may in turn have severe consequences not only for the women but also for her family/employer/NHS in regards to the women's coping mechanisms. Importantly this lack of quality care impacts directly on the cost to the NHS by repeated attendance to the GP/A&E/emergency services/emergency admissions.
National priorities	None identified
Current evidence base	Poor and inconsistent
Equality	Risks for women restricted within their working environment, travelling in pain with heavy bleeding and consequently loss of employment due to these factors. Delayed diagnosis prevents some women, nearing 40yrs of age from access to fertility services

Table 25: Research recommendation PICO table

Criterion	Explanation
Population	Women from 16yrs onwards
Intervention	Data collection: EHP-30 (the main research questionnaire for endometriosis) looks at mental health – the 5 core components are: pain, control and powerlessness, social support, emotional well-being and self-image
Prognostic or risk factor	Women with symptoms of endometriosis
Comparator (without the risk factor)	• Women with asymptomatic endometriosis • Women without endometriosis
Outcome	• Health related quality of life • Mental wellbeing • Activities of daily living
Study design	A prospective multi-centre study collecting prospective community (GP) service and hospital data.
Timeframe	2 years

ISBN 978-1-4731-2661-9

8 Risk of cancer of the reproductive organs

Review question: Do women with endometriosis have an increased risk of cancer of thereproductive organs and do they need to be monitored or referred accordingly?

8.1 Introduction

Cancer of reproductive organs (cancer of the cervix, body of uterus, Fallopian tubes and ovaries) is an important cause of death and morbidity for women in the UK. A national screening programme is established for cervical cancer. Previous studies have suggested an association between endometriosis and cancer of the ovary. However, other factors may also play a part, for example, endometriosis is associated with infertility and women who remain childless are recognised to have an increased risk of ovarian cancer. This can be a source of anxiety for women with endometriosis and their health professionals may be uncertain as to whether additional surveillance (for example, with pelvic ultrasound) should be offered to women with endometriosis.

The objective of this systematic review is:

- To determine whether there is an increased risk of cancer of reproductive organs (i.e. incidence) in women with endometriosis compared with those without endometriosis.
- The following amendments were made to the initial protocol (for full details, see review protocol in Appendix D):
  - The risk of developing cancer of reproductive organs was reviewed to enable a comparison to be made. Prevalence figures were not often reported (or only in cross-sectional studies that were excluded from the protocol). The majority of studies reported the incidence of cancer of reproductive organs
  - Studies were included but downgraded if they did not adjust for the Committee-specified confounders (severity of endometriosis, age, subfertility, family history and smoking). This is due to no studies adjusting for all of these confounders.
  - Case control studies were excluded as there were sufficient retrospective cohort studies found that met the inclusion criteria.

See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I and study evidence tables in Appendix G.

8.2 Description of clinical evidence

Seventeen papers reporting 15 studies (Aris 2010, Buis 2013, Brinton 1997, 2004, 2005a, 2005, Chang 2014, Kobayashi 2007/2008, Kok 2015, Lee 2015, Melin 2006, 2007, Mogensen 2016, Stewart 2013, Wang 2014 and Yu 2015) were included in this clinical review. Three systematic reviews (Kim 2014, Heidemann 2014 and Zafrakas 2014) were used to cross check the studies, but were not included in the review due to them having different inclusion criteria in their protocols. Evidence from these studies is summarised in the clinical evidence tables below (Table 26 to Table 32). See also the study selection flow chart in Appendix B, forest plots in Appendix D, study evidence tables in Appendix E and exclusion list in Appendix G.

All of the included studies are retrospective cohort studies that reported either a standardised incidence ratio (SIR), hazard ratio (HR) or relative risk (RR) of the following cancers of reproductive organs:

- cervical cancer (3 studies)
- carcinoma in situ (CIS) of the cervix (1 study)
- endometrial cancer (6 studies)

ISBN 978-1-4731-2661-9

- ovarian cancer (14 studies)
- borderline ovarian tumour (2 studies)
- fallopian tube cancer (1 study)
- uterus not otherwise specified/ uterine cancer (4 studies).

The results of the studies were not pooled together due to the following differences in the studies:

- Population: geographical location (Canada, the Netherlands, Sweden, Denmark, Taiwan, the USA, Japan, western Australia), age group inclusion criteria.
- Comparison groups: population-wide comparison, matched (age, calendar year) population controls, subfertile population.
- Diagnosis: single or combination of methods (questionnaire, medical records, database or registry with coding for outpatient appointments, inpatient stays, surgery and histology).
- Prevalent and incident figures of cancers of reproductive organs.
- Adjustment for different confounders, including those specified as major confounders by the Committee:
  - subfertility (pregnancy/parity/gravidity: 5 studies; infertility: 4 studies)
  - family history (1 study)
  - hormonal treatment use (oral contraceptives: 2 studies, in vitro fertilisation (IVF): 2 studies, other: 1 study)
  - no studies adjusted for the severity of endometriosis or smoking
  - all of the studies adjusted for age.
  - No meta-analysis was performed due to differences in the studies such as statistical analysis, confounder adjustment and comparison group populations.

8.2.1 Summary of included studies

A summary of the studies that were included in this review are presented in the evidence profiles (Table 26 to Table 32).

The studies used a variety of data sources, the details of which are summarised below:

1. National registries or databases:
   a. Taiwan National Health Insurance Research Database (NHIRD): computerised databases that have file/original data on claims reimbursements from the national insurance system. It was started in 1995 with >99% residents enrolled in the programme. From December 2010, >99% of the population was covered and almost all of the medical hospitals/clinics in Taiwan were included. The Longitudinal Health Insurance Database used in the included studies contains 1 million randomly selected individuals (4.5% of Taiwanese population) with anonymised data. Data includes details of medical orders, procedures and medical diagnoses based on International Classification of Diseases (ICD), 9th edition. This was used in the following studies: Chang 2014, Kok 2015, Lee 2015, Wang 2014 and Yu 2015.
   b. Registry for 'Catastrophic Illness Patients and National Cancer Registration system' (Taiwan): Chang 2014, Lee 2015 and Wang 2014.
   c. Danish Cancer Register, Hospital Discharge Register (Denmark): Brinton 2005.
   d. Danish National Patient Register: Mogensen 2016.
   e. National Swedish Registry, National Causes of Death Registry (Sweden) and National Swedish Cancer Registry: Brinton 1997, Melin 2006, 2007.
   f. Swedish Multi-gGeneration Register: Melin 2007.
   g. National Death Index, Cancer Registries of the US: Brinton 2004, 2005a.
   h. Shizuoka Cancer Registry (Japan): Kobayashi 2008a, 2008b.

ISBN 978-1-4731-2661-9

i. Whole population linked hospital and registry data (Western Australia data linkage system), Hospital Morbidity Data System, Reproductive Technology Register: Stewart 2013.
j. PALGA (Dutch Pathology Registry): This contains all the cytological and histological diagnoses made in the Netherlands. Individual pathology laboratories submit their data. There has been nationwide coverage since 1989: Buis 2013.
k. Netherlands Cancer Registry: Buis2013.

2. Questionnaires: Kobayashi 2008/2008, Buis 2013, Brinton 2004, 2005a.
3. Medical records: Aris 2010, Buis 2013, Brinton 2004, 2005a:
   a. CIRESSS: Centre Informatise de Recherche Evaluative en Services et Soins de Sante: patient clinical and pathological records system (ICD coding) that covers all the residents of the Estrie region of Quebec, Canada (300,383 residents): Aris 2010.

8.2.2 Clinical evidence profile

The clinical evidence profiles for this review question (risk of cancers of reproductive organs) are presented in Table 26 to Table 32.

ISBN 978-1-4731-2661-9

8.2.3 Summary tables of cancer risk

Table 26: Cervical cancer

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Swedish National Register							
Brinton 1997 RC, SNR 1969–1983 All endometriosis patients included No age restriction (range 12–82 years)	N=20,686 endometriosis	11	15.24 Population	11.4 (range 1 to 21)	SIR: 0.72 (0.4 to 1.3)	Very high	• Only 1 ICD code for endometriosis diagnosis • 55.6% truncated follow-up due to uncertainty of whether 1/both ovaries remained e.g. after hysterectomy or oophorectomy • Population likely to have missed some cases (non-hospital admissions, milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60%–85% coverage) • Post-hoc analysis by site of endometriosis (data not shown as also n values were not given for denominators) • Wide CIs
Melin 2006 RC, SNR 1969–2000 New diagnoses of endometriosis No age restriction (average 39.4 (SD 10.4) at entry	N=64,492 endometriosis	51	80.18 Population	12.7 (528,441 person years**)	SIR:0.64 (0.47 to 0.84)	High	• Population likely to have missed some cases (non-hospital admissions, milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60%–85% coverage)

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Melin 2007 RC, SNR1969–2002 New diagnoses of endometriosis No age restriction (average 39.5 (SD 10.5) at entry	N=63,630 endometriosis	49	Not given	13.4 (792,013 person years**)	SIR 0.71 (0.53 to 0.94)	High	• Population likely to have missed some cases (non-hospital admissions) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60%–85% coverage) •

(a) RC: retrospective cohort; N: number of participants in study; SNR: Swedish National Registry; RR: relative risk ratio; HR: hazard ratio; SIR: standardised incidence ratio; ICD:
(a) International Classification of
(b) Diseases; CI: confidence interval; SD: standard deviation
*(a) *Observed: The number of cancer cases in the study sample; Expected: The estimated number of cancer cases in the sample if the same was from a population without*
(b) Endometriosis (calculated using prevalence from a registry or external source.
*(c) **Person-years: the total number of years at risk across all participants (number of participants x years of follow-up). This accounts for different lengths of follow-up among*
***(d)** Different individuals.*

Table 27: Cancer in situ of the cervix

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Melin 2006 RC, SNR 1969–2000 New diagnoses of endometriosis No age restriction (average 39.4 (SD 10.4) at entry	N= 64,492 endometriosis	523	584.5	12.7 years (508,447 person years**)	SIR 0.89 (0.82 to 0.97)	High	• Population likely to have missed some cases (non-hospital admissions; milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage)

(a) RC: retrospective cohort; N: number of participants in study; SNR: Swedish National Registry; RR: relative risk ratio; HR: hazard ratio; SIR: standardised incidence ratio; SD:
(b) Standard deviation.
*(c) *Observed: The number of cancer cases in the study sample; Expected: The estimated number of cancer cases in the sample if the same was from a population without*
(d) Endometriosis (calculated using prevalence from a registry or external source.

ISBN 978-1-4731-2661-9

*(e) **Person-years: the total number of years at risk across all participants (number of participants x years of follow-up). This accounts for different lengths of follow-up among*
(f) *Different individuals.*

Table 28: Endometrial cancer

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/Control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Swedish National Register							
Brinton 1997 RC, SNR 1969–1983 All endometriosis patients included No age restriction (range 12–82 years)	N=20,686 endometriosis	12	10.97 Population	11.4 (range 1 to 21)	SIR 1.09 (0.6 to 1.9)	Very high	• Only 1 ICD code for endometriosis diagnosis • 55.6% truncated follow-up due to uncertainty of whether 1/both ovaries remained, e.g. after hysterectomy or oophorectomy • Population likely to have missed some cases (non-hospital admissions, milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage) • Post-hoc analysis by site of endometriosis (data not shown as also n values were not given for denominators)
Melin 2006 RC, SNR 1969–2000 New diagnoses of endometriosis No age restriction (average 39.4 (SD 10.4) at entry	N= 64,492 endometriosis	92	77.37	12.7 years (427,114 person years**)	SIR 1.19 (0.96 to 1.46)	High	• Population likely to have missed some cases (non-hospital admissions, milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage)

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/Control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Melin 2007 RC, SNR1969–2002 New diagnoses of endometriosis No age restriction (average 39.5 (SD 10.5) at entry	N=63,630 endometriosis	97	Not given	13.4 (792,013 person years**)	SIR 1.14 (0.93 to 1.39)	High	• Population likely to have missed some cases (non-hospital admissions) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage)
National Health Insurance Research Database (NHIRD) of Taiwan							
Kok 2015, RC, NHIRD, Claims between 2002 and2005 followed up until 31 December, 2008 Newly diagnosed endometriosis Age: >20 years	N=2,266 endometriosis N=9,064 comparison group (age and index matched 1:4)	12	5	Mean NR 9,842 person years** in the endometriosis cohort, 36,274 in the comparison cohort	HR 4.05 (1.20 to 13.66) Ovarian endometriosis HR 3.23 (0.54 to 19.27) Pure ovarian endometriosis – none	Very high Very high	• Adjusted for age, diabetes mellitus, chronic kidney disease, liver cirrhosis, rheumatoid arthritis, use of medroxyprogesterone acetate, norethindrone acetate, danazol and GnRH agonist • Unclear drop out/lost to follow-up but patients were censored at this point • Only age was controlled for out of the major confounders listed by the Committee • Women <3 times evaluated or for a follow-up period of <2 months were excluded (potentially milder cases excluded) • Post-hoc analysis by site of endometriosis • No censoring for women who had a hysterectomy etc. after their index date • Very wide Cis
Yu 2015, RC, NHIRD Jan	N=15,488 endometriosis	104 (0.7%)	288 (0.2%)	10 year follow-up	HR 2.83	High	• Adjusted for age, urbanisation level, monthly income, geographic

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/Control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
1997–Dec 2000 with 10 year follow-up Unclear if just new or includes old diagnoses of endometriosis prior to study start date Age: no inclusion criteria described	N=123,904 age/sex matched controls				(1.49 to 5.35)		region, hypertension, hyperlipidaemia, obesity and diabetes mellitus • No description of any exclusions for women with hysterectomy etc. • Unclear drop out/lost to follow-up/ no description of censoring • Women <2outpt apt within initial year of endometriosis diagnosis given by a gynaecologist were excluded (potentially milder cases excluded) • Only age was controlled for out of the major confounders listed by the Committee • Wide Cis
• The Danish National Patient Register							
Mogensen 2016, RC, Danish National Patient Register, 1977–2012, mean follow-up 4.1 years Women with a diagnosis of endometriosis Unclear if just new or includes old diagnoses Age: no age restriction. Mean age at diagnosis 59 years	N=43,784 endometriosis	118	55.34	Median: 4.1 years	SIR 2.13 (1.77 to 2.55) Subgroup analysis by age at first endometriosis (years) <30: SIR 0.62 (0.17 to 1.59) 30–39: SIR 1.81 (1.26 to 2.53) 40–49: SIR 1.23 (0.80 to 1.80)	Very high	• Study was limited to only women who were hospitalised for endometriosis • Only age was considered as a confounding factor • Very limited baseline characteristics were provided in the paper

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/Control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
					≥50: SIR 1.75 (0.93 to 2.99)		

(a) RC: retrospective cohort; N: number of participants in study; SNR: Swedish National Registry; NHIRD: National Health Insurance Research Database; SIR: standardised
(b) Incidence ratio; RR: relative risk ratio, HR: hazard ratio; ICD: International Classification of Diseases; CI: confidence interval; SD: standard deviation.
*(c) *Observed: The number of cancer cases in the study sample; Expected: The estimated number of cancer cases in the sample if the same was from a population without*
(d) endometriosis (calculated using prevalence from a registry or external source.
*(e) **Person-years: the total number of years at risk across all participants (number of participants x years of follow-up). This accounts for different lengths of follow-up among*
(f) different individuals.

Table 29: Ovarian cancer

Study, design and data source	Number of participants (N)	Observed/case group* (cancer in women with endometriosis)	Expected/control group* (cancer in women without endometriosis)	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Aris 2010 RC, Estrie region of Quebec CIRESSS database 1997–2006 All endometriosis diagnoses No age restriction	N=2,521 endometriosis	41	251	NR 9 year study length	RR 1.6 (1.12 to 2.09)	Very high	• Adjusted for age, no pregnancies, family history of ovarian cancer, family origin, OC use, tubal ligation, hysterectomy and breastfeeding • Unpublished n value for the comparison group • No baseline characteristics apart from age were given • Only age and family history of ovarian cancer were controlled for out of the major confounders listed by the Committee • Imprecise (CIs cross upper MID)
Brinton 2004/2005A, RC, USA, 1965-1988, 5 reproductive	N=1,919 endometriosis	13	5.2 expected (US population,	Median- 35,196 person years**	SIR 2.48 1.3 to 4.2	Very high	• 20% lost to follow-up • Excluded 1st year follow-up data

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group* (cancer in women with endometriosis)	Expected/control group* (cancer in women without endometriosis)	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
endocrinology centres, cancer registries, National Death Index, questionnaires, 1965–1988 Endometriosis included in selection criteria (seen >1/referred by physician for infertility advice) No age restriction	N=6,510 infertility population (unclear, as some women must have had >1 cause for infertility)		155,624 person years*) infertile population comparison figure unclear		RR compared to the infertile population (95% CI): 1.26 (0.6 to 2.6)		• Adjusted for age at follow-up and calendar year • Infertility comparison adjusted for age at follow-up, calendar time, study site, gravidity at entry and cause of infertility • Women seeking treatment for infertility population (?more severe endometriosis) • Very limited baseline characteristics given • 31% self-reported ovarian cancer • Infertile population comparison-very imprecise (CIs cross both MIDs)
Brinton 2005, RC, Denmark, Danish Cancer Registry, Hospital Discharge Register and random subpopulation from Central Population Register First diagnosis of endometriosis patients included	N=2,491 endometriosis N=99,421 population comparison Unclear n values published	50	2,441	NR Split into <1, 1–4, >5 years	RR 1.69 (1.27 to 2.25)	Very high	• Adjusted for calendar time (per 5 years), parity, number of births, age at first birth • N values differ to those reported by Kim (2014) Systematic Review. Unclear. Figures could have come directly from contacting the authors • Unclear denominators and appropriateness of weighting • Very limited baseline characteristics • Only age was controlled for out of the major confounders listed by the Committee

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group* (cancer in women with endometriosis)	Expected/control group* (cancer in women without endometriosis)	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
No age restriction							
Buis 2013, RC Netherlands, OMEGA study linked to Dutch Pathology database and Netherlands Cancer Registry, self-reported questionnaire 1989–2007 Prevalent and incident cases of endometriosis included No age restriction specified	N=3,657 with endometriosis N=5,247 without endometriosis Ovarian endometriosis N=49	16	2	NR	HR 12.7 2.9 to 55.5 Ovarian endometriosis HR: 15.0 (3.1 to 72.4) – only adjusted for age	High Very high	• Adjusted for age, OC use, child, IVF • Generalisability of results due to subfertile population • All cancer cases are included from after the index date in main analysis • Mixed data collection methods (22% by self-reported questionnaire) • Only age was controlled for out of the major confounders listed by the Committee • Post-hoc analysis by site of endometriosis • Very wide CIs
Kobayashi 2007,2008, RC, Japan, Shizuoka Cohort Study on Endometriosis and Ovarian Cancer Programme, Shizuoka Cancer Registry, risk factor	N=6,398 with ovarian endometrioma (US) Compared to prefecture-wide rates of ovarian cancer	46	5.14	12.8 years 79,102 person years**	SIR 8.95 (4.12 to 15.3)	Very high	• Adjusted for age and calendar year only • Population ovarian endometrioma detected by ultrasound • Risk of misclassification and selection bias • Repeated ultrasound every 3-5 months (detection bias) • Very large CI

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group* (cancer in women with endometriosis)	Expected/control group* (cancer in women without endometriosis)	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
questionnaire 1985–1995 All those with evidence on US of ovarian endometrioma Age: 20–59 years							
Stewart 2013, RC, western Australia, 1982–2002 Whole population linked hospital and registry data (WA data linkage system), Hospital Morbidity Data System, Reproductive Technology Register, 1982–2002 All endometriosis diagnoses. It is indexed from infertility admission Age 20–40 years	N=2,978 (1,914 undergoing fertility treatment but not IVF, 1,064 having IVF) N=21,646 in the whole cohort (14,907 gave birth, 6,739 did not give birth)	NR	NR	NR Total cohort 17 years (366,041 person years**)	HR 2.33 (1.02 to 5.35) Subgroup: Birth In women that gave birth HR (95% CI): 1.52 (0.34 to 6.75) In women who did not give birth HR (95% CI): 3.11 (1.13 to 8.57)	Very high	• Total no. ovarian cancer=38 in a population of 21,646 • Adjusted for: birth, IVF, age at the start of follow-up, socioeconomic status • Subgroup analysis not specified in the methods • Population – infertility problems • Only total n value for ovarian cancer was reported, not by groups (endometriosis versus no endometriosis) • Large CI
Swedish National Register							

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group* (cancer in women with endometriosis)	Expected/control group* (cancer in women without endometriosis)	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Brinton 1997 RC, SNR 1966–-1983	N=20,686 endometriosis	29	15.11 Population	11.4 (range 1 to 21)	SIR 1.92 (1.3 to 2.8)	High	• Only 1 ICD code for endometriosis diagnosis • 55.6% truncated follow-up due to uncertainty of whether 1/both ovaries remained, e.g. after hysterectomy or oophorectomy • Population likely to have missed some cases (non-hospital admissions, milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage) • Post-hoc analysis by site of endometriosis (data not shown as also n values were not given for denominators) • Wide CIs
Melin 2006 RC, SNR 1969–2000 New diagnoses of endometriosis No age restriction (average 39.4 (SD 10.4) at entry	N=64,492 endometriosis N=25,430 ovarian endometriosis	122	85.09	12.7 years (444,931 person years**)	SIR 1.43 (1.19 to 1.71) Ovarian endometriosis SIR: 1.77 (1.38 to 2.24)	Very high	• Population likely to have missed some cases (non-hospital admissions, milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage) • Post-hoc analysis by site of endometriosis

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group* (cancer in women with endometriosis)	Expected/control group* (cancer in women without endometriosis)	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Melin 2007 RC, SNR1969–2002 New diagnoses of endometriosis No age restriction (average 39.5 (SD 10.5) at entry	N=63,630 endometriosis N=24,955 ovarian endometriosis	134	Not given	13.4 (792,013 person years**)	SIR 1.37 (1.14 to 1.62) Ovarian endometriosis SIR: 1.59 (1.26 to 1.98)	Very high	• Population likely to have missed some cases (non-hospital admissions) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage) • Post-hoc analysis by site of endometriosis
National Health Insurance Research Database (NHIRD) of Taiwan							
Chang 2014, RC, NHIRD, Registry for Catastrophic Illness Patients, 2000–2009 Age: 20–51 years newly diagnosed endometriosis	N=7,537 endometriosis (5,468 with surgical confirmation) N=15,074 comparison cohort (matched by age, index year, obstetric history, SES, work and urbanisation)	15 (2 non surgically confirmed endo, 13 surgically confirmed endo)	9	Mean NR 45,364 person years* in the endometriosis group 91,279 person years** comparison cohort	HR 3.28 (1.37 to 7.85) Surgically confirmed endometriosis: HR3.87 (1.58 to 9.47) Non surgically confirmed endometriosis: HR 1.64 0.35 to 7.80)	High	• Adjusted for age, SES, work, urbanisation, PID, infertility status, CVD, DM, chronic liver disease, rheumatic disease and Charlson Comorbidity Index • Post-hoc surgical diagnosis subgroup analysis • Unclear drop out/lost to follow-up/no description of censoring • Women <3 outpatient appointment within initial year of endometriosis diagnosis and without surgical confirmation were excluded (potentially milder cases excluded) • Only age and infertility were controlled for out of the major confounders listed by the CommitteeWide CIs

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group* (cancer in women with endometriosis)	Expected/control group* (cancer in women without endometriosis)	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Kok 2015, RC, NHIRD, Claims between 2002 and 2005 followed up until 31 December, 2008 Newly diagnosed endometriosis Age: >20 years	N=2,266 endometriosis N=9,064 comparison group (age and index matched 1:4)	13	9	Mean NR 9,842 person years** in the endometriosis cohort, 36,274 in the comparison cohort	HR 4.56 (1.72 to 12.11) Ovarian endometriosis HR 4.37 (1.07 to 17.83) Pure ovarian endometriosis HR 5.59 (0.67 to 46.48)	Overall : High Ovarian endometriosis: Very high Pure ovarian endometriosis: Very high	• Adjusted for age, diabetes mellitus, chronic kidney disease, liver cirrhosis, rheumatoid arthritis, use of medroxyprogesterone acetate, norethindrone acetate, danazol and GnRH agonist • Cases evaluated less than 3 times or for a follow-up period less than 2 months were excluded(n=3,099), (potentially milder cases excluded) • No censoring for women who have hysterectomy etc. after their index date • Unclear drop out/lost to follow-up but patients were censored at this point • Only age was controlled for out of the major confounders listed by the Committee • Post-hoc analysis by site of endometriosis
Lee 2015, RC, NHIRD and Registry for Catastrophic Illness Patients, 1996–2010 Age 20–51 years All cases of endometriosis	N=73,724 (recall diagnosis of endometriosis, to N=3,782 (tissue proven ovarian endometrioma) Comparison: 165,661 (no recall endometriosis)	166 recall 47 tissue proven	182 recall comparison group 301 tissue proven comparison group	Ranged from 3,228,799 to 3,409,338 person years** depending on diagnostic criteria	Epithelial ovarian cancer HR 1.90 (1.51 to 2.37) recall endo HR 18.57 (13.37 to 25.79) tissue proven endo	High	• Adjusted for PID, infertility status, Charlson co-morbidity index, age • Women who had a hysterectomy, bilateral salpingo-oophorectomy and bilateral oophorectomy were excluded, except those women with a diagnosis of EOC during the follow-up

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group* (cancer in women with endometriosis)	Expected/control group* (cancer in women without endometriosis)	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
(prevalent and incident)	235,703 (no tissue proven endometriosis) Both groups had to have at least 1 genealogical visit after 2000						• Presume 1st year of EOC was excluded as the paper only presents EOC values from 2001 to 2010 • Unclear drop out/lost to follow-up/inadequate basic information but patients were censored at this point • Only age and infertility were controlled for out of the major confounders listed by the Committee
Wang 2014, RC, NHIRD, Registry for Catastrophic Illness (National Cancer Registry), 2000–2010 Age: 20–51 years at entry Newly diagnosed endometriosis	N=5,945 new surgico-pathological diagnosis of endometriosis N=23,780 comparison cohort (matched on age, year, SES, work, obstetric history, frequency of gynae/obstetric provider's outpatient visits and urbanisation)	39	36	Mean NR 33, 519 person years* (women with endometriosis group) 135,408 person years** (comparison group)	Invasive epithelial ovarian cancer HR 5.62 (3.46 to 9.14) Subgroup analysis by age: <30 years HR: 3.34 (0.54 to 20.60) 30–39 years HR: 19.41 (5.02 to 75.10)	High	• Adjusted for PID, infertility status, CVD, DM, chronic liver disease and rheumatic disease • Study does not exclude diagnoses within the 1st year of the study • 29/39 EOC diagnosed in 1st year endometriosis group and 22/36 in the control group • Post-hoc age subgroup analysis. • Unclear drop out/lost to follow-up but patients were censored at this point Only age and infertility were controlled for out of the major confounders listed by the Committee

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group* (cancer in women with endometriosis)	Expected/control group* (cancer in women without endometriosis)	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
					40–49 years HR: 3.41 (1.76 to 6.61) ≥50 years HR: 9.63 (3.27 to 28.37)		
• The Danish National Patient Register							
Mogensen 2016, RC, Danish National Patient Register, 1977–-2012, mean follow-up 4.1 years Women with a diagnosis of endometriosis. Unclear if just new or includes old diagnoses Age: no age restriction. Mean age at diagnosis 59 years	N=45,356 endometriosis	221	142.64	Median 10.75 years	SIR 1.55 (1.35 to 1.77) Subgroup analysis by age at first endometriosis (years) <30: SIR 1.27 (0.71 to 2.10) 30–39: SIR 1.44 (1.10 to 1.85) 40–49: SIR 1.06 (0.83 to 1.34) ≥50: SIR 2.27 (1.61 to 3.10)	Very high	• Study was limited to only women who were hospitalised for endometriosis • Only age was considered as a confounding factor • Very limited baseline characteristics were provided in the paper

(a) RC: retrospective cohort; N: number of participants in study; SNR: Swedish National Register, SES: socioeconomic status; PID: pelvic inflammatory disease; CVD:

ISBN 978-1-4731-2661-9

(b) cardiovascular disease; DM: diabetes mellitus ,NHIRDs: National Health Insurance Research Database; SIR: standardised incidence ratio; RR: relative risk ratio; HR: hazard
(c) ratio; ICD: International Classification of Diseases; IVF: in vitro fertilisation; OC: oral contraceptive; CI: confidence Interval; MID: minimal important difference;
(d) Development Group; SD: standard deviation.
*(e) *Observed: The number of cancer cases in the study sample; Expected: The estimated number of cancer cases in the sample if the same was from a population without*
(f) Endometriosis (calculated using prevalence from a registry or external source.
*(g) **Person-years: the total number of years at risk across all participants (number of participants x years of follow-up). This accounts for different lengths of follow-up among*
(h) Different individuals.

Table 30: Borderline ovarian tumour

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk if bias	Comments
Buis 2013, RC Netherlands, OMEGA study linked to Dutch Pathology database and Netherlands Cancer Registry, self-reported questionnaire 1989–2007 Prevalent and incident cases of endometriosis included No age restriction specified	N=3,657 with endometriosis N=5,247 without endometriosis N=49 ovarian endometriosis	10	3	NR	HR 5.5 1.5 to 20.4 Ovarian endometriosis HR: 8.9 (2.2 to 35.7) only adjusted for age	High Very high	• Adjusted for age, OC use, child, IVF • Generalisability of results due to subfertile population • All cancer cases are included from after the index date in main analysis • Mixed data collection methods (22% by self-reported questionnaire) • Only age was controlled for out of the major confounders listed by the Committee • Post-hoc analysis by site of endometriosis • Very wide CIs
Brinton 2005, RC, Denmark, Danish Cancer Registry, Hospital Discharge Register and random subpopulation	N=2,491 endometriosis N=99,421 population comparison Unclear n values published	12	848	NR Split into <1, 1–4, >5 years	RR 1.22 (0.69 to 2.17)	Very high	• Adjusted for calendar time (per 5 years), parity, number of births, age at first birth • N values differ to those reported by Kim 2014 systematic review. Unclear. Figures could have come directly from contacting the authors

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk if bias	Comments
from Central Population Register All endometriosis patients included No age restriction							• Unclear denominators and appropriateness of weighting • Very limited baseline characteristics • Only age was controlled for out of the major confounders listed by the Committee • Very imprecise (both CIs cross the MIDs)

(a) CI: confidence interval; N: number of participants in study; RC: retrospective cohort; SIR: standardised incidence ratio; RR: relative risk ratio; HR: hazards Ratio; MID: minimal
(b) Important difference; OC: oral contraceptive; IVF: in vitro fertilisation.
*(c) *Observed: The number of cancer cases in the study sample; Expected: The estimated number of cancer cases in the sample if the same was from a population without*
(d) Endometriosis (calculated using prevalence from a registry or external source.

Table 31: Fallopian tube cancer

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Melin 2006 RC, SNR 1969–2000 New diagnoses of endometriosis No age restriction (average 39.4 (SD 10.4) at entry	N= 64,492 endometriosis	10	8.32	12.7 years (766,498 person years**)	SIR 1.20 (0.58 to 2.21)	Very high	• Population likely to have missed some cases (non-hospital admissions, milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage) • Wide CIs

(a) RC: retrospective cohort; N: number of participants in study; SNR: Swedish National Registry; SIR: standardised incidence ratio; HR: hazard ratio; RR: relative risk ratio; CI:
(b) confidence interval; SD: standard deviation
*(c) *Observed: The number of cancer cases in the study sample; Expected: The estimated number of cancer cases in the sample if the same was from a population without*
(d) Endometriosis (calculated using prevalence from a registry or external source.
*(e) **Person-years: the total number of years at risk across all participants (number of participants x years of follow-up). This accounts for different lengths of follow-up among*
(f) different individuals.

ISBN 978-1-4731-2661-9

Table 32: Uterus not otherwise specified / uterine cancer

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
Brinton 2004/2005A, RC, USA, 1965–1988, 5 reproductive endocrinology centres, cancer registries, National Death Index, questionnaires, 1965–1988 Endometriosis included in selection criteria (seen >1/referred by physician for infertility advice) No age restriction	N=1,919 endometriosis N=6,510 infertility population (unclear, as some women must have had >1 cause for infertility) Above figures taken from Brinton 2004. Unclear figures published in Brinton 2005 as total n=8,422 but methods are quoted to be from the Brinton 2004 paper	NR	Infertile population comparison figure unclear Total for both groups =39	Median- 35,196 person years**	RR compared to the infertile population (95% CI): 0.82 (0.3 to 1.9)	Very high	• 20% lost to follow-up • Excluded 1st year follow-up data • Infertility comparison adjusted for age at follow-up, calendar time, study site, gravidity at entry and cause of infertility • Women seeking treatment for infertility (possibly more severe endometriosis) • Very limited baseline characteristics given • N values unclear • Infertile population comparison-very imprecise (CIs cross both MIDs)
Brinton 2005, RC, Denmark, Danish Cancer Registry, Hospital Discharge Register and random subpopulation from Central Population Register All endometriosis patients included	N=2,491 endometriosis N=99,421 population comparison Unclear n values published	9	1389	Not recorded Split into <1, 1-4, >5 years	RR 1.23 (0.63 to 2.38)	Very high	• Adjusted for calendar time (per 5 years), parity, number of births, age at first birth • N values differ to those reported by Kim 2014 Systematic Review. Unclear. Figures could have come directly from contacting the authors • Unclear denominators and appropriateness of weighting • Very limited baseline characteristics

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
No age restriction							• Only age was controlled for out of the major confounders listed by the Committee • Very imprecise (CIs cross both MIDs)
Swedish National Register							
Brinton 1997 RC, SNR 1969–1983 All endometriosis patients included No age restriction (range 12–82 years)	N=20,686 endometriosis	1	1.69 Population	11.4 (range 1 to21)	SIR 0.59 (0.00 to 3.3)	Very high	• Only 1 ICD code for endometriosis diagnosis • 55.6% truncated follow-up due to uncertainty of whether 1/both ovaries remained, e.g. after hysterectomy or oophorectomy • Population likely to have missed some cases (non-hospital admissions, milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage) • Post-hoc analysis by site of endometriosis (data not shown as also n values were not given for denominators) • Wide CIs
Melin 2006 RC, SNR 1969–2000 New diagnoses of endometriosis No age restriction (average 39.4	N=64,492 endometriosis	11	10.33	12.7 years (427,220 person years**)	SIR 1.06 (0.53 to 1.90)	Very high	• Population likely to have missed some cases (non-hospital admissions, milder cases) • Very limited baseline characteristic data • Population comparison (only age and calendar year adjustment, 60–85% coverage) • Wide CIs

ISBN 978-1-4731-2661-9

Study, design and data source	Number of participants (N)	Observed/case group*	Expected/control group*	Mean follow-up (years)	RR/HR/SIR (95%CI)	Risk of bias	Comments
(SD 10.4) at entry							

(a) RC: retrospective cohort, SNR: Swedish National Registry, SIR: standardised incidence ratio; RR: relative risk ratio; CI: confidence interval; MID: minimal important difference;
(b) HR: hazard ratio; ICD: International Classification of Disease
*(c) *Observed: The number of cancer cases in the study sample; Expected: The estimated number of cancer cases in the sample if the same was from a population without*
(d) Endometriosis (calculated using prevalence from a registry or external source.
*(e) **Person-years: the total number of years at risk across all participants (number of participants x years of follow-up). This accounts for different lengths of follow-up among*
(f) Different individuals.

ISBN 978-1-4731-2661-9

8.3 Economic evidence

No health economic studies were found relevant to this question, and therefore no health economic modelling was conducted for this question.

8.4 Clinical evidence statements

8.4.1 Cervical cancer

Three studies with very high to high risk of bias with 20,686 to 64,492 women with endometriosis compared with the rest of the Swedish population found the standardised incidence ratios (SIRs) ranged from 0.64 to 0.72, with variable uncertainty. This would suggest that there is not an increased risk of cervical cancer in women with endometriosis.

8.4.2 Cancer in situ of the cervix

One study with moderate risk of bias with 64,492 women with endometriosis was compared to the rest of the Swedish population and found a reduced SIR of 0.89, with little uncertainty. This would suggest that there is not an increased risk of CIS of the cervix in women with endometriosis.

8.4.3 Endometrial cancer

Three studies with very high to high risk of bias with 20,686 to 63,630 women with endometriosis compared with the rest of the Swedish population found the SIRs ranged from 1.09 to1.19 with variable uncertainty. One study with very high risk of bias with 43,734 women hospitalised with endometriosis compared with the rest of the Danish population found an increased risk of endometrial cancer in the women with endometriosis. The SIR was 2.13 (1.77–2.55).

Two studies with very high to high risk of bias based in Taiwan, looked at 2,266 and 15,488 women with endometriosis compared with 9,064 and 123,904 women without endometriosis and found an increased hazard ratio (HR) of 4.05 and 2.83 respectively, with large confidence intervals (Cis). The differences between the results of the Swedish and Taiwanese studies could be due to a variety of confounding factors (geographical variations, detection differences, statistical analysis and major confounder adjustment). Overall it is unclear whether there is an increased risk of endometrial cancer in women with endometriosis.

8.4.4 Ovarian cancer

14 studies with very high to high risk of bias with a population of women with endometriosis ranging from 1,919 to 73,724 and a comparison group population of 5,247 to 235,703 (when reported) suggest an increased risk of ovarian cancer in women with endometriosis. Although the studies vary in size, confounder adjustment, statistical analysis (relative risk (RR), HR, SIR) and comparison group populations (population wide, matched, infertile, geography), they all indicate an increased risk of ovarian cancer with variable certainty of the size of the risk.

8.4.5 Borderline ovarian tumour

Two studies with very high and high risk of bias compared women with endometriosis (n=2,491, n=3,657) with those without endometriosis (n=99,421, n=5247) in a Danish and subfertile population, respectively. The Danish population study did not demonstrate any

ISBN 978-1-4731-2661-9

clinical evidence of an increased risk of borderline ovarian tumour in those with endometriosis. However, compared with the subfertile population, the women with endometriosis were suggested to have an increased risk of borderline ovarian tumour, the degree of which was uncertain.

Overall, it is unclear whether there is an increased risk of borderline ovarian tumour in women with endometriosis.

8.4.6 Fallopian tube cancer

One study with very high risk of bias of 64,492 women with endometriosis who were compared with the Swedish population, demonstrated no clinical evidence of an increased risk of fallopian tube cancer with high uncertainty.

8.4.7 Uterine otherwise not specified/uterine cancer

Four studies of very high risk of bias showed no clinical difference in uterine otherwise not specified/uterine cancer between women with endometriosis (n=1,919–64,492) and women without endometriosis (number in the population was not clearly reported but was up to 99,421), with high uncertainty.

8.5 Evidence to recommendations

8.5.1 Relative value placed on the outcomes considered

The Committee agreed that the risk of developing cancers of reproductive organs in women with endometriosis compared with women without endometriosis was considered to be the only critical outcome for decision-making.

8.5.2 Consideration of clinical benefits and harms

The Committee noted that many women with endometriosis ask questions about whether or not the condition is associated with an increased risk of cancer.

Even though very large population-based studies were identified, the Committee were cautious about drawing conclusions from the results because the evidence base was generally of low to very low quality.

The Committee agreed that the evidence did not show an overall increased risk in uterine and cervical cancers. The pattern of the risk of endometrial cancer was heterogeneous and therefore no clear conclusions could be drawn.

Committee members focused mainly on results relating to a risk of ovarian cancer. They agreed that these data seemed to be relatively consistent in indicating an increased risk in women with endometriosis compared to women without endometriosis. However, they discussed that the evidence was mainly of very low to low quality and that due to various limitations an absolute risk could not be derived from these data. They recognised that there needs to be a balance between women being fully informed about their condition (including related risks), with rationales for not encouraging unnecessary treatments. The Committee also took into account that there is no national screening available for ovarian cancer and that there is no clear management plan that would help to reduce a possible small increased risk. Since an absolute risk could not be quantified, it was decided that recommendations would not aid decision making, would cause anxiety in women with endometriosis and could be misconstrued, for example women seeking treatments (such as removal of the ovaries) that this small risk increase would not warrant.

ISBN 978-1-4731-2661-9

The Committee discussed whether a research recommendation should be made. It was decided not to prioritise this because it would require longitudinal follow-up accounting for all possible confounding factors and would therefore require a registry type dataset of all women in the UK. This was considered to be a very ambitious project that would be unlikely to be taken up.

8.5.3 Consideration of economic benefits and harms

The frequency of clinical reviews to assess the need and timing of referral to specialist services has implications for health care resources. If the frequency of review is too great then additional resources will be used for insufficient gain. Alternatively, if the referral criteria are too loose (i.e. the tests too sensitive) then women who do not need to use specialist services will be referred there, presenting a direct cost to these services. Conversely, if the frequency or sensitivity of the tests is insufficient to identify women with these conditions then this will carry an economic cost; possibly healthcare loss relating to late treatment of the conditions, possibly direct financial cost relating to the increased complexity of treating the more advanced condition, or possibly both.

However, in the absence of clinical evidence it is difficult to suggest an optimum frequency. It was therefore not thought appropriate to attempt a health economic synthesis of the evidence, as the Committee could not provide a selection of competing alternatives with differing opportunity costs, which are required for health economic analysis. Additionally, the downstream costs of referral to specialist services are difficult to estimate, in view of the fact that the management of these conditions requires specialist knowledge not possessed by the Committee. Consequently, this question was not prioritised for economic analysis and no health economic evidence was identified.

8.5.4 Quality of evidence

Risk of bias was assessed as being very high to high according to the critical appraisal tool for prevalence studies. The level of description and measurement was often very different between the studies and possible confounders were not systematically adjusted for. Baseline characteristics were often not described at all, or were only poorly described. It was also often unclear whether all cases of endometriosis would have been captured using only the International Classification of Disease (ICD) codes. Only using ICD codes would be likely to miss less severe cases of cancer. The uncertainty around the effect was often very large, which makes it difficult to be confident about the findings.

Due to these concerns regarding risk of bias in the included studies, the Committee interpreted these data with caution.

8.5.5 Other considerations

The Committee also discussed the associated NICE guidance for the recognition and referral for suspected cancer, which is applicable to this section (Suspected cancer: recognition and referral – NICE guideline NG12). Even though the assessment of the signs and symptoms of cancers of reproductive organs is outside the scope of this guideline, ‘monitoring and referral’ is within the scope. They were particularly interested in the site-specific cancer guidance related to gynaecological cancers.

8.5.6 Key conclusions

The Committee concluded that no recommendations should be made based on the available evidence. The most consistent results were related to a possible small increased risk of ovarian cancer in women with endometriosis compared to women without endometriosis. However, because of evidence limitations and the inability to quantify this risk in absolute terms, the Committee decided against making recommendations after considerable debate.

ISBN 978-1-4731-2661-9

They decided this because the potential harms associated with misinterpretation or over-interpretation of any recommendation based on this data would outweigh any benefits conferred by women being specifically informed about this data. This may lead to unnecessary procedures. The Committee agreed that for other types of cancer of reproductive organs the evidence was negative or inconclusive.

8.6 Recommendations

No recommendation was made.

ISBN 978-1-4731-2661-9

9 Diagnosis

9.1 Introduction: diagnostic testing

A diagnostic delay of 5-10 years between presentation and diagnosis of endometriosis is not unusual and this can have devastating effects on women's quality of life. Delayed diagnosis is mainly due to the non-specific nature of the associated symptoms and the need to verify the disease surgically. In addition, there may be little relationship between the severity of the symptoms and the extent of disease, further complicating successful diagnosis.

It is important that women with endometriosis are assessed and a diagnosis made in a timely manner, to prevent delay in effective treatment. While medical treatment may be commenced empirically, it is essential to be as confident as possible that the underlying diagnostic assumptions are correct and to identify any findings that require more urgent treatment. If surgery is considered, the accuracy of pre-operative diagnostic tests is important to determine correct care and timely intervention. An accurate impression of the extent of disease enables women to be treated in the appropriate setting where all the required specialist services are available. Long-term ineffective care carries with it significant morbidity for the woman.

The main imaging modalities utilised in diagnosing and mapping endometriosis are ultrasound (abdominal, vaginal and rectal) and MRI imaging. While both investigation modalities are operator dependant, there is potential for more inter-observer variation in ultrasound than MRI, which may be reported by a second radiologist, providing some quality control.

This chapter reviews the efficacy of all the diagnostic modalities for identifying endometriosis. In addition, the cost effectiveness of each test modality and their place in the diagnostic hierarchy will be determined.

9.2 Ultrasound

Review question: What is the accuracy of ultrasound in diagnosing endometriosis?

9.2.1 Introduction

The 'gold standard' for diagnosis of endometriosis has been considered to be laparoscopy with biopsy which allows histological confirmation of suspicious lesions. Endometriosis might be suspected and empirically managed in community (GP) services, but a definitive diagnosis is usually made after gynaecological referral and surgery which requires a general anaesthetic and a period of recovery. Imaging is also widely used. Ultrasound technology has developed in recent years and magnetic resonance imaging has become more readily available.

The aim of this review was to evaluate accuracy of ultrasound for the diagnosis of endometriosis in women with suspected endometriosis.

For full details see review protocol in Appendix D.

9.2.2 Description of clinical evidence

No test-and-treat trials were found, therefore no clinical or patient-reported outcomes such as quality of life were identified.

Three studies were included in this review. Evidence was available from 1 Cochrane systematic review (Nisenblat 2016) and 2 further observational studies (Bahr 2006 and

ISBN 978-1-4731-2661-9

Sayasneh 2015). 32 studies within the Cochrane systematic review were relevant and results from these are included here (Abrao 2007; Bazot 2009; Bergamini 2010; Dessole 2003; Eskenazi 2001; Falco 2011; Fedele 1998; Ferrero 2011; Ghezzi 2005; Goncalves 2010; Grasso 2010; Guerriero 1996a; Guerriero 1996b; Guerriero 2007; Guerriero 2008; Guerriero 2014; Holland 2010; Hudelist 2011; Hudelist 2013; Leon 2014; Mangler 2013; Menada 2008; Pascual 2010; Piessens 2014; Piketty 2009; Reid 2014; Ribeiro 2008; Said 2014; Savelli 2011; Scarella 2013; Reid 2013; Ubaldi 1998).

Of the included studies, 14 were from Italy (Bergamini 2010; Dessole 2003; Eskenazi 2001; Falco 2011; Fedele 1998; Ferrero 2011; Grasso 2010; Guerriero 1996a; Guerriero 1996b; Guerriero 2007; Guerriero 2008; Guerriero 2014; Menada 2008; Savelli 2011), 3 from Brazil (Abrao 2007; Goncalves 2010; Ribeiro 2008), France (Bahr 2006; Bazot 2009; Piketty 2009) and Australia (Piessens 2012; Reid 2013; Reid 2014), 2 from UK (Holland 2010; Sayasneh 2015), Austria (Hudelist 2011; Hudelist 2013) and Chile (Leon 2014; Scarella 2013) and 1 each from Belgium (Ubaldi 1998), Germany (Managler 2013), Spain (Pascual 2010), Switzerland (Ghezzi 2005) and Egypt (Said 2014).

The size of the population in each of the studies ranged from 33 (Grasso 2010) to 722 (Ghezzi 2005).

With regard to the types of ultrasound used, studies reported the following methods:

- transvaginal ultrasonography (TVUS) (Abrao 2007; Bazot 2009; Bergamini 2010; Dessole 2003; Eskenazi 2001; Falco 2011; Ghezzi 2005; Guerriero 1996a; Guerriero 1996b; Guerriero 2007; Holland 2010; Hudelist 2011; Hudelist 2013; Menada 2008; Piketty 2009; Mangler 2013; Reid 2013; Savelli 2011; Sayasneh 2015; Ubaldi 1998),
- transrectal ultrasonography (TRUS) (Bahr 2006; Bazot 2009; Bergamini 2010; Fedele 1998; Piketty 2009; Said 2014),
- rectal water contrast transvaginal ultrasonography (RWC-TVUS) (Ferrero 2011; Menada 2008),
- transvaginal ultrasonography with bowel preparation (TVUS-BP) (Goncalves 2010; Piessens 2014; Scarella 2013),
- sonovaginography (SVG) (Dessole 2003; Reid 2014),
- tenderness-guided TVUS (tg-TVUS) (Guerriero 2008; Guerriero 2014),
- TVUS kissing ovaries sign (Ferrero 2011), 3-dimensional transvaginal ultrasonography (3D-TVUS) (Grasso 2010),
- introital 3-dimensional ultrasound (3D-US) (Pascual 2010),
- SVG+TVUS-BP (Leon 2014).

Seven studies compared more than 1 ultrasound method in the same cohort of women (Bazot 2009; Bergamini 2010; Dessole 2003; Dessole 2003; Guerriero 2014; Menada 2008; Piketty 2009).

This review reports diagnostic accuracy outcomes such as sensitivity and specificity.

Evidence from the included studies are summarised in the clinical GRADE evidence profile below (Table 33). Modified GRADE was used to assess the quality of the evidence diagnostic outcomes. See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots and ROC plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G.

9.2.3 Summary of included studies

A summary of the studies that were included in this review is presented in Table 33.

ISBN 978-1-4731-2661-9

Table 33: Summary of included studies

Study	Index test/reference standard	Population	Type of endometriosis/outcomes
Abrao 2007 (Nisenblat 2016 CSR) Brazil	• TVUS • Laparoscopy and histology	Women with clinically suspected endometriosis N=104 (consecutive)	DIE sites: rectovaginal septum, rectosigmoid involvement Sensitivity and specificity
Bahr 2006 France	• TRUS • Surgery (not specified) and histology	Women suspected of having deep pelvic endometriosis on the basis of outpatient history and/or clinical symptoms with a mass palpable on bimanual examination that might infiltrate the rectal wall. N=37 (consecutive)	DIE site: bowel Sensitivity and specificity
Bazot 2009 (Nisenblat 2016 CSR) France	• TVUS • TRUS • Laparoscopy/laparotomy and histology	Women referred with clinical evidence of pelvic endometriosis. N=92 (consecutive)	DIE sites: uterosacral ligaments, rectovaginal septum, vaginal wall, rectosigmoid involvement Ovarian endometriosis Sensitivity and specificity
Bergamini 2010 (Nisenblat 2016 CSR) Italy	• TVUS • TRUS • Laparoscopy or laparotomy and histology	Women scheduled for surgery because of signs and symptoms of severe posterior deep infiltrating endometriosis. N=61 (consecutive)	DIE site: rectosigmoid involvement Sensitivity and specificity
Dessole 2003 (Nisenblat 2016 CSR) Italy	• TVUS • SVG • Laparoscopy or laparotomy and histology	Women scheduled for laparotomy or laparoscopy because of rectovaginal endometriosis suspected on the basis of patient history and/or clinical examination. N=46	DIE site: posterior DIE Sensitivity and specificity
Eskenazi 2001 (Nisenblat 2016 CSR) Italy	• TVUS • Laparoscopy or laparotomy and histology	Women scheduled to undergo a laparoscopy or laparotomy for pelvic pain, infertility, tubal ligation or masses of the adnexus or uterus. N=90	Pelvic endometriosis Sensitivity and specificity
Falco 2011 (Nisenblat 2016 CSR) Italy	• TVUS • Laparoscopy and histology	Women scheduled for laparoscopy with ≥ 1 symptom suggestive for the presence of endometriosis. N=128	Pelvic endometriosis DIE sites: posterior DIE, uterosacral ligaments, rectovaginal septum, vaginal wall, rectosigmoid involvement

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Type of endometriosis/outcomes
			Sensitivity and specificity
Fedele 1998 (Nisenblat 2016 CSR) Italy	• TRUS • Laparoscopy or laparotomy and histology	Women scheduled for laparoscopy or laparotomy for pelvic endometriosis, suspected on basis of history and objective findings (not specified). N=140	DIE sites: uterosacral ligaments, rectovaginal septum, vaginal wall, rectosigmoid involvement Sensitivity and specificity
Ferrero 2011 (Nisenblat 2016 CSR) Italy	• RWC-TVS • Laparoscopy and histology	Women with suspected deep pelvic endometriosis. N=96	DIE sites: rectosigmoid involvement, bowel (ileum-rectum) Sensitivity and specificity
Ghezzi 2005 (Nisenblat 2016 CSR) Switzerland	• TVUS (kissing ovaries sign) • Laparoscopy and histology	Premenopausal women with adnexal mass or with clinical signs suggestive of pelvic endometriosis who were scheduled for laparoscopic surgery. N=722 (consecutive)	Pelvic endometriosis Sensitivity and specificity
Goncalves 2010 (Nisenblat 2016 CSR) Brazil	• TVUS-BP • Laparoscopy+ histology	Women submitted to laparoscopy on suspicion of endometriosis. N=194 (consecutive)	DIE site: rectosigmoid involvement Sensitivity and specificity
Grasso 2010 (Nisenblat 2016 CSR) Italy	• 3D-TVUS • Laparoscopy and histology	Women with clinical suspicion of pelvic endometriosis. N=33	DIE DIE site: bladder* Sensitivity and specificity
Guerriero 1996a (Nisenblat 2016 CSR) Italy	• TVUS • Laparoscopy or laparotomy and histology	Women scheduled for laparoscopy or laparotomy for a persistent ovarian mass. N=118 (consecutive)	Ovarian endometriosis Sensitivity and specificity
Guerriero 1996b (Nisenblat 2016 CSR) Italy	• TVUS • Laparoscopy or laparotomy and histology	Women who were submitted to laparoscopy or laparotomy because of the presence of a persistent adnexal mass. N=101 (consecutive)	Ovarian endometriosis Sensitivity and specificity
Guerriero 2007 (Nisenblat 2016 CSR)	• TVUS • Laparoscopy and histology	Women scheduled for laparoscopic surgery for rectovaginal endometriosis, suspected on the basis of patient history of pelvic pain and/or clinical examination. N=50 (consecutive)	DIE site: posterior DIE Ovarian endometriosis Sensitivity and specificity
Guerriero 2008	• tg-TVUS • Laparoscopy and histology	Women scheduled for surgery for clinically suspected endometriosis (on	DIE sites: anterior DIE, uterosacral ligaments, rectovaginal septum,

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Type of endometriosis/outcomes
(Nisenblat 2016 CSR) Italy		the basis of patient history of pelvic pain and/or clinical examination). N=88 (consecutive)	vaginal wall, rectosigmoid involvement, bladder* Sensitivity and specificity
Guerriero 2014 (Nisenblat 2016 CSR) Italy	• Tg-TVUS • 3D-US • Laparoscopy or laparotomy and histology	All premenopausal women with clinical suspicion of deep endometriosis who were scheduled for surgery in our department. N=202 (consecutive)	DIE sites: posterior DIE, rectosigmoid involvement Sensitivity and specificity
Holland 2010 (Nisenblat 2016 CSR) UK	• TVUS • Laparoscopy	Women with clinically suspected or proven pelvic endometriosis. N=201 (consecutive)	Any DIE DIE sites: posterior DIE, pouch of Douglas Pelvic endometriosis Sensitivity and specificity
Hudelist 2011 (Nisenblat 2016 CSR) Austria	• TVUS • Laparoscopy and histology	Women with suspected endometriosis attending 1 of 3 pelvic pain clinics who were referred to the pelvic pain clinic for laparoscopy because of suspected endometriosis on the basis of clinical history and the referring physician's clinical findings, or were self-referred (coming to the pain clinic without seeing any gynaecologist before this time for their current problems). N=153	DIE sites: uterosacral ligaments, rectovaginal septum, vaginal wall, pouch of Douglas rectosigmoid involvement bladder* ovarian endometriosis Sensitivity and specificity
Hudelist 2013 (Nisenblat 2016 CSR) Austria	• TVUS • Laparoscopy and histology	Women attending pelvic pain clinic with suspected endometriosis and scheduled for laparoscopy on the basis of clinical examination and TVUS findings. N=142 (consecutive)	DIE site: rectosigmoid involvement Sensitivity and specificity
Leon 2014 (Nisenblat 2016 CSR) Chile	• SVG+TVUS-BP • Laparoscopy and histology	Women with clinical suspicion of DIE based on clinical symptoms (chronic pelvic pain, deep dyspareunia, dyschezia, catamenial rectal bleeding, catamenial hematuria) or physical pelvic examination findings (non-mobile uterus, posterior vaginal fornix nodules, a painful pelvic examination). N=110	DIE sites: pouch of Douglas, bladder* Sensitivity and specificity

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Type of endometriosis/outcomes
Mangler 2013 (Nisenblat 2016 CSR) Germany	• TVUS • Laparoscopy or laparotomy and histology	Patients with suspected or known rectovaginal endometriosis who were operated on at the study authors' institution. Endometriosis suspected on the basis of clinical symptoms, abnormal gynaecological examination or other imaging tests, or known through previous operations. N=79 (consecutive)	DIE site: rectosigmoid involvement Sensitivity and specificity
Menada 2008 (Nisenblat 2016 CSR) Italy	• TVUS • RWC-TVUS • Laparoscopy and histology	Women with suspected rectovaginal endometriosis on the basis of pain symptoms and/or gynaecological examination. N=90	DIE site: rectovaginal septum Sensitivity and specificity
Pascual 2010 (Nisenblat 2016 CSR) Spain	• 3D-US • Laparoscopy and histology	Women with clinically suspected endometriosis based on patient history of pelvic pain and/or clinical examination. N=39 (consecutive)	DIE site: rectovaginal septum Sensitivity and specificity
Piessens 2014 (Nisenblat 2016 CSR) Australia	• TVUS-BP • Laparoscopy and histology	Women with clinically suspected endometriosis referred to TVUS. N=205 (prospective)	DIE sites: vaginal wall pouch of Douglas, bowel (ileum-rectum), bladder* Ovarian endometriosis Sensitivity and specificity
Piketty 2009 (Nisenblat 2016 CSR) France	• TVUS • TRUS • Laparoscopy or laparotomy and histology	Women suffering from pelvic pain (alone or associated with infertility) who underwent complete surgical exeresis of deeply infiltrating endometriosis, which was suspected in all cases preoperatively (questioning, clinical examination, imaging). N=134	DIE site: bowel (ileum-rectum) Sensitivity and specificity
Reid 2013 (Nisenblat 2016 CSR) Australia	• TVUS • Laparoscopy and histology	Women with a history of chronic pelvic pain and/or endometriosis and scheduled for operative laparoscopy. N=100	DIE sites: uterosacral ligaments, rectovaginal septum, pouch of Douglas, rectosigmoid involvement Sensitivity and specificity
Reid 2014 (Nisenblat 2016 CSR)	• SVG • Laparoscopy and histology	Women who presented to pelvic pain clinic with	DIE sites: posterior DIE,

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Type of endometriosis/outcomes
Australia		symptoms suggestive of endometriosis. N=220 (consecutive)	uterosacral ligaments, rectovaginal septum, vaginal wall, pouch of Douglas, rectosigmoid involvement Sensitivity and specificity
Ribeiro 2008 (Nisenblat 2016 CSR) Brazil	• TRUS • Laparoscopy and histology	Women with clinically suspected deeply infiltrating endometriosis referred to gynaecological endoscopy and endometriosis clinic N=37 (consecutive)	DIE site: rectosigmoid involvement Sensitivity and specificity
Said 2014 (Nisenblat 2016 CSR) Egypt	• TVUS • Laparoscopy and histology	Women with any symptoms suggestive of endometriosis who were booked for laparoscopy N=142 (consecutive)	Pelvic endometriosis Sensitivity and specificity
Savelli 2011 (Nisenblat 2016 CSR) Italy	• TVUS • Laparoscopy and histology	Women with results of pelvic examination or symptoms suggestive of DIE of the posterior compartment N=94 (consecutive)	DIE sites: posterior DIE, rectosigmoid involvement Sensitivity and specificity
Scarella 2013 (Nisenblat 2016 CSR) Chile	• TVUS-BP • Laparoscopy or laparotomy + histology	Women with chronic pelvic pain and/or suspected endometriosis N=100 (consecutive)	DIE DIE sites: uterosacral ligaments, rectovaginal septum Sensitivity and specificity
Sayasneh 2015	• TVUS • Laparoscopy and histology	Women referred because of suspected or confirmed pelvic mass observed on ultrasound examination N=313 (consecutive)	Ovarian endometriosis Sensitivity and specificity
Ubaldi 1998 (Nisenblat 2016 CSR) Belgium	• TVUS • Laparoscopy and histology	Women who had been referred for diagnostic or operative laparoscopy for infertility, chronic pelvic pain and/or adnexal masses N=133	Ovarian endometriosis Sensitivity and specificity

N: number of participants in study; DIE: deeply infiltrating endometriosis; CSR: Cochrane systematic review TVUS: transvaginal ultrasonography; TRUS: transrectal ultrasonography; RWC-TVUS: rectal water contrast transvaginal ultrasonography; SVG: sonovaginography; TVUS-BP: transvaginal ultrasonography with bowel preparation; 3D-TVUS: 3-dimensional transvaginal ultrasonography; tg-TVUS: tenderness-guided TVUS; 3D-US: introital 3-dimensional ultrasound
**bladder data from the original study, calculated by the technical team*

9.2.4 Clinical evidence profile

The clinical evidence profile for this review question is presented in Table 34.

ISBN 978-1-4731-2661-9

Table 34: Summary clinical evidence profile for diagnosis of endometriosis using ultrasound

Sensitivity (95%CI)	Specificity (95% CI)	Site of endometriosis as diagnosed using ultrasound	No of participants (no. of studies)	Quality of the evidence (GRADE)13
62% (18 to 94)	93% (78 to 99)	Pelvic1 (TVUS, tg-TVUS, kissing ovaries sign)	1222 (5)	⊕⊖⊖⊖ Very low
88% (70 to 97)	95% (85 to 99)	Bowel2 (TVUS, RWC-TVUS, TVUS-BP)	314 (3)	⊕⊖⊖⊖ Very low
88% (47 to-100) 96% (89 to 99)	97% (82 to 100) 100% (94 to 100)	Bowel2 (TRUS)	171 (2)	⊕⊖⊖⊖ Very low
78% (37 to 97)	90% (58 to 99)	DIE3 (TVUS, TVUS-BP, 3D-TVUS)	282 (3)	⊕⊖⊖⊖ Very low
73% (55 to 87)	91% (76 to 98)	Posterior DIE4 (TVUS, tg-TVUS, SVG)	853 (7)	⊕⊖⊖⊖ Very low
91% (75 to 98) 87% (78 to 93)	86% (57 to 98) 94% (87 to 97)	Posterior DIE4 (SVG and 3D-TVUS)	248 (2)	⊕⊖⊖⊖ Very low
33% (13 to 59)	100% (95 to 100)	Anterior DIE5 (TVUS)	88 (1)	⊕⊕⊖⊖ Low
66% (33 to 90)	98% (95 to 99)	Rectovaginal[6] (TVUS, TVUS-BP, tg-TVUS, introital 3D-US, SVG)	983 (10)	⊕⊖⊖⊖ Very low
97% (90 to 100)	100% (84 to 100)	Rectovaginal[6] (RWC-TVUS)	90 (1)	⊕⊕⊖⊖ Low
18% (2 to 52) 97% (85 to 100)	95% (88% to 99%) 96% (91 to 99)	Rectovaginal[6] (TRUS)	232 (2)	⊕⊖⊖⊖ Very low
89% (80 to 95)	96% (93 to 98)	Rectosigmoid[7] (TVUS, TVUS-BP, tg-TVUS, RWC-TVUS, SVG)	1615 (14)	⊕⊖⊖⊖ Very low
91% (82 to 96)	97% (92 to 99)	Rectosigmoid[7] (3D-TVUS)	202 (1)	⊕⊕⊖⊖ Low
90% (77 to 98)	93% (79 to 99)	Rectosigmoid[7] (TRUS)	330 (4)	⊕⊖⊖⊖ Very low
63% (45 to 79)	96% (91 to 98)	Uterosacral ligament[8] (TVUS, tg-TVUS, TVUS-BP, SVG)	714 (7)	⊕⊖⊖⊖ Very low
48% (37 to 59) 80% (44 to 97)	44% (14 to 79) 98% (93 to 100)	Uterosacral ligament[8] (TRUS)	232 (2)	⊕⊖⊖⊖ Very low
57% (26 to 84)	98% (94 to 100)	Vaginal wall involvement[9] (TVUS, TVUS-BP, tg-TVUS, SVG)	679 (6)	⊕⊖⊖⊖ Very low
7% (1 to 22) 100% (79 to 100)	100% (94 to 100) 100% (97 to 100)	Vaginal wall involvement[9] (TRUS)	232 (2)	⊕⊖⊖⊖ Very low
83% (71 to 91)	97% (93 to 99)	Pouch of Douglas[10] (TVUS, TVUS-BP, SVG+TVUS-BP)	755 (6)	⊕⊖⊖⊖ Very low

ISBN 978-1-4731-2661-9

Sensitivity (95%CI)	Specificity (95% CI)	Site of endometriosis as diagnosed using ultrasound	No of participants (no. of studies)	Quality of the evidence (GRADE)13
35% (13-63)	98% (96 to 100)	Bladder[11] (TVUS, TVUS-BP, tg-TVUS, 3D-TVUS, SVG+TVUS-BP)	383 (5)	⊕⊖⊖⊖ Very low
90% (83 to 96)	96% (93 to 98)	Ovarian[12] (TVUS, TVUS-BP, tg-TVUS)	1066 (9)	⊕⊕⊖⊖ Low
89% (74 to 97)	77% (64 to 87)	Ovarian[12] (TRUS)	92 (1)	⊕⊕⊖⊖ Low

CI: confidence interval; TVUS: transvaginal ultrasonography; TRUS: transrectal ultrasonography; RWC-TVUS: rectal water contrast transvaginal ultrasonography; SVG: sonovaginography; TVUS-BP: transvaginal ultrasonography with bowel preparation; 3D-TVUS: 3-dimensional transvaginal ultrasonography
Endometriosis sites as defined in Nisenblat Cochrane Systematic Review 2016:
1 Endometriotic lesions, deep or superficial, located at any site in pelvic/abdominal cavity: on the peritoneum, fallopian tubes, ovaries, uterus, bowel, bladder or Pouch of Douglas
2 Endometriotic lesions infiltrating at least the muscular layer of the intestinal wall ileum - rectum; predominantly affects rectosigmoid colon
3 Deep endometriotic lesions extending more than 5 mm under the peritoneum located at any site of pelvic/abdominal cavity
4 Deep endometriotic lesions involve ≥ 1 site of the posterior pelvic compartment (uterosacral ligament, rectovaginal septum, vaginal wall, and bowel) and/or obliterate Pouch of Douglas
5 Deep endometriotic lesions located at any site of the anterior pelvic compartment (bladder ± anterior pouch)
6 Deep endometriotic implants infiltrate the retroperitoneal area between posterior wall of vaginal mucosa and anterior wall of rectal muscularis
7 Endometriotic lesions infiltrating at least the muscular layer of the rectosigmoid colon; the most common form of bowel endometriosis
8 Endometriotic lesions infiltrate uterosacral ligaments unilaterally or bilaterally
9 Endometriotic lesions infiltrate vaginal wall, particularly posterior vaginal fornix
10 Defined when the peritoneum of the Pouch of Douglas is only partially or no longer visible during surgery and occurs as a result of adhesion formation; can be partial or complete, respectively
11 Endometriotic lesions infiltrating bladder muscularis propria
12 Ovarian cysts lined by endometrial tissue (endometrioma)
13 Reasons for downgrading the evidence can be found in Appendix J.6

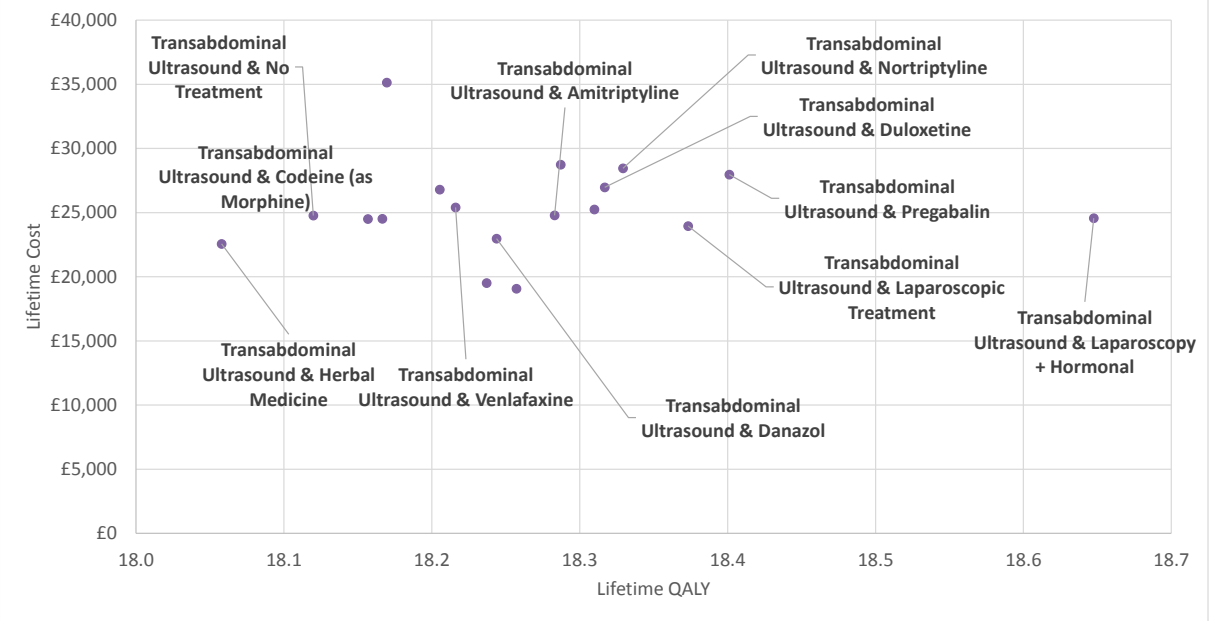
9.2.5 Economic evidence

A significant source of dissatisfaction with the current treatment pathway for endometriosis relates to the slow diagnosis and treatment of the condition. Consequently a de novo economic model was constructed to consider the optimal diagnosis and treatment strategies to attempt to increase the speed of accurate diagnosis in a cost-effective way. However, as the choice of diagnostic test depends in part on the choice of treatment (which is itself influenced by the availability of other diagnostic tests) it does not make sense to consider the 'cost-effectiveness' of one particular diagnostic strategy as though this were independent from the cost-effectiveness of other such strategies.

Figure 8 demonstrates how ultrasound interacts with various treatment options and Table 35 tabulates the same data. Ultrasound is highly likely to be cost-effective vs no treatment, and cost-effective when given in combination with the main treatment options of hormonal treatment and surgery.

ISBN 978-1-4731-2661-9

Figure 8: Costs and Lifetime QALYs of offering various treatment options in combination with ultrasound



Source: Economic model

Table 35: Costs and Lifetime QALYs of offering various treatment options in combination with ultrasound (showing only non-dominated strategies)

Treatment	Cost	QALY	ICER	Probability cost-effective vs no treatment (£20,000 / QALY)	Probability cost-effective vs no treatment (£30,000 / QALY)
Empirical Diagnosis & No Treatment	£22,752.60	18.120	Base Case	N/A	N/A
Transabdominal Ultrasound & Combined Oral Contraceptive Pill	£19,073.04	18.257	-£26,840.11	89.0%	89.0%
Transabdominal Ultrasound & Laparoscopic Treatment	£23,948.36	18.373	Extendedly Dominated	86.8%	87.9%
Transabdominal Ultrasound & Laparoscopy + Hormonal	£24,562.05	18.648	£14,058.31	85.7%	87.9%

ISBN 978-1-4731-2661-9

9.2.6 Clinical evidence statements

9.2.6.1 Pelvic endometriosis

Very low quality evidence from 5 studies (n=1,222, includes TVUS, tg-TVUS and kissing ovaries sign) found that the pooled sensitivity and specificity of ultrasound was 62% (18% to 94%) and 93% (78% to 99%).

9.2.6.2 Bowel endometriosis

Very low quality evidence from 3 studies (n=314, includes TVUS, RWC-TVUS and TVUS-BP) found the pooled sensitivity and specificity of 88% (70% to 97%) and 95% (85% to 99%). Very low quality evidence from 2 studies (n=171, includes TRUS) showed sensitivity and specificity of 88% (47% to 100%) and 96% (89% to 99%) and 97% (82% to 100%) and 100% (94% to 100%), respectively.

9.2.6.3 Deeply infiltrating endometriosis (DIE)

Very low quality evidence from 3 studies (n=282, includes TVUS, TVUS-BP and 3D-TVUS) found that the pooled sensitivity and specificity of ultrasound was 78% (37% to 97%) and 90% (58% to 99%).

9.2.6.4 Posterior DIE

Very low quality evidence from 7 studies (n=853, includes TVUS, tg-TVUS and SVG) showed that the pooled sensitivity and specificity of ultrasound was 73% (55% to 8%7) and 91% (76% to 98%). Another 2 studies (n=248, includes SVG and 3D-TVUS) found sensitivity of 91% (75% to 98%) and 7% (78% to 93%) and specificity of 86% (57% to 98%) and 94% (87% to 97%), respectively.

9.2.6.5 Anterior DIE

Low quality evidence from 1 study (n=88) found sensitivity and specificity of TVUS of 33% (13% to 59%) and 100% (95% to 100%).

9.2.6.6 Rectovaginal endometriosis

Very low quality evidence from 10 studies (n=983, includes TVUS, TVUS-BP, tg-TVUS, introital 3D-US and SVG) found that the pooled sensitivity and specificity of ultrasound was 66% (33% to 90%) and 98% (95% to 99%). Low quality evidence from 1 study (n=90) that used RWC-TVUS reported sensitivity of 97% (90% to 100%) and specificity of 100% (84% to 100%). Very low quality evidence from 2 studies (n=232, includes TRUS) found that the sensitivity and specificity of ultrasound was 18% (2% to 52%) and 97% (85% to 100%) and 95% (88% to 99%) and 96% (91% to 99%), respectively.

9.2.6.7 Rectosigmoid endometriosis

Very low quality evidence from 14 studies (n=1615, includes TVUS, TVUS-BP, tg-TVUS, RWC-TVUS and SVG) found that the pooled sensitivity and specificity of ultrasound was 89% (80% to 95%) and 96% (93% to 98%), respectively. 1 study (n=202, includes 3D-TVUS) reported sensitivity of 91% (82% to 96%) and specificity of 97% (92% to 99%). Evidence was of low quality. Very low quality evidence from 4 studies (n=330, includes TRUS) found the pooled sensitivity and specificity of 90% (77% to 98%) and 93% (79% to 99%).

ISBN 978-1-4731-2661-9

9.2.6.8 Uterosacral ligament endometriosis

Very low quality evidence from 7 studies (n=714, includes TVUS, tg-TVUS, TVUS-BP and SVG) found that the pooled sensitivity of ultrasound was 63% (45% to 79%) and the pooled specificity was 96% (91% to 98%). 2 studies (n=232, includes TRUS) reported sensitivity and specificity of 48% (37% to 59%) and 80% (44% to 97%) and 44% (14% to 79%) and 98% (93% to 100%), respectively.

9.2.6.9 Vaginal wall involvement

Very low quality evidence from 6 studies (n=679, includes TVUS, TVUS-BP, tg-TVUS and SVG) found that the pooled sensitivity and specificity of ultrasound was 57% (26% to 84%) and 98% (94% to 100%). Very low quality evidence from a further 2 studies (n=232) that used TRUS reported sensitivity of 7% (1% to 22%) and 100% (79% to 100%) and specificity of 100% (94% to 100%) and 100% (97% to 100%), respectively.

9.2.6.10 Pouch of Douglas

Very low quality evidence from 6 studies (n=755, includes TVUS, TVUS-BP and SVG+TVUS-BP) found that the pooled sensitivity and specificity of ultrasound was 83% (71% to 91%) and 97% (93% to 99%).

9.2.6.11 Bladder endometriosis

Very low quality evidence from 5 studies (n=383, includes TVUS, TVUS-BP, tg-TVUS, 3D-TVUS and SVG+TVUS-BP) reported the pooled sensitivity of 35% (13% to 63%) and specificity of 98% (96% to 100%).

9.2.6.12 Ovarian endometriosis

Low quality evidence from 9 studies (n=1066, includes TVUS, TVUS-BP and tg-TVUS) showed the pooled sensitivity of 90% (83% to 96%) and specificity of 96% (93% to 98%). One study (n=92, includes TRUS) reported sensitivity of 89% (74% to 97%) and specificity of 77% (64% to 87%).

9.2.7 Evidence to recommendations

9.2.7.1 Relative value placed on the outcomes considered

Sensitivity and specificity were considered proxies for patient outcomes (indicating a benefit from a true negative or true positive finding) and were prioritised as critical outcomes for this review. Although the Committee did not specify clinically important thresholds for these 2 diagnostic measures, the imprecision of estimates was assessed according to the confidence region around the pooled estimate in the summary ROC plots. Inconclusive results and test complications were also considered by the Committee.

Quality of life and other patient-reported outcomes were considered critical by the Committee but these data were not identified by the review.

9.2.7.2 Consideration of clinical benefits and harms

The consequences of testing are of great importance to women and delay in diagnosis of endometriosis due to false negative results is a well-recognised issue in this population. Not having a diagnosis, or having an incorrect negative diagnosis, can cause emotional distress. Women may assume, or be told, that their pain symptoms (such as dysmenorrhoea) are normal and assume that it is their inability to cope that is having a debilitating impact on their everyday lives. As such, a correct positive diagnosis of endometriosis may provide relief for

ISBN 978-1-4731-2661-9

women and improve their emotional wellbeing by validating their symptoms as arising from a pathological cause and providing reassurance that management via an appropriate care pathway will be initiated. A correct negative diagnosis establishes that a woman's symptoms are not due to endometriosis which enables the opportunity to promptly pursue investigation for other causes.

The Committee considered the accuracy of diagnosis that ultrasound scanning could provide. It should be noted that the clinical evidence in the review referred to studies from specialist and not community settings. In a community setting, many ultrasonographers have a general ultrasound certification, rather than specialist expertise in reviewing endometriosis. The Committee considered this likely to influence the accuracy of diagnosis and discussed how results of imaging need to be interpreted in light of the practitioner's level of training. They further noted that imaging reports may not be very specific to endometriosis and the GP (unless the GP had an interest in gynaecological issues) would then have to refer further to a gynaecologist.

For ultrasound services in specialist endometriosis services (centres), the Committee concluded that the health professional performing the procedure would have to be experienced in ultrasound with a specialist interest in endometriosis as it is not part of standard training. These ultrasound scans take the tenderness and mobility of tissues into account when interpreting the scan.

The Committee also acknowledged how the current use of ultrasound in the UK may involve different types of services. A referral for an ultrasound scan does not necessarily mean that a gynaecologist or the gynaecological service will see or treat the women. In current NHS practice, the referral could mean that results are interpreted by the ultrasonographer and then sent back to the GP without any further direct involvement. The Committee agreed that this practice could still be useful but highlighted that a negative ultrasound does not guarantee endometriosis is absent and if symptoms persist a further ultrasound by a more specialist scanning service should be considered. The evidence available was drawn from testing the different endometriosis sites. The Committee noted that overall the specificity was consistently high, however the sensitivity was heterogeneous.

Communication was considered to be highly important, especially regarding the GP communicating the diagnosis of endometriosis to women with suspected endometriosis.

The Committee concluded that in addition to changes in technology, training of the practitioner could also impact on imaging results, as well as the quality of the examination itself. However, it was agreed that the training of healthcare professionals was outside the scope of the guideline.

9.2.7.3 Consideration of economic benefits and harms

The model identifies ultrasound as being a useful intermediate step between empirical diagnosis (treating based on symptoms rather than definitive diagnosis) and laparoscopic confirmation, which tends to be quite expensive. This makes the model important for identifying whether the switchover from empirical diagnosis (which would be preferred at low cost/QALY thresholds) to laparoscopic confirmation (which would be preferred at higher cost/QALY thresholds) allows ultrasound to be the most cost-effective treatment at some intermediate thresholds. In the main health economic model, the strategies of MRI and ultrasound respond similarly to sensitivity analysis, and have similar cost and accuracy profiles (ultrasound less sensitive but more specific, and slightly cheaper). As sensitivity was a critical driver of cost-effectiveness, ultrasound tended to be extendedly dominated by MRI.

The Committee disagreed with the findings of the model, stressing that in their opinion the NHS Reference Cost overpriced a transabdominal ultrasound and underpriced a Pelvic MRI. As these values were used in the model this translated to an effective 'switching' of MRI and ultrasound in the order of cost-effectiveness. It is possible for both claims about the cost of

ISBN 978-1-4731-2661-9

MRI to be accurate at once; if clinicians operate in an environment where time on an MRI machine is scarce then this may genuinely shift the shadow price of an MRI scan for these clinicians while not altering the expected marginal cost of performing a scan (i.e. the cost of the machine divided by the number of scans it can be expected to do in its lifetime). Consequently the Committee agreed to leave the model unchanged to allow for a rationalisation of the price of MRI in the future, but make recommendations based on their clinical expertise of the price of an MRI.

The Committee are therefore recommending more ultrasound than is current practice. However, each of these ultrasound displaces a more expensive MRI. While there is disagreement about exactly how much money this saves the NHS (around £100 per scan based on NHS Reference Costs and around £400 per scan based on Committee opinion), there is no disagreement that this will therefore represent a net saving to the NHS and not a significant resource impact.

A fuller discussion of the economic benefits and harms of diagnostic strategies is located in the Health Economic Appendix K.

9.2.7.4 Quality of evidence

The quality of the evidence was very low to low according to GRADE criteria. This was mainly due to risk of bias (often the patient selection was not consecutive or random, not all patients were included in the analysis or studies were not blinded), inconsistency (particularly in relation to sensitivity estimates) and imprecision (with a high level of uncertainty as indicated by the confidence region in the pooled analysis).

The Committee discussed the validity of including studies published prior to 2003, as these would have used older ultrasound technology that may not be used in current practice. However, as a cut-off date had not been included in the protocol, these older studies were not excluded from the review. Many of the older studies would have focused on imaging of hard tissue, whereas more recent studies focus on soft tissue imaging because of the advancement in technology.

The Committee also noted differences in the terminology of defining endometriosis sites, for example, posterior pelvic endometriosis as a term used by clinicians, but which may refer to many sites.

9.2.7.5 Other considerations

Although the evidence showed that both ultrasound and MRI were reliable tests for identifying site specific endometriosis in a specialist setting, MRI could not be compared with ultrasound as women with endometriosis would not be sent for an MRI scan initially. However, if the ultrasound was inconclusive or negative, but deep endometrioses involving the bowel, bladder or ureter were suspected then women might be referred for an MRI scan. The Committee noted that there was also a cost implication, as MRI is a more expensive test than ultrasound.

The Committee also made a recommendation for women who may not be able to tolerate a transvaginal scan or where a transvaginal scan was not appropriate, for example, in women who have not had intercourse. In these circumstances a transabdominal ultrasound may be performed (with a full bladder) to visualise the pelvis; however, the Committee discussed the limitations of transabdominal scanning which is less accurate than transvaginal ultrasound. It was noted that a transrectal scan might be considered for women who could not tolerate a transvaginal scan as transrectal scanning is thought to have a similar accuracy to transvaginal scanning, but the Committee declined to make a recommendation on this. The Committee discussed how there were not many centres of expertise in transrectal or transperineal scanning for endometriosis in the UK so it would be an extremely expensive recommendation, and further noted that they did not want to give the impression that a

ISBN 978-1-4731-2661-9

woman who declined a transvaginal scan should be pressured into a transrectal or transperineal scan as women might find both options unacceptable.

9.2.7.6 Key conclusions

The Committee agreed that avoiding a delay in diagnosis is most important. If women suspected of having endometriosis had a negative ultrasound, endometriosis could not be ruled out as there was no certainty that these women would not have endometriosis and further investigation would need to be considered if symptoms persisted. They discussed that this also applies to abdominal or pelvic examination as well as MRI (please see section 9.4) and agreed that it was important that if suspicions remain women should be referred for further assessment. They therefore made an overarching recommendation to highlight this.

The evidence showed that a well performed ultrasound scan (in a specialist endometriosis service) accurately identified site specific endometriosis (for example, endometrioma, rectovaginal and rectocervical disease), but where endometriosis is superficial and spread across different sites throughout the pelvis it is less accurate.

9.2.8 Recommendations

General principle

17. **Do not exclude the possibility of endometriosis if the abdominal or pelvic examination, ultrasound or MRI are normal. If clinical suspicion remains or symptoms persist, consider referral for further assessment and investigation.**

Ultrasound

18. **Consider transvaginal ultrasound:**
    - to investigate suspected endometriosis even if the pelvic and/or abdominal examination is normal
    - to identify endometriomas and deep endometriosis involving the bowel, bladder or ureter.

19. **If a transvaginal scan is not appropriate, consider a transabdominal ultrasound scan of the pelvis.**

9.3 Biomarkers

9.3.1 Biomarker Cancer Antigen 125 (CA-125)

Review question: What is the accuracy of serum CA-125 in diagnosing endometriosis?

9.3.1.1 Introduction

A non-invasive diagnostic test for endometriosis could provide easier and quicker diagnosis and might allow the effects of treatment to be monitored. Numerous biochemical markers have been proposed and if these prove to be sufficiently accurate, a blood test could provide a safer and cheaper method of diagnosis that is accessible in community (GP) services. Biomarkers can be used to determine the prevalence of a condition in the population. Depending on their sensitivity and specificity, they may help inform the likelihood of the diagnosis suggested by other tests, or help exclude other conditions. They can also be

ISBN 978-1-4731-2661-9

utilised to determine the recurrence of a condition prior to symptoms returning. The possible usefulness of various biomarkers will be sought from the literature and their applicability to various clinical situations will be determined.

The aim of this review was to evaluate the accuracy of serum CA-125 for the diagnosis of endometriosis in women with suspected endometriosis.

For full details see review protocol in Appendix D, the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G.

Description of clinical evidence

One study was included in this review (Cochrane systematic review by Nisenblat 2016). 25 studies within the Cochrane systematic review were relevant (Barbati 1994; Bilibio 2014; Chen 1998; Colacurci 1996; Fedele 1989; Fereira 1994; Franchi 1993; Gagne 2003; Guerriero 1996; Hallamaa 2012; Harada 2002; Hornstein 1995; Koninckx 1996; Kurdoglu 2009; Lanzone 1991; Maiorana 2007; Martinez 2007; Mohamed 2013; Molo 1994; Muscatello 1992; Patton 1986; Somigliana 2004; Vigil 1999; Yang 1994; Zeng 2005) (Table 36). Studies that reported the results based on cancer antigen 125 (CA-125) cut-off threshold of ≥35 U/ml were included in the review. One study (Guerriero 1996) assessed serum CA-125 plasma levels in the diagnosis of endometrioma.

Of the included studies, 9 were from Italy (Barbati 1994; Colacurci 1996; Fedele 1989; Franchi 1993; Guerriero 1996; Lanzone 1991; Maiorana 2007; Muscatello 1992; Somigliana 2004), three from USA (Hornstein 1995; Molo 1994; Patton 1986), 2 from China (Yang 1994; Zeng 2005) and 1 each from Portugal (Fereira 1994), Belgium (Koninckx 1996), Finland (Hallamaa 2012), Spain (Martinez 2007), Turkey (Kurdoglu 2009), Canada (Gagne 2003), Brazil (Bilibio 2014), Chile (Vigil 1999), Japan (Harada 2002), Taiwan (Chen 1998) and Egypt (Mohamed 2013). In the majority of studies women were undergoing laparoscopy for various indications such as infertility, pelvic pain, pelvic or adnexal mass, dysmenorrhoea or a combination of these. The majority of studies provided details of performance of the CA-125 test.

The size of the population in each of the studies ranged from 35 (Molo 1994) to 368 (Gagne 2003) participants.

This review reports diagnostic accuracy outcomes such as sensitivity and specificity. No test-and-treat trials were identified, therefore no clinical or patient-reported outcomes such as quality of life were identified.

Evidence from the included studies are summarised in the clinical GRADE evidence profile below (Table 37 and Table 38). Modified GRADE was used to assess quality of outcomes. See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots and ROC plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G.

9.3.1.2 Summary of included studies

A summary of the studies that were included in this review are presented in Table 36.

Table 36: Summary of included studies

Study	Index test/reference standard	Population	Outcomes	Comments
Barbati 1994 (Nisenblat 2016 CSR)	• Serum CA-125 • Laparoscopy or laparotomy	Women undergoing laparotomy or diagnostic	Sensitivity and specificity	Serum CA-125 levels were measured by immunoradiometric

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Outcomes	Comments
Italy		laparoscopy for infertility or pelvic pain N=45		'one step' sandwich assay (IRMA CA-125 II K, Sorin Biomedica, Italy)
Bilibio 2014 (Nisenblat 2016 CSR) Brazil	• Serum CA-125 • Laparoscopy and histology	Women who underwent laparoscopy for infertility, pelvic pain or tubal ligation N=97	Sensitivity and specificity	Serum CA-125 levels were measured with Roche Diagnostics GmbH, Mannheim, Germany
Chen 1998 (Nisenblat 2016 CSR) Taiwan	• Serum CA-125 • Laparoscopy and histology	Women undergoing laparoscopy for dysmenorrhoea N=157 (consecutive)	Sensitivity and specificity	Serum CA-125 levels were measured by immunoradiometric assay ELISA-CA-125 II kit (GIF-SUR-YVETTE CEDEX, France)
Colacurci 1996 (Nisenblat 2016 CSR) Italy	• Serum CA-125 • Laparoscopy	Women undergoing laparoscopy for infertility N=45	Sensitivity and specificity	Serum CA-125 levels were measured by immunoradiometric 'two-step method' (IRMA-mat, Byk-Stangtee Diagnostic GmbH&Co Kgy, Dietzenbach)
Fedele 1989 (Nisenblat 2016 CSR) Italy	• Serum CA-125 • Laparoscopy and histology	Women undergoing laparoscopy for infertility, pelvic pain or both N=264	Sensitivity and specificity	Serum CA-125 levels were measured by immunoradiometric assay (Sorin Biomedica, Saluggia VC, Italy)
Fereira 1994 (Nisenblat 2016 CSR) Portugal	• Serum CA-125 • Laparoscopy or laparotomy and histology	Women scheduled for laparoscopy or laparotomy for investigation of infertility N=54	Sensitivity and specificity	Serum CA-125 levels were measured by ELISA (Cobas Core CA-125 II, EIA Roche 1992)
Franchi 1993 (Nisenblat 2016 CSR) Italy	• Serum CA-125 • Laparoscopy or laparotomy	Women of reproductive age undergoing laparotomy or laparoscopy for pelvic mass N=120	Sensitivity and specificity	Serum CA-125 levels were measured by radioimmunoassay
Gagne 2003 (Nisenblat 2016 CSR) Canada	• Serum CA-125 • Laparoscopy or laparotomy	Women scheduled to undergo laparoscopy or laparotomy N=368 (random)	Sensitivity and specificity	Serum CA-125 levels were measured by using a one step-sandwich radioimmunoassay (Fujirebio America Inc.)

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Outcomes	Comments
Guerriero 1996 (Nisenblat 2016 CSR) Italy	• Serum CA-125 • Laparoscopy or laparotomy and histology	Women undergoing laparoscopy or laparotomy for persistent adnexal mass N=101 (consecutive)	Endometrioma Sensitivity and specificity	Serum CA-125 levels were measured by immunoradiometric assay (CIS Bio International, Gif sur Yvette, France), limit of detection 0.5 U/ml
Hallamaa 2012 (Nisenblat 2016 CSR) Finland	• Serum CA-125 • Laparoscopy and histology	Women undergoing laparoscopy for suspected endometriosis or tubal ligation N=180	Sensitivity and specificity	Serum CA-125 levels were measured by ELISA (Fujirebio Diagnostics inc, Malvern, PA, USA)
Harada 2002 (Nisenblat 2016 CSR) Japan	• Serum CA-125 • Laparoscopy or laparotomy	Women who underwent laparotomy or laparoscopy with the preoperative diagnosis of infertility, myoma uteri, adenomyosis or endometriosis N=123	Sensitivity and specificity	Serum CA-125 levels were measured by enzyme immunoassay (TFB Co,Tokyo, Japan)
Hornstein 1995 (Nisenblat 2016 CSR) USA	• Serum CA-125 • Laparoscopy	Women with the preoperative diagnosis of endometriosis, pelvic pain, or infertility recruited from 2 fertility units N=123	Sensitivity and specificity	Serum CA-125 levels were measured by immunoradiometric assay (Centocor, Malvern, PA, USA)
Koninckx 1996 (Nisenblat 2016 CSR) Belgium	• Serum CA-125 • Laparoscopy	Women scheduled for laparoscopy for suspected endometriosis N=61 (consecutive)	Sensitivity and specificity	Serum CA-125 levels were measured by second generation IRMA kit (CA-125 II, Centocor, Malvern, Pa)
Kurdoglu 2009 (Nisenblat 2016 CSR) Turkey	• Serum CA-125 • Laparoscopy or laparotomy and histology	Women undergoing laparoscopy or laparotomy or various indications N=179	Sensitivity and specificity	Procedure of the index test not reported
Lanzone 1991 (Nisenblat 2016 CSR) Italy	• Serum CA-125 • Laparoscopy	Women undergoing laparoscopy for infertility or pelvic pain during luteal	Sensitivity and specificity	Serum CA-125 levels were measured by radioimmunoassay (CIS Diagnostici)

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Outcomes	Comments
		phase of the cycle N=270 (consecutive)		
Maiorana 2007 (Nisenblat 2016 CSR) Italy	• Serum CA-125 • Laparoscopy	Women who underwent laparoscopy for infertility, ovarian cyst or suspected endometriosis N=86	Sensitivity and specificity	Serum CA-125 levels were measured by enzyme immunoassay
Martinez 2007 (Nisenblat 2016 CSR) Spain	• Serum CA-125 • Laparoscopy	Women undergoing laparoscopy for various indications N=128	Sensitivity and specificity	Serum CA-125 levels were measured by enzyme immunoassay and were expressed in arbitrary units based on a primary reference standard
Mohamed 2013 (Nisenblat 2016 CSR) Egypt	• Serum CA-125 • Laparoscopy	Women referred for laparoscopy for unexplained primary infertility, chronic pelvic pain or both N=60	Sensitivity and specificity	Serum CA-125 levels were measured by ELISA kit for Can-Ag CA-125 (Fujirebio Diagnostics, Inc, Goteborg, Sweden)
Molo 1994 (Nisenblat 2016 CSR) USA	• Serum CA-125 • Laparoscopy	Women undergoing laparoscopy for infertility investigation N=35 (consecutive)	Sensitivity and specificity	Serum CA-125 levels were measured by radioimmunoassay (Contocor Inc, Malvern, PA)
Muscatello 1992 (Nisenblat 2016 CSR) Italy	• Serum CA-125 • Laparoscopy	Women who underwent laparoscopy for infertility, pelvic pain or both N=119	Sensitivity and specificity	Serum CA-125 levels were measured by using a commercially available radioimmunoassay (CIS Diagnostici)
Patton 1986 (Nisenblat 2016 CSR) USA	• Serum CA-125 • Laparoscopy and histology	Women who underwent laparoscopy N=113	Sensitivity and specificity	Serum CA-125 levels were measured by using radioimmunoassay (RIA)
Somigliana 2004 (Nisenblat 2016 CSR) Italy	• Serum CA-125 • Laparoscopy	Women who underwent gynaecologic laparoscopy for benign gynaecological pathologies N=80 (consecutive)	Sensitivity and specificity	Serum CA-125 levels were measured by commercially available chemiluminescent immunometric assay (Roche Diagnostics GmbH, Germany)

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Outcomes	Comments
Vigil 1999 (Nisenblat 2016 CSR) Chile	• Serum CA-125 • Laparoscopy and histology	Women who underwent laparoscopy for dysmenorrhoea and pelvic pain not responding to medical management, with or without infertility N=49	Sensitivity and specificity	Serum CA-125 levels were measured by the IRMA-COUNT OM-MA method
Yang 1994 (Nisenblat 2016 CSR) China	• Serum CA-125 • Laparoscopy	Women who underwent laparoscopy for infertility or suspected endometriosis N=42	Sensitivity and specificity	Serum CA-125 levels were measured by emission immunoassay kit (Syntron Biotech Co, USA)
Zeng 2005 (Nisenblat 2016 CSR) China	• Serum CA-125 • Laparoscopy or laparotomy	Women undergoing laparoscopy or laparotomy for pelvic pain, infertility or both N=58	Sensitivity and specificity	Serum CA-125 levels were measured by chemiluminescence assay

N: number of participants in study; CSR: Cochrane systematic review

9.3.1.3 Clinical evidence profile

The clinical evidence profile for this review question is presented in Table 37 and Table 38.

Table 37: Summary clinical evidence profile for diagnosis of endometriosis

Sensitivity (95%CI)	Specificity (95% CI)	No. of participants (no. of studies)	Quality of the evidence (GRADE)
38% (30 to 47)	92% (89 to 94)	2491 (24)	⊕⊖⊖⊖ Very low[1]

CI: confidence interval
15 studies did not use a consecutive or random sample, 10 studies did not pre-specify the threshold used and 5 studies did not include all patients in the analysis; unclear whether in 12 studies a consecutive or random sample of patients was used; unclear whether 3 studies avoided inappropriate exclusions; unclear whether in 13 studies the index test results was interpreted without knowledge of the results of the reference standard and whether in 4 studies the reference standard results were interpreted without knowledge of the results of the index test; unclear whether in 10 studies the reference standard was likely to correctly classify the target condition. In 8 studies there was high/unclear applicability concern in terms of population in so called "two-gate" design studies (according to Nisenblat 2016 Cochrane systematic review, a "two-gate" design study includes participants sampled from distinct populations with respect to clinical presentation; the same study includes participants with a clinical suspicion of having the target condition (e.g. women with pelvic pain) and also participants in whom the target condition is not suspected (e.g. women admitted for tubal ligation). "Two-gate" studies were included only where all cases and controls belonged to the same population with respect to the reference standard). Serious inconsistency.

Table 38: Summary clinical evidence profile for diagnosis of endometrioma

Sensitivity (95%CI)	Specificity (95% CI)	No. of participants (no. of studies)	Quality of the evidence (GRADE)
59% (39-76)	79% (68-88)	101 (1)	⊕⊕⊕⊖ Moderate[1]

CI: confidence interval
1 Unclear whether the index test result was interpreted without knowledge of the results of the reference standard

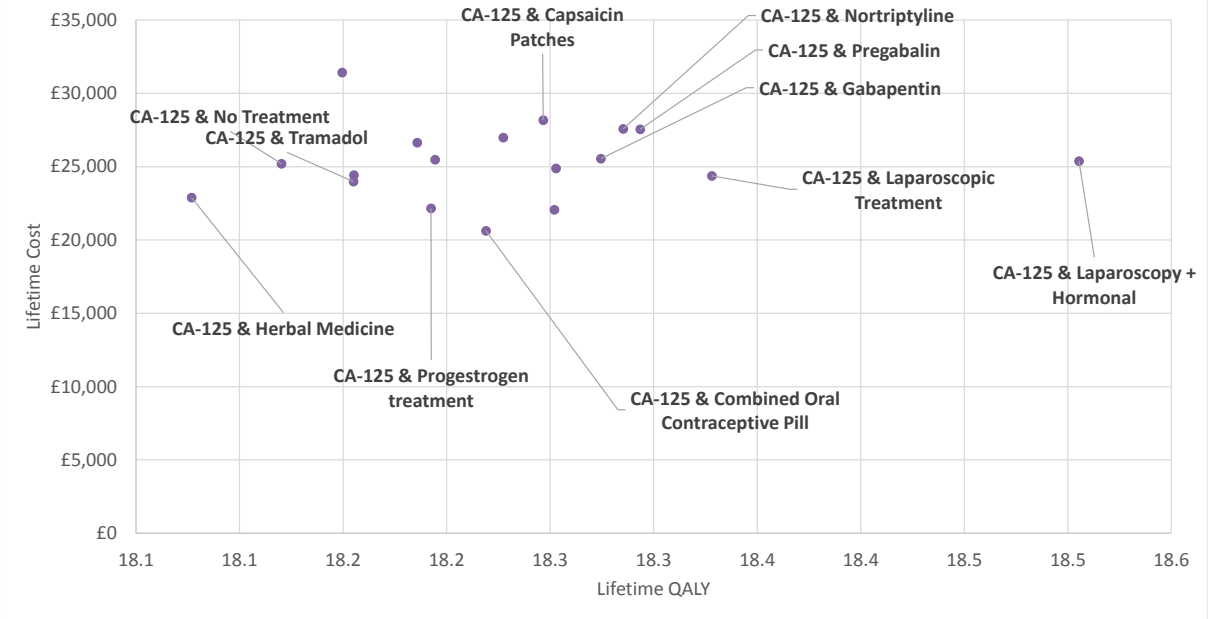
ISBN 978-1-4731-2661-9

9.3.1.4 Economic evidence

A significant source of dissatisfaction with the current treatment pathway for endometriosis relates to the slow diagnosis and treatment of the condition. Consequently a de novo economic model was constructed to consider the optimal diagnosis and treatment strategies to attempt to increase the speed of accurate diagnosis in a cost-effective way. However, as the choice of diagnostic test depends in part on the choice of treatment (which is itself influenced by the availability of other diagnostic tests) it does not make sense to consider the 'cost-effectiveness' of one particular diagnostic strategy as though this were independent from the cost-effectiveness of other such strategies. With CA-125 in particular, it would be of huge value to clinicians to have a cheap and non-invasive strategy to diagnose endometriosis, even if that strategy could only be used to justify the use of more expensive tests later. However the economic model suggests that CA-125 is currently too inaccurate to be used in this way.

Figure 9 demonstrates how CA-125 interacts with various treatment options and Table 39 tabulates the same data. Although CA-125 is relatively cost-effective and relatively likely to be cost-effective against no treatment, the average lifetime QALYs are quite low relative to more accurate diagnostic tests. Consequently in combination with other diagnostic tests, CA-125 tends to be dominated. In the case of infertile women, CA-125 and laparoscopic treatment is on the cost-effectiveness envelope, but is extendedly dominated by the same treatment with a more accurate diagnostic strategy.

Figure 9: Costs and Lifetime QALYs of offering various treatment options in combination with CA-125



Source: Economic model

ISBN 978-1-4731-2661-9

Table 39: Costs and Lifetime QALYs of offering various treatment options in combination with CA-125 (showing only non-dominated strategies)

Treatment	Cost	QALY	ICER	Probability cost-effective vs no treatment (£20,000 / QALY)	Probability cost-effective vs no treatment (£30,000 / QALY)
Empirical Diagnosis & No Treatment	£22,752.60	18.120	Base Case	N/A	N/A
CA-125 & Combined Oral Contraceptive Pill	£20,623.23	18.219	-£21,519.76	53.8%	53.8%
CA-125 & Danazol	£22,067.12	18.252	Extendedly Dominated	47.3%	48.4%
CA-125 & Laparoscopic Treatment	£24,377.37	18.328	Extendedly Dominated	76.9%	82.4%
CA-125 & Laparoscopy + Hormonal	£25,381.47	18.505	£13,084.94	70.3%	75.8%

9.3.1.5 Clinical evidence statements

Very low quality evidence from 24 studies (n=2491) showed that sensitivity and specificity of serum CA-125 in detecting endometriosis was 38% (30% to 47%) and 92% (89% to 94%).

Moderate quality evidence from 1 study (n=101) showed that sensitivity and specificity of serum CA-125 in detecting endometrioma was 59% (39% to 76%) and 79% (68% to 88%).

9.3.1.6 Evidence to recommendations

9.3.1.6.1 Relative value placed on the outcomes considered

As sensitivity and specificity reflect patient outcomes, these were prioritised as critical outcomes for this review. Although the Committee did not specify clinically important thresholds for these 2 diagnostic measures, the imprecision of estimates were assessed according to the confidence region around the pooled estimate in the ROC plots. Inconclusive results and test complications were also considered by the Committee. Quality of life was prioritised as an outcome if this were available. The Committee was particularly interested in the sensitivity of the test as high sensitivity would mean the cheaper CA-125 test (compared to imaging) is suitable for ruling out endometriosis at the first-line before a second, more specific test, is used to rule endometriosis in.

9.3.1.6.2 Consideration of benefits and harms

For the agreed cut-off of ≥35U/ml, the Committee agreed that serum CA-125 is not sensitive nor accurate enough to identify endometriosis. The Committee recognise that there are many women who have symptoms of endometriosis but do not have raised serum CA-125. In other words, the number of false negative results would be very high.

The specificity of serum CA-125 was high which means that women who have signs and symptoms of endometriosis but do not have a raised serum CA-125, are likely to be confirmed as not having endometriosis.

ISBN 978-1-4731-2661-9

Serum CA-125 may not be a sensitive marker, but a positive result will indicate women who truly have endometriosis. However, in current practice, women would not be diagnosed based on CA-125 testing alone. If they had signs and symptoms and an incidentally raised CA-125 levels, they would usually be referred for further diagnostic procedures such as an ultrasound scan, MRI or laparoscopy. The Committee therefore agreed that this test does not add anything to the diagnostic strategy, apart from a possible delay and additional costs for further unnecessary referral and investigation.

The Committee also discussed the possibility of making a recommendation to use this test in community (GP) services. However, this would potentially lead to many women being falsely reassured that they did not have endometriosis due to the large number of false negative results. Therefore the Committee discourage use of CA-125 testing in this setting.

9.3.1.6.3 Consideration of economic benefits and harms

Although CA-125 is by far the cheapest test available to diagnose endometriosis, its low accuracy means that it is failing to detect many cases of endometriosis, and accidentally diagnosing many cases of non-endometriosis

The Committee considered that CA-125 might be used in combination with other tests – either incidental information from the test could be used to help diagnose women who may have endometriosis, or the test itself could be used as a 'rule out' test to limit the number of women who needed to be diagnosed using more expensive methods. In both of these cases, the Committee decided that the information was potentially too misleading so despite the potential economic benefits of a cheap screening test recommended against paying too much account to the CA-125 results

The Committee recommended against using CA-125 to diagnose endometriosis, in line with current NHS practice. Consequently these recommendations are unlikely to carry a high resource impact.

A fuller discussion of the economic benefits and harms of diagnostic strategies is located in the Health Economic Appendix K.

9.3.1.6.4 Quality of evidence

The quality of the evidence was very low according to GRADE criteria. This was mainly due to risk of bias (often the patient selection was not consecutive or random, not all patients were included in the analysis or the serum CA-125 cut-off was not pre-specified) and inconsistency (particularly related to sensitivity estimates).

9.3.1.6.5 Other considerations

The Committee also discussed whether further evidence would reduce the uncertainty around the results; however, they concluded that there are many studies that investigate the diagnostic accuracy of serum CA-125 with a fairly consistent pattern of low sensitivity. The Committee therefore did not prioritise this topic for further research. The Committee considered whether additional recommendations were necessary for adolescent women but concluded that none were required.

9.3.1.6.6 Key conclusions

The Committee concluded that the serum CA-125 test would have too many false negative results to promote usage in clinical practice. However, if an incidental finding of raised serum CA-125 is reported in combination with signs and symptoms (for example, following investigation for ovarian cancer), it does raise the likelihood of women having endometriosis and further investigations would then be warranted.

ISBN 978-1-4731-2661-9

9.3.1.7 Recommendations

20. Do not use serum CA125 to diagnose endometriosis.

21. If a coincidentally reported serum CA125 level is available, be aware that:

- a raised serum CA125 (that is, 35 IU/ml or more) may be consistent with having endometriosis
- endometriosis may be present despite a normal serum CA125 (less than 35 IU/ml).

9.3.2 Biomarker Human Epididymis protein 4 (HE-4)

Review question: What is the accuracy of HE-4 in diagnosing endometriosis?

9.3.2.1 Introduction

HE-4 is a serum biomarker which has been used to detect epithelial ovarian cancer, often in conjunction with serum CA-125 testing. It is not currently used within the NHS as a diagnostic test for endometriosis. However it is an emerging technology that is sometimes offered to women outside the NHS.

The aim of this review was to evaluate the accuracy of HE-4 for the diagnosis of endometriosis in women with suspected endometriosis.

For full details see review protocol in Appendix D.

9.3.2.2 Description of clinical evidence

One study (Zhang 2014) was included in the review that examined the effectiveness of HE-4 (at a cut-off threshold of 114pM) to diagnose endometriosis or endometrioma in women (n=68) who had been diagnosed with pelvic mass and were scheduled for surgery (Table 40).

No test-and-treat trials were identified, therefore no clinical or patient-reported outcomes such as quality of life were reported.

Evidence from the included study is summarised in the clinical GRADE evidence profile below. See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, full GRADE profile in Appendix J and study evidence tables in Appendix G.

9.3.2.3 Summary of included studies

A summary of the studies that were included in this review are presented in Table 40.

Table 40: Summary of included studies

Study	Index test or reference standard	Population	Outcomes
Zhang 2014 China	• HE-4 • Surgery and histology	Women diagnosed with pelvic mass who were scheduled for surgery (N=68)	Specificity in detection of endometriosis or endometrioma

N: number of participants in study

ISBN 978-1-4731-2661-9

9.3.2.4 Clinical evidence profile

The clinical evidence profile for this review question is presented in Table 41.

Table 41: Summary clinical evidence profile for diagnosis of endometriosis

Sensitivity (95%CI)	Specificity (95% CI)	Test	No of participants (studies)	Quality of the evidence (GRADE)
0%	98% (90 to 100)	HE-4	68 (1)	⊕⊖⊖⊖ Very low [1]

CI: confidence interval

1 study was not blinded; unclear whether a consecutive or random sample was used, whether inappropriate exclusions were avoided and whether there was an appropriate interval between index test and reference standard.

9.3.2.5 Economic evidence

No economic evidence was found on the use of HE-4 in the diagnosis of endometriosis. As the clinical review found that the sensitivity of HE-4 as a biomarker was 0%, it was excluded from the health economic model on the grounds that it would significantly distort average results by failing to find any patient with endometriosis.

9.3.2.6 Clinical evidence statements

Very low quality evidence from 1 study (n=68) showed that at a cut off threshold of 114pM, specificity of HE-4 in diagnosing endometriosis/endometrioma in women with diagnosis of pelvic mass was 98% (90% to 100%) and sensitivity was 0%.

9.3.2.7 Evidence to recommendations

9.3.2.7.1 Relative value placed on the outcomes considered

As sensitivity and specificity reflect patient outcomes, these were prioritised as critical outcomes for this review. Inconclusive results and test complications were also considered by the Committee. Quality of life was prioritised as an outcome if this were available from test and treat RCTs.

The Committee was particularly interested in the sensitivity of HE-4 testing to rule out endometriosis as high specificity may indicate a useful and cheap (compared to imaging) first-line test.

9.3.2.7.2 Consideration of clinical benefits and harms

The clinical benefit of HE-4 as a diagnostic test is similar to that for other biomarker tests, in that it is cheap to perform. However, as it would not be used as a definitive diagnostic test, there would be other associated costs. If the test was positive, a diagnosis would require further diagnostic confirmation and, if it was negative, it may incur costs because women would have a delay in diagnosis which may lead to disease progression.

9.3.2.7.3 Consideration of economic benefits and harms

The HE-4 test is extremely cheap, but has no ability to detect endometriosis in a patient who actually has it. Consequently, the health economic model would find it to be incredibly expensive; more expensive than offering no treatment. For this reason it was excluded from the analysis, as it significantly distorted graphs and tables of final results.

ISBN 978-1-4731-2661-9

In real life there may be value in using the test as a cheap way to rule out endometriosis in patients in whom there is uncertainty about the diagnosis, although given that the usual concern is about whether an ovarian mass is a cancer or endometrioma it may be that the cost and QALY impact of a misdiagnosis are sufficiently severe that a more reliable technique is indicated; the economic model was not set up to answer this question.

As the Committee chose not to recommend the technique, this is not a departure from current practice in the NHS and consequently the resource impact will be minimal.

9.3.2.7.4 Quality of evidence

The evidence was limited to 1 small study with serious methodological flaws; the quality was very low. It examined the association between HE-4 level and different gynaecological pathologies and the assessment of HE-4 was performed after the diagnosis of endometriosis or endometrioma was already known. The cut-off was chosen based on the distribution of the sample, rather than specified *a priori*, which increases the risk of bias. The hypothesis was not specified *a priori* so it is unclear whether the authors were intending to use raised HE-4 levels to diagnose endometriosis or low HE-4 levels to exclude it.

The Committee noted that there was high specificity in the study which might indicate this test was useful for ruling in a diagnosis of endometriosis. Also the very low (0%) sensitivity may be useful for ruling out endometriosis in cases where there is uncertainty whether a complex ovarian mass is a potential ovarian malignancy or endometrioma. This may help ensure women are seen by the most appropriate specialist. However, because of the study limitations described above, there was considerable uncertainty in the available evidence to base a recommendation on this finding as well as that for sensitivity.

9.3.2.7.5 Other considerations

The Committee noted that HE-4 was not used in current clinical practice for the detection of endometriosis and if used for the detection of ovarian cancer, testing would be in the context of parallel serum CA-125 testing. As the Committee did not recommend serum CA-125 testing for women with suspected endometriosis, this further persuaded the Committee that no clinical or research recommendation should be made to support HE-4 testing.

9.3.2.7.6 Key conclusions

The Committee concluded that there was no evidence to support a recommendation for HE-4 for the diagnosis of endometriosis or endometrioma in women with suspected endometriosis.

9.3.2.8 Recommendations

No recommendation was made.

9.3.3 Biomarkers in endometrial tissues (the nerve fibre marker Protein Gene Product 9.5)

Review question: What is the accuracy of biomarkers in endometrial tissue, such as the nerve fibre marker Protein Gene Product 9.5 (PGP 9.5) in diagnosing endometriosis?

9.3.3.1 Introduction

Nerve fibres are present in the basal layer of the endometrial lining and grow with blood vessels into the functional layer as it grows during each menstrual cycle. It has been postulated that these small nerve fibres may be associated with menstrual pain. The nerve fibres are not identifiable with routine histological staining therefore immunohistochemistry, using PGP 9.5, is required to detect them.

ISBN 978-1-4731-2661-9

The aim of this review was to evaluate the accuracy of the nerve fibre marker PGP 9.5 for the diagnosis of endometriosis in women with suspected endometriosis. Although it is not current NHS practice to use this test, if it provided a sufficiently accurate diagnosis it may present a relatively non-invasive technique to diagnose abdomino-pelvic endometriosis.

For full details see review protocol in Appendix D.

9.3.3.2 Description of clinical evidence

One Cochrane systematic review (Gupta 2016) was included. Eight studies within the Cochrane review were relevant (Al-Jefout 2007; Al-Jefout 2009; Bokor 2009; Elgafor el Sharkwy 2013; Leslie 2013; Makari 2012; Meibody 2011; Yadav 2013) (Table 42).

Of the included studies, 3 were from Australia (Al-Jefout 2007; Al-Jefout 2009; Leslie 2013) and 1 each from Belgium (Bokor 2009), Lithuania (Makari 2012), Iran (Meibody 2011, India (Yadav 2013) and Egypt (Elgafor el Sharkwy 2013).

This review reports diagnostic accuracy outcomes such as sensitivity and specificity. No test-and-treat trials were identified, therefore no patient-reported outcomes such as quality of life were reported. The size of the population in each of the studies ranged from 20 (Makari 2012) to 114 (Elgafor el Sharkwy 2013). Studies included women undergoing laparoscopy for suspected endometriosis or for infertility, pelvic pain or both. Menstrual cycle phase details were available for 6 studies (Table 42).

Evidence from the included study is summarised in the clinical GRADE evidence profile below (Table 43. Modified GRADE was used to assess quality of outcomes. See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G.

9.3.3.3 Summary of included studies

A summary of the studies that were included in this review are presented in Table 42.

Table 42: Summary of included studies

Study	Index test/reference standard	Population	Outcomes	Comments
Al-Jefout 2007 (Gupta 2016 CSR) Australia	• Endometrial nerve fibres PGP 9.5 • Laparoscopy and histology	Reproductive-aged women undergoing laparoscopy for suspected endometriosis or infertility N=37	Sensitivity and specificity	Menstrual cycle phase not specified
Al-Jefout 2009 (Gupta 2016 CSR) Australia	• Endometrial nerve fibres PGP 9.5 • Laparoscopy and histology	Reproductive-aged women undergoing laparoscopy for infertility, pelvic pain or both N=103	Sensitivity and specificity	Menstrual cycle phase n=15; proliferative n=39; mid-cycle n=14; secretory n=31
Bokor 2009 (Gupta 2016 CSR) Belgium	• Endometrial neural marker PGP 9.5 • Laparoscopy and histology	Endometrial samples selected from tissue bank, which were collected from women undergoing laparoscopies for infertility, pain or both N=40	Sensitivity and specificity	All women in secretory phase of menstrual cycle

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Outcomes	Comments
Elgafor el Sharkwy 2013 (Gupta 2016 CSR) Egypt	• Endometrial nerve fibres PGP 9.5 • Laparoscopy	Women undergoing laparoscopy for infertility, pelvic pain or both N=114	Sensitivity and specificity	All women in follicular cycle phase
Leslie 2013 (Gupta 2016 CSR) Australia	• Endometrial functional layer nerve fibres PGP 9.5 • Laparoscopy and histology	Women undergoing laparoscopy for suspected endometriosis N=68	Sensitivity and specificity	Menstrual cycle phase n=25 in proliferative, n=19 in secretory cycle phase; n=24 unclear/hormonal treatment; Endometrial sampling was usually performed using a metal curette. 9 women were receiving oral contraceptive treatment and 2 women were receiving gonadotrophin-releasing hormone antagonists at the time of surgery
Makari 2012 (Gupta 2016 CSR) Lithuania	• Endometrial nerve fibres PGP 9.5 • Laparoscopy and histology	Women that presented for laparoscopy for infertility, pelvic pain or both N=20	Sensitivity and specificity	N=15 in proliferative and n=5 in secretory cycle phase
Meibody 2011 (Gupta 2016 CSR) Iran	• Endometrial small nerve fibres in eutopic endometrium PGP 9.5 • Laparoscopy/laparotomy and histology	Women undergoing laparoscopy or laparotomy for infertility or pelvic pain N=27	Sensitivity and specificity	All women in proliferative cycle phase
Yadav 2013 (Gupta 2016 CSR) India	• Endometrial nerve fibres PGP 9.5 • Laparoscopy and histology	Women who underwent laparoscopy for infertility or pelvic pain or suspected endometriosis N=60	Sensitivity and specificity	Cycle phase not specified

N: number of participants in study; CSR: Cochrane systematic review

9.3.3.4 Clinical evidence profile

The clinical evidence profile for this review question is presented in Table 43.

ISBN 978-1-4731-2661-9

Table 43: Summary clinical evidence profile for diagnosis of endometriosis

Sensitivity (95%CI)	Specificity (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
88% (69 to 98)	81% (69 to 91)	429 (8)	⊕⊖⊖⊖ Very low[1]

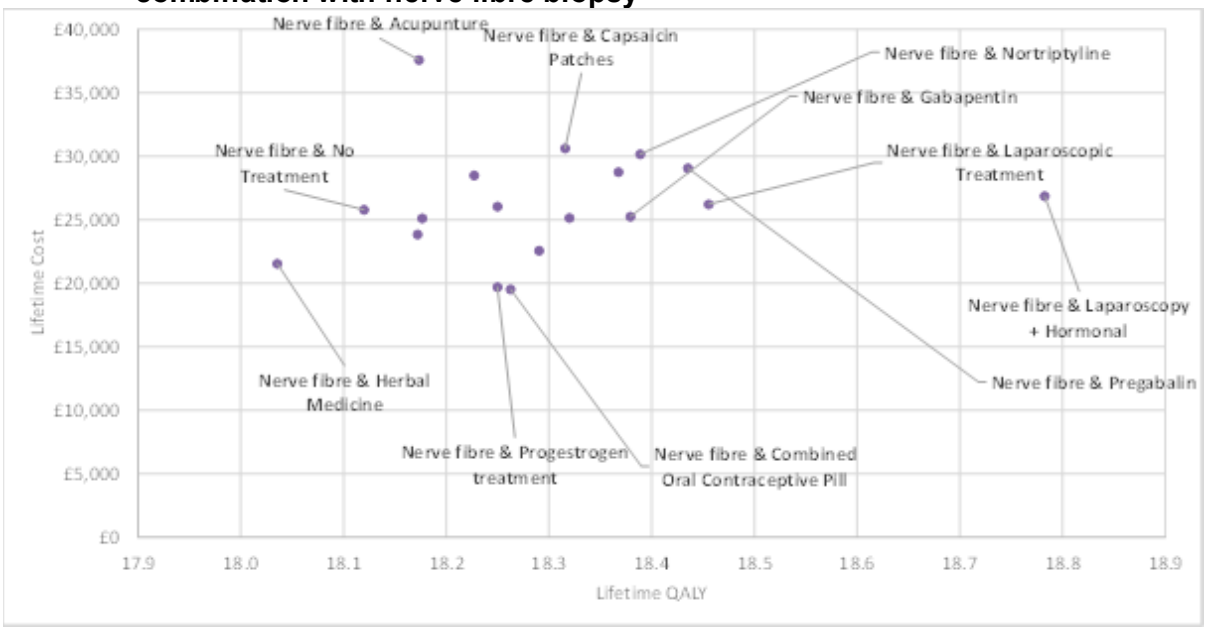
CI: confidence interval
1 5 studies did not use a consecutive or random sample, 1 study did not pre-specified the threshold used and 1 study did not include all patients in the analysis; unclear whether in 1 study a consecutive or random sample of patients was used; unclear whether 2 studies were blinded. In 3 studies there was high/unclear applicability concern in terms of population in so called "two-gate" design studies (according to Gupta 2016 Cochrane systematic review, a "two-gate" design study includes participants sampled from distinct populations with respect to clinical presentation; the same study includes participants with a clinical suspicion of having the target condition (e.g. women with pelvic pain) and also participants in whom the target condition is not suspected (e.g. women admitted for tubal ligation). "Two-gate" studies were included only where all cases and controls belonged to the same population with respect to the reference standard). Serious inconsistency and imprecision.

9.3.3.5 Economic evidence

A significant source of dissatisfaction with the current treatment pathway for endometriosis relates to the slow diagnosis and treatment of the condition. Consequently a de novo economic model was constructed to consider the optimal diagnosis and treatment strategies to attempt to increase the speed of accurate diagnosis in a cost-effective way. However, as the choice of diagnostic test depends in part on the choice of treatment (which is itself influenced by the availability of other diagnostic tests) it does not make sense to consider the 'cost-effectiveness' of one particular diagnostic strategy as though this were independent from the cost-effectiveness of other such strategies.

Figure 10 demonstrates how nerve fibre biopsy interacts with various treatment options and Table 44 tabulates the same data. In particular, they demonstrate how uncertain the findings on nerve fibres are in practice; even a strategy which is highly cost-effective on average at £20,000 / QALY (such as nerve fibre biopsy and laparoscopic treatment with adjunct hormonal therapy) still only has a 31% chance of being more cost-effective than doing nothing.

Figure 10: Costs and Lifetime QALYs of offering various treatment options in combination with nerve fibre biopsy



Source: Economic model

ISBN 978-1-4731-2661-9

Table 44: Costs and Lifetime QALYs of offering various treatment options in combination with nerve fibre biopsy (showing only non-dominated strategies)

Treatment	Cost	QALY	ICER	Probability cost-effective vs no treatment (£20,000 / QALY)	Probability cost-effective vs no treatment (£30,000 / QALY)
Empirical Diagnosis & No Treatment	£22,752.60	18.120	Base Case	N/A	N/A
Nerve fibre & Combined Oral Contraceptive Pill	£19,528.03	18.263	-£22,635.17	81.3%	91.2%
Nerve fibre & Danazol	£22,583.04	18.290	Extendedly Dominated	52.7%	78.0%
Nerve fibre & Amitriptyline	£25,146.88	18.320	Extendedly Dominated	79.1%	82.4%
Nerve fibre & Gabapentin	£25,258.85	18.379	Extendedly Dominated	37.4%	65.9%
Nerve fibre & Laparoscopic Treatment	£26,222.93	18.455	Extendedly Dominated	34.1%	58.2%
Nerve fibre & Laparoscopy + Hormonal	£26,875.57	18.783	£4,006.35	30.8%	72.5%

9.3.3.6 Clinical evidence statements

Very low quality evidence from 8 studies (n=429) reported that sensitivity and specificity of PGP 9.5 for detection of endometriosis was 88% (69% to 98%) and 81% (69% to 91%).

9.3.3.7 Evidence to recommendations

9.3.3.7.1 Relative value placed on the outcomes considered

As sensitivity and specificity are proxies for patient level outcomes, these were prioritised as critical outcomes for this review. No test and treat randomised controlled trials which would directly report patient level outcomes (such as health related quality of life) were identified. Inconclusive results and test complications were also considered by the Committee.

9.3.3.7.2 Consideration of clinical benefits and harms

The Committee agreed that their priority was early detection and treatment of endometriosis to prevent disease progression and to enable early clinical management.

The Committee discussed the importance of reducing the likelihood of a false negative diagnosis which could result in the woman not receiving effective management and the potential additional negative psychological impact of a false negative diagnosis if a woman was experiencing painful symptoms. However they noted that a false positive result might lead to unnecessary treatment (and associated costs) and also result in a negative psychological impact.

ISBN 978-1-4731-2661-9

9.3.3.7.3 Consideration of economic benefits and harms

The model did not identify nerve fibre biopsy as being notably likely to be cost-effective at any reasonable cost/QALY threshold. This is due in part to the fact that it is somewhat more expensive than other tests without compensating accuracy and due in another part to the fact that better tests exist which are preferred at the NICE threshold of £20,000 / QALY.

The Committee described how some very knowledgeable specialists in endometriosis believed the test might evolve into a cheap replacement for more expensive scans with a little more development; should the evidence for nerve fibres change it would not require very much more accuracy before it became cost-effective as a potential substitute for MRI or ultrasound. However on the evidence that was available during development, the Committee believed the economic model was finding the correct results.

As the Committee are not recommending using nerve fibres in the diagnosis of endometriosis and this is a relatively new technique (that is not yet current practice) these recommendations do not carry a significant resource impact.

A fuller all discussion of the economic benefits and harms of diagnostic strategies is located in the Health Economic Appendix K.

9.3.3.7.4 Quality of evidence

The quality of the evidence was very low according to GRADE criteria. This was due to risk of bias (often the patient selection was not consecutive or random or it was unclear whether studies were blinded), inconsistency as well as imprecision with a high level of uncertainty as indicated by the confidence region in the pooled analysis. Although the Committee did not specify clinically important thresholds for sensitivity and specificity diagnostic measures, the imprecision of estimates were assessed according to the confidence region around the pooled estimate in the ROC plot. One study differed from all other studies in terms of population as it included women on hormonal treatment at the time of surgery. Studies also differed regarding the timing of the index test as women were in various phases of the menstrual cycle.

9.3.3.7.5 Other considerations

The Committee agreed that neither a recommendation nor a research recommendation would be appropriate at this point. They discussed and agreed that PGP 9.5 was not specific as a diagnostic tool for identifying endometriosis. The following points were raised and agreed by the Committee:

- Nerve fibres can be found in normal tissues and furthermore increased density in non-endometriotic pathologies such as adenomyosis. It is therefore not a specific test.
- Currently appropriate samples may need the functional layer to be present which means that the procedure is not completely non-invasive.
- PGP 9.5 is not usually utilised in most laboratories and keeping it ‘just in case’ would mean both greater expense and danger of degeneration due to infrequent use.
- It would mean a change in current practice (to a practice that is currently not validated) with the methodology being expensive and not available everywhere. The available evidence has not included any costings of this.
- The methodology is not standardised.
- Despite already having been researched for 9 to 10 years the technique has not been adopted because it is not used specifically to identify endometriosis.

They therefore agreed that there was insufficient validation and evaluation of this method which requires standardisation in terms of sample taking and size and the laboratory methodologies which are not universal across laboratories in the UK.

ISBN 978-1-4731-2661-9

The option of a research recommendation was discussed but the Committee agreed that this methodology would never be specific enough to diagnose endometriosis and therefore further research would, most likely, not provide evidence that would support a positive or negative recommendation for this.

9.3.3.7.6 Key conclusions

The Committee decided not to make a recommendation or a research recommendation based on their discussion about PGP 9.5. This is mainly due to the fact that this methodology in not specific as a diagnostic tool to detect endometriosis. It was agreed that as a method of testing it requires standardisation in methodology, it is not routinely used in current practice, it is not conclusively validated and utilised in most laboratories and is expensive.

9.3.3.8 Recommendations

No recommendation was made.

9.3.3.9 Research recommendations

No research recommendation was made.

9.4 Magnetic resonance imaging (MRI)

Review question: What is the accuracy of MRI in diagnosing endometriosis?

9.4.1 Introduction

MRI is a non-invasive test for the diagnosis of endometriosis and, if it is accurate, it could lead to the diagnosis without the need for a surgical procedure or it could decrease the need for it.

The aim of this review was to evaluate the accuracy of MRI for the diagnosis of endometriosis in women with suspected endometriosis.

For full details see review protocol in Appendix D.

9.4.2 Description of clinical evidence

Two studies were included in this review. Evidence was available from 1 Cochrane systematic review (Nisenblat 2016) and 1 observational study (Arrive 1989). Seventeen studies within the Cochrane systematic review were relevant (Abrao 2007, Ascher 1995, Bazot 2009, Bazot 2013, Biscaldi 2014, Chamie 2009, Ha 1994, Hottat 2009, Grasso 2010, Manganaro 2012a, Manganaro 2012b, Manganaro 2013, Okada 1995, Stratton 2003, Sugimura 1993, Takeuchi 2005, Thomeer 2014). Three studies compared more than 1 MRI method in the same cohort of women (Acher 1995, Bazot 2013, Ha 1994) (Table 45).

Of the included studies, 5 were from Italy (Grasso 2010, Biscaldi 2014, Manganaro 2012a, Manganaro 2012b, Manganaro 2013), 3 from USA (Arrive 1989, Ascher 1995, Stratton 2003), 2 from France (Bazot 2009, Bazot 2013), Brazil (Abrao 2007, Chamie 2009) and Japan (Sugimura 1993, Takeuchi 2005), and 1 each from the Netherlands (Thomeer 2014), Belgium (Hottat 2009), Japan (Okada 1995) and Korea (Ha 1994).

The size of the population in each of the studies ranged from 19 (Manganaro 2012b) to 260 (Biscaldi 2014).

The majority of studies used T1/T2-w MRI (Abrao 2007, Arrive 1989, Asher 1995, Ha 1994, Stratton 2003, Sugimura 1993), other studies used T1/T2-w + fat- suppressed/Gd MRI

ISBN 978-1-4731-2661-9

(Ascher 1995, Bazot 2009, Chamie 2009, Grasso 2010, Stratton 2003), T1/T2-w + fat-suppressed MRI (Ascher 1995, Ha 1994, Takeuchi 2005), T1/T2-w + fat- suppressed, jelly method MRI (Biscaldi 2014), 2D FSE T2-w MRI (Bazot 2013), T1-w fat-saturated MRI (Okada 1995), 3.0T MRI (Hottat 2009, Manganaro 2012a, Manganaro 2012b, Manganaro 2013, Thomeer 2014) or 3D MRI (Bazot 2013).

This review reports diagnostic accuracy outcomes such as sensitivity and specificity. No test-and-treat trials were identified, therefore no patient-reported outcomes such as quality of life were reported.

Evidence from the included studies are summarised in the clinical GRADE evidence profile below (Table 46). Modified GRADE was used to assess quality of outcomes. See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots and ROC plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G.

9.4.3 Summary of included studies

A summary of the studies that were included in this review are presented in Table 45.

Table 45: Summary of included studies

Study	Index test/reference standard	Population	Type of endometriosis/outcomes
Abrao 2007 (Nisenblat 2016 CSR) Brazil	• MRI (T1/T2-w) • Laparoscopy and histology	Women with clinically suspected endometriosis N=104	DIE sites: rectovaginal septum, recto-sigmoid involvement Sensitivity and specificity
Arrive 1989 USA	• MRI (T1/T2-w) • Laparoscopy or laparotomy and histology	Women with clinically suspected endometriosis N=30	Pelvic endometriosis Sensitivity and specificity
Ascher 1995 (Nisenblat 2016 CSR) USA	• MRI (T1/T2-w, T1/T2-w + fat- suppressed, T1/T2-w + fat-suppressed/Gd) • Laparoscopy or laparotomy	Women with clinically suspected endometriosis who were scheduled for surgery N=38	Pelvic endometriosis Sensitivity and specificity
Bazot 2009 (Nisenblat 2016 CSR) France	• MRI (T1/T2-w + fat-suppressed/Gd) • Laparotomy or laparoscopy and histopathology	Women referred with clinical evidence of pelvic endometriosis N=92	DIE DIE sites: rectovaginal septum, rectosigmoid involvement, uterosacral ligament, vaginal, ovarian endometriosis Sensitivity and specificity
Bazot 2013 (Nisenblat 2016 CSR) France	• MRI (2D FSE T2-w, 3D) • Laparotomy or Laparoscopy and histology	Women referred for pelvic MRI because of clinical suspicion of endometriosis N=110	DIE DIE sites: rectosigmoid involvement, uterosacral ligament, vaginal, pouch of Douglas Sensitivity and specificity
Biscaldi 2014	• MRI (jelly method 1/T2-w + fat- suppressed)	Women referred to endometriosis centre N=260	DIE site: rectosigmoid involvement Sensitivity and specificity

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Type of endometriosis/outcomes
(Nisenblat 2016 CSR) Italy	• Laparoscopy and histology		
Chamie 2009 (Nisenblat 2016 CSR) Brazil	• MRI (T1/T2-w + fat-suppressed/Gd) • Laparoscopy and histology	Women who had a history and findings of a physical exam consistent with endometriosis N=92	DIE sites: rectovaginal septum, rectosigmoid involvement, vaginal, ureteral, bladder* Sensitivity and specificity
Grasso 2010 (Nisenblat 2016 CSR) Italy	• MRI (T1/T2-w + fat-suppressed/Gd) • Laparoscopy and histology	Women with clinically suspected endometriosis N=33	DIE Sensitivity and specificity
Ha 1994 (Nisenblat 2016 CSR) Korea	• MRI (T1/T2-w, T1/T2-w + fat-suppressed) • Laparoscopy	Women with clinically suspected endometriosis N=31	Pelvic endometriosis Sensitivity and specificity
Hottat 2009 (Nisenblat 2016 CSR) Belgium	• MRI (3.0T) • Laparoscopy or laparotomy and histology	Women with clinically suspected endometriosis N=106	DIE DIE sites: rectosigmoid involvement, uterosacral ligament, vaginal, pouch of Douglas, anterior DIE Sensitivity and specificity
Manganaro 2012a (Nisenblat 2016 CSR) Italy	• MRI (3.0T) • Laparoscopy	Women with clinical ± sonographic suspicion of endometriosis N=46	Pelvic endometriosis DIE DIE site: uterosacral ligament ovarian endometriosis Sensitivity and specificity
Managaro 2012b (Nisenblat 2016 CSR) Italy	• MRI (3.0T) • Laparoscopy	Women with clinical ± sonographic suspicion of endometriosis N=19	DIE site: pouch of Douglas sensitivity and specificity
Manganaro 2013 (Nisenblat 2016 CSR) Italy	• 3.0T MRI • Laparoscopy and histology	Women with suspected USL DIE based on clinical symptoms, abnormal gynaecological examination or transvaginal ultrasound findings N=42	DIE site: uterosacral ligament Sensitivity and specificity
Okada 1995 (Nisenblat 2016 CSR) Japan	• T1-w fat-supressed • Laparoscopy or laparotomy and histology	Women visiting outpatient department with suspected endometriosis based on clinical presentation (symptoms and pelvic examination), transvaginal ultrasonography and/or blood test for Ca-125 N=74	Pelvic endometriosis Sensitivity and specificity

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Type of endometriosis/outcomes
Stratton 2003 (Nisenblat 2016 CSR) USA	• MRI (T1/T2-w, T1/T2-w + fat-suppressed/Gd) • Laparoscopy and histology	Women with pelvic pain, who were otherwise in good health, were evaluated to exclude other causes of pain N=58	Pelvic endometriosis Sensitivity and specificity
Sugimura 1993 (Nisenblat 2016 CSR) Japan	• MRI (T1/T2-w) • Laparoscopy or laparotomy and histology	Women with clinically suspected endometriosis N=35	Pelvic endometriosis Sensitivity and specificity
Takeuchi 2005 (Nisenblat 2016 CSR) Japan	• MRI (T1/T-w + fat-suppressed) • Laparoscopy and histology	Women scheduled to undergo laparoscopy for suspected rectovaginal endometriosis based on clinical symptoms, rectal/pelvic examination findings and preoperative sonographic examination results N=31	DIE Sensitivity and specificity
Thomeer 2014 (Nisenblat 2016 CSR) Netherlands	• MRI (3.0T) • Laparoscopy	Women with clinically suspected endometriosis scheduled to undergo laparoscopy N=40	Pelvic endometriosis Sensitivity and specificity

N: number of participants in study; CSR, Cochrane systematic review; DIE, deeply infiltrating endometriosis
**bladder data from the original study*
MRI types as defined in Nisenblat Cochrane Systematic Review 2016:

- *T1/T2-w MRI: includes axial spin-echo or gradient echo T1-weighted (T1-w) images followed by fast spin-echo (FSE)/turbo spin-echo (TSE) images or fast relaxation fast-spin echo (FR-FSE) T2-w images*
- *T1/T2-w + fat-supressed MRI: includes T1-w imaging using chemical fat suppression, which aids in the differentiation of lipid and haemorrhagic pathologies*
- *T1/T2-w + fat-supressed MRI/Gd: includes gradient echo T1 images with and without fat suppression followed by FSE or FR-FSE T2-w images before and after intravenous injection of the paramagnetic contrast agent gadolinium*
- *Jelly method 1/T2-w + fat- suppressed: involves pre-treatment of participants for MRI by simultaneous injection of ultrasonographic gel into the vagina (˜ 50 mL) and into the rectum (150 mL gel 50% diluted with water). Another technique evolves introduction of 300-400 mL of diluted ultrasonographic gel (1:8 dilution) for rectosigmoid distension without use of intravaginal gel*
- *3D MRI: includes 3D coronal single-slab (containing all the slices) MRI, entitled 'CUBE' with FSE T2-w images. The technique involves using variable flip angle refocusing, auto-calibrating, 2D accelerated parallel imaging and nonlinear view ordering to produce high-resolution volumetric image data sets and to reduce imaging time by using multi-planar reformations*
- *3.0T MRI: 3.0Tesla Magnetom system with a multi-channel phased-array surface body-coil*

9.4.4 Clinical evidence profile

The clinical evidence profile for this review question is presented in Table 46.

ISBN 978-1-4731-2661-9

Table 46: Summary clinical evidence profile for diagnosis of endometriosis

Sensitivity (95%CI)	Specificity (95% CI)	Site of endometriosis (MRI test)	No of participants (studies)	Quality of the evidence (GRADE)[13]
77% (62 to 88)	72% (53 to 87)	Pelvic[1] (T1-/T2-w, T1-w+fat-supressed, T-1/T2-w + fat-suppressed/Gd and 3.0T MRI)	333 (8)	⊕⊖⊖⊖ Very low
86% (64 to 97) 76% (56 to 90)	50% (19 to 81) 100% (16 to 100)	Pelvic[1] (T1-/T2-w + fat-suppressed and fat-suppressed MRI)	62 (2)	⊕⊖⊖⊖ Very low
81% (58 to 95)	50% (19 to 81)	Pelvic[1] (T-1/T2-w + fat-suppressed/Gd MRI)	31 (1)	⊕⊖⊖⊖ Very low
96% (90 to 99)	86% (54 to 98)	DIE[2] (T-1/T2-w + fat-suppressed/Gd and 3.0T MRI)	212 (4)	⊕⊖⊖⊖ Very low
89% (65 to 99) 94% (71 to 100)	20% (1 to 72) 100% (77 to 100)	Posterior DIE[3] (2D FSE T2-w MRI and jelly method T1-/T2-w + fat-suppressed)	54 (2)[14]	⊕⊖⊖⊖ Very low
100% (81 to 100)	20% (1 to 72)	Posterior DIE[3] (3D MRI)	23 (1)[14]	⊕⊖⊖⊖ Very low
75% (35 to 97)	100% (89 to 100)	Anterior[4] DIE (3.0T MRI)	41 (1)	⊕⊖⊖⊖ Very low
75% (35 to 95)	88% (43 to 99)	Rectovaginal[5] (T-1/T2-w + fat-suppressed/Gd MRI)	288 (3)	⊕⊖⊖⊖ Very low
91% (79 to 97)	96% (92 to 99)	Rectosigmoid[6] (T-1/T2-w + fat-suppressed/Gd, 2D FSE T2-w, jelly method (T1-/T2-w + fat-suppressed) and 3.0T MRI)	662 (6)	⊕⊖⊖⊖ Very low
85% (55 to 98)	90% (55 to 100)	Rectosigmoid[6] (3D MRI)	23 (1)	⊕⊖⊖⊖ Very low
88% (77 to 96)	84% (62 to 96)	Uterosacral ligament[7] (T-1/T2-w + fat-suppressed/Gd, 2D FSE T2-w and 3.0T MRI)	241 (5)	⊕⊖⊖⊖ Very low
88% (64 to 99)	33% (4 to 78)	Uterosacral ligament[7] (3D MRI)	23 (1)	⊕⊖⊖⊖ Very low
75% (50 to 92)	94% (83 to 99)	Vaginal wall involvement[8] (T-1/T2-w + fat-suppressed/Gd, 2D FSE T2-w and 3.0T MRI)	248 (4)	⊕⊖⊖⊖ Very low
80% (28 to 99)	100% (81 to 100)	Vaginal wall involvement8 (3D MRI)	23 (1)	⊕⊖⊖⊖ Very low
89% (75 to 97)	91% (76 to 98)	Pouch of Douglas[9] (Jelly method (T1-/T2-w + fat-	154 (5)	⊕⊖⊖⊖ Very low

ISBN 978-1-4731-2661-9

Sensitivity (95%CI)	Specificity (95% CI)	Site of endometriosis (MRI test)	No of participants (studies)	Quality of the evidence (GRADE)[13]
		suppressed), 2D FSE T2-w and 3.0T MRI)		
71% (42 to 92)	100% (66% to 100%)	Pouch of Douglas[9] (3D MRI)	23 (1)	⊕⊖⊖⊖ Very low
50% (16 to 84)	100% (96 to 100)	Ureteral[10] (T1-/T2-w + fat-suppressed/Gd MRI)	92 (1)	⊕⊖⊖⊖ Very low
23% (5 to 54)	100% (95 to 100)	Bladder[11] (T1-/T2-w + fat-suppressed/Gd MRI)	92 (1)	⊕⊖⊖⊖ Very low
93% (78 to 99)	92% (73 to 99)	Ovarian[12] (T1-/T2-w + fat-suppressed/Gd and 3.0T MRI)	179 (3)	⊕⊖⊖⊖ Very low

CI: confidence interval
Endometriosis sites as defined in Nisenblat Cochrane Systematic Review 2016:
1 Endometriotic lesions, deep or superficial, located at any site in pelvic/abdominal cavity: on the peritoneum, fallopian tubes, ovaries, uterus, bowel, bladder or Pouch of Douglas
2 Deep endometriotic lesions extending more than 5 mm under the peritoneum located at any site of pelvic/abdominal cavity
3 Deep endometriotic lesions involve ≥ 1 site of the posterior pelvic compartment (uterosacral ligament, rectovaginal septum, vaginal wall, bowel) and/or obliterate Pouch of Douglas
4 Deep endometriotic lesions located at any site of the anterior pelvic compartment (bladder ± anterior pouch)
5 Deep endometriotic implants infiltrate the retroperitoneal area between posterior wall of vaginal mucosa and anterior wall of rectal muscularis
6 Endometriotic lesions infiltrating at least the muscular layer of the rectosigmoid colon; the most common form of bowel endometriosis
7 Endometriotic lesions infiltrate uterosacral ligaments unilaterally or bilaterally
8 Endometriotic lesions infiltrate vaginal wall, particularly posterior vaginal fornix
9 Defined when the peritoneum of the Pouch of Douglas is only partially or no longer visible during surgery and occurs as a result of adhesion formation; can be partial or complete, respectively
10 Endometriotic lesions involving ureters
11 Endometriotic lesions infiltrating bladder muscularis propria
12 Ovarian cysts lined by endometrial tissue (endometrioma)
13 Reasons for downgrading the evidence can be found in Appendix J.10
14 The specificity in Bazot 2013 study may be due to a different (pre-selected) population: a substantial proportion of women had endometriosis already

9.4.5 Economic evidence

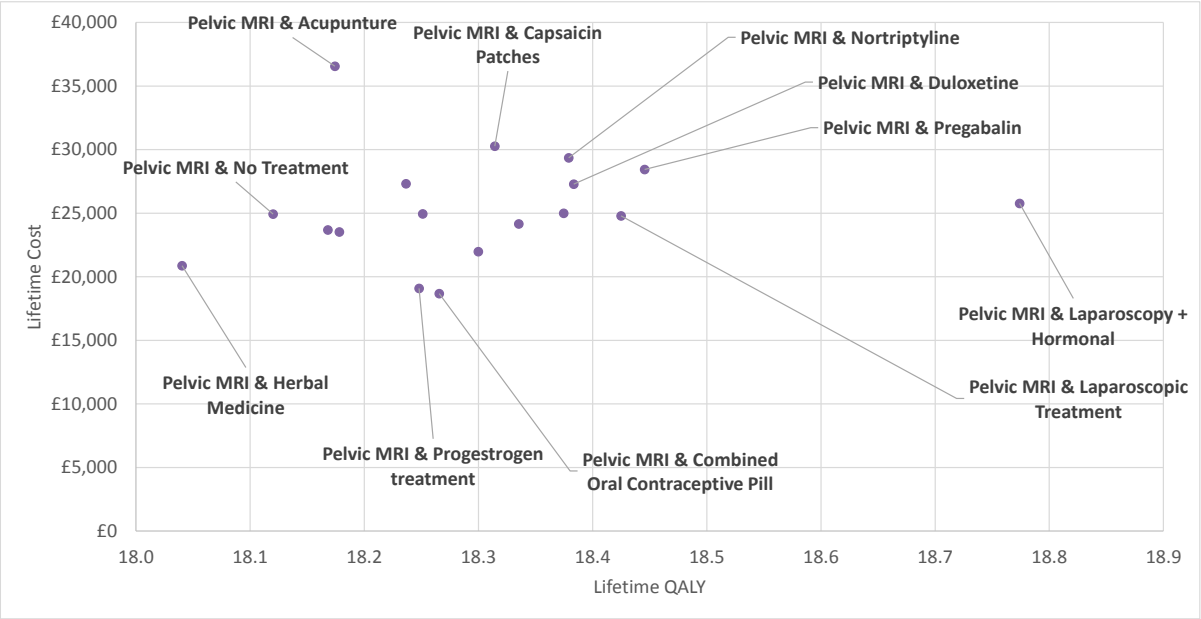
A significant source of dissatisfaction with the current treatment pathway for endometriosis relates to the slow diagnosis and treatment of the condition. Consequently a de novo economic model was constructed to consider the optimal diagnosis and treatment strategies to attempt to increase the speed of accurate diagnosis in a cost-effective way. However, as the choice of diagnostic test depends in part on the choice of treatment (which is itself influenced by the availability of other diagnostic tests) it does not make sense to consider the 'cost-effectiveness' of one particular diagnostic strategy as though this were independent from the cost-effectiveness of other such strategies.

Figure 11 demonstrates how Pelvic MRI interacts with various treatment options and Table 47 tabulates the same data. The findings confirm the intuitive belief that offering an expensive test like MRI should only be done if the treatment is expensive (or risky) enough to make it worthwhile paying for the extra specificity and sensitivity of an MRI. Consequently the incremental benefit of MRI is highest for the most expensive treatment, laparoscopic surgery and adjunct hormonal treatment. Nevertheless this enormous cost-effectiveness almost disappears when other diagnostic strategies are considered in tandem; in the full model MRI

ISBN 978-1-4731-2661-9

is only borderline cost-effective because more cost-effective options exist for both major treatment groups recommended by the health economic model.

Figure 11: Costs and Lifetime QALYs of offering various treatment options in combination with MRI



Source: Economic model

Table 47: Costs and Lifetime QALYs of offering various treatment options in combination with MRI (showing only non-dominated strategies)

Treatment	Cost	QALY	ICER	Probability cost-effective vs no treatment (£20,000 / QALY)	Probability cost-effective vs no treatment (£30,000 / QALY)
Empirical Diagnosis & No Treatment	£22,752.60	18.120	Base Case	N/A	N/A
Pelvic MRI & Combined Oral Contraceptive Pill	£18,674.17	18.266	-£28,032.93	82.4%	90.1%
Pelvic MRI & Danazol	£21,968.90	18.300	Extendedly Dominated	81.3%	91.2%
Pelvic MRI & Amitriptyline	£24,157.06	18.335	Extendedly Dominated	85.7%	94.5%
Pelvic MRI & Laparoscopic Treatment	£24,783.78	18.425	Extendedly Dominated	90.1%	91.2%
Pelvic MRI & Laparoscopy + Hormonal	£25,772.03	18.774	£3,681.35	86.8%	89.0%

ISBN 978-1-4731-2661-9

9.4.6 Clinical evidence statements

9.4.6.1 Pelvic endometriosis

Eight studies (n=333, includes conventional (T1-/T2-w), T1-w+fat-suppressed, T-1/T2-w + fat-suppressed/Gd and 3.0T MRI) reported that the pooled sensitivity and specificity of MRI was 77% (62% to 88%) and 72% (53% to 87%). Two studies (n=62, includes T1-/T2-w + fat-suppressed and fat-suppressed MRI) showed sensitivity and specificity of 86% (64% to 97%) and 76% (56% to 90%), 50% (19% to 81%) and 100% (16% to 100%), respectively. One study (n=31, includes T-1/T2-w + fat-suppressed/Gd MRI) found sensitivity of 81% (58% to 95%) and specificity of 50% (19% to 81%). Evidence was of very low quality.

9.4.6.2 DIE

Very low quality evidence from 4 studies (n=212, includes T-1/T2-w + fat-suppressed/Gd and 3.0T MRI) found that the pooled sensitivity and specificity of MRI was 96% (90% to 99%) and 86% (54% to 98%).

9.4.6.3 Posterior DIE

Very low quality evidence from 2 studies (n=54, includes Jelly method (T1-/T2-w + fat-suppressed) and 2D FSE T2-w MRI) reported that the sensitivity and specificity of MRI was 89% (65% to 99%), 94% (71% to 100%), and 20% (1% to 72%) and 100% (77% to 100%), respectively. Very low quality from 1 study (n=23, includes 3D MRI) found sensitivity of 100% (81% to 100%) and specificity of 20% (1% to 72%).

9.4.6.4 Anterior DIE

Very low quality evidence from 1 study (n=41, includes 3.0T MRI) reported the sensitivity of MRI in diagnosing anterior DIE of 75% (35% to 97%) and specificity of 100% (89% to 100%).

9.4.6.5 Rectovaginal endometriosis

Very low quality evidence from 3 studies (n=288, includes T-1/T2-w + fat-suppressed/Gd MRI) found that the pooled sensitivity and specificity of MRI was 75% (35% to 95%) and 88% (43% to 99%).

9.4.6.6 Rectosigmoid endometriosis

Very low quality evidence from 6 studies (n=662, includes T-1/T2-w + fat-suppressed/Gd, 2D FSE T2-w, jelly method (T1-/T2-w + fat-suppressed) and 3.0T MRI) found that the pooled sensitivity and specificity of MRI was 91% (79% to 97%) and 96% (92% to 99%). Very low quality evidence from 1 study (n=23, includes 3D MRI) reported sensitivity of 85% (55% to 98%) and specificity of 90% (55% to 100%).

9.4.6.7 Uterosacral ligament endometriosis

Very low quality evidence from 5 studies (n=241, includes T-1/T2-w + fat-suppressed/Gd, 2D FSE T2-w and 3.0T MRI) found that the pooled sensitivity of MRI was 88% (77% to 96%) and the pooled specificity was 84% (62% to 96%). Very low quality evidence from 1 study (n=23, includes 3D MRI) reported sensitivity and specificity of 88% (64% to 99%) and of 33% (4% to 78%).

9.4.6.8 Vaginal wall involvement

Very low quality evidence from 4 studies (n=248, includes T-1/T2-w + fat-suppressed/Gd, 2D FSE T2-w and 3.0T MRI) found that the pooled sensitivity and specificity of MRI was 75%

ISBN 978-1-4731-2661-9

(50% to 92%) and 94% (83% to 99%). Very low quality evidence from 1 study (n=23, includes 3D MRI) reported sensitivity of 80% (28% to 99%) and specificity of 100% (81% to 100%).

9.4.6.9 Pouch of Douglas

Very low quality evidence from 5 studies (n=154, includes jelly method (T1-/T2-w + fat-suppressed), 2D FSE T2-w and 3.0T MRI) found that the pooled sensitivity and specificity of MRI was 89% (75% to 97%) and 91% (76% to 98%). Very low quality evidence from 1 study (n=23, includes 3D MRI) reported sensitivity of 71% (42% to 92%) and specificity of 100% (66% to 100%).

9.4.6.10 Ureteral endometriosis

Very low quality evidence from 1 study (n=92, includes T1-/T2-w + fat-suppressed/Gd MRI) reported sensitivity of 50% (16% to 84%) and specificity of 100% (96% to 100%).

9.4.6.11 Bladder endometriosis

Very low quality evidence from 1 study (n=92, includes T1-/T2-w + fat-suppressed/Gd MRI) found sensitivity of 23% (5% to 54%) and specificity of 100% (95% to 100%).

9.4.6.12 Ovarian endometriosis

Very low quality evidence from 3 studies (n=179, includes T1-/T2-w + fat-suppressed/Gd and 3.0T MRI) found that the pooled sensitivity and specificity of MRI was 93% (78% to 99%) and 92% (73% to 99%).

9.4.7 Evidence to recommendations

9.4.7.1 Relative value placed on the outcomes considered

As sensitivity and specificity reflect patient outcomes, these were prioritised as critical outcomes for this review. Although the Committee did not specify clinically important thresholds for these 2 diagnostic measures, the imprecision of estimates were assessed according to the confidence region around the pooled estimate in the ROC plots. Inconclusive results and test complications were also considered by the Committee. Quality of life was prioritised as an outcome if this was available from test and treat RCTs.

9.4.7.2 Consideration of clinical benefits and harms

The Committee agreed that their priority was early detection and treatment of endometriosis to prevent disease progression and to enable early clinical management.

The Committee discussed the importance of reducing the likelihood of a false negative diagnosis which could result in the woman not receiving effective management and the potential additional negative psychological impact of a false negative diagnosis if a woman was experiencing painful symptoms. However they noted that a false positive result might lead to unnecessary treatment and also result in a negative psychological impact.

The Committee also discussed the relative benefits and harms associated with MRI scanning. They concluded that laparoscopy although invasive is necessary as the gold standard test for identification of endometriosis. The benefit of MRI would be as an additional non-invasive informative test for surgery because it would identify the involvement and depth of endometriosis prior to surgery.

ISBN 978-1-4731-2661-9

The Committee considered that the value of an MRI was dependent on the proper interpretation and reporting of the results and that this should be performed by a healthcare professional appropriately trained in interpretation of MRI scans for endometriosis.

9.4.7.3 Consideration of economic benefits and harms

The model identifies Pelvic MRI as being a useful intermediate step between empirical diagnosis (treating based on symptoms rather than definitive diagnosis) and laparoscopic confirmation, which tends to be quite expensive. This makes the model important for identifying whether the switchover from empirical diagnosis (which would be preferred at low cost/QALY thresholds) to laparoscopic confirmation (which would be preferred at higher cost/QALY thresholds) allows Pelvic MRI to be the most cost-effective treatment at some intermediate thresholds. In the main health economic model the strategy of Pelvic MRI & Laparoscopy and adjunct hormonal therapy is borderline cost-effective and in the case of infertile women Pelvic MRI & Laparoscopic treatment is comfortably cost-effective at the usual threshold of £20,000 / QALY.

The Committee disagreed with the findings of the model, stressing that in their opinion the NHS Reference Cost overpriced a transabdominal ultrasound and underpriced a Pelvic MRI. As these values were used in the model this translated to an effective 'switching' of MRI and ultrasound in the order of cost-effectiveness. It is possible for both claims about the cost of MRI to be accurate at once; if clinicians operate in an environment where time on an MRI machine is scarce then this may genuinely shift the shadow price of an MRI scan for these clinicians while not altering the expected marginal cost of performing a scan (i.e. the cost of the machine divided by the number of scans it can be expected to do in its lifetime). Consequently the Committee agreed to leave the model unchanged to allow for a rationalisation of the price of MRI in the future, but make recommendations based on their clinical expertise of the price of an MRI.

The health economic importance of avoiding false negatives was discussed by the Committee. There was considerable discussion around the evidence which suggested MRI was more sensitive than ultrasound. Eventually it was concluded that the cost of a change in practice to all-MRI as a first-line treatment was not supported by the health economic evidence, especially in the light of disagreement about the true cost of an MRI.

The Committee and model were in agreement that an MRI was the most cost-effective method of assessing the extent of deep endometriosis.

As the Committee are recommending fewer MRIs than current practice, these recommendations are likely to have a small negative resource impact as some marginal MRIs are converted into ultrasound scans.

A fuller discussion of the economic benefits and harms of diagnostic strategies is located in the Health Economic Appendix K.

9.4.7.4 Quality of evidence

The quality of the evidence was very low according to GRADE criteria. This was mainly due to risk of bias (often the patient selection was not consecutive or random, or not all patients were included in the analysis), inconsistency (particularly related to specificity estimates) as well as imprecision with a high level of uncertainty as indicated by the confidence region in the pooled analysis.

The Committee commented that studies conducted in the 1990s or earlier would use MRI scanning techniques that may not be used in current practice due to advancement of technology. However, as this cut off was not specified in the protocol, older studies were included in the review.

ISBN 978-1-4731-2661-9

The Committee acknowledged that specificity was particularly variable across studies suggesting that even if a woman did have a negative MRI test result, this would not indicate very much to a clinician. The high level of imprecision expressing uncertainty around the pooled effect estimates (indicated by wide confidence regions) was also discussed.

As the evidence was limited to the detection of deep infiltrating endometriosis, the Committee considered that their recommendations should not extend to earlier or more superficial disease. The Committee concluded that the evidence showed that MRI was a good test for deep endometriosis involving the bowel, bladder or ureter, but should not be used as the first diagnostic or investigative test in women with suspected endometriosis. The Committee were unable to specify where an MRI would be in the patient pathway, because its use would also depend on judgements regarding the symptoms and signs and the clinical examination of the woman at presentation (for example severe chronic pelvic pain, severe dysmenorrhea, endometriomas or deep nodules).

9.4.7.5 Other considerations

The Committee believed that MRI should not be restricted to specialist endometriosis services, however it was also noted that if there was no specialist endometriosis service, hospitals would struggle to provide the service. The Committee considered whether any additional recommendations were necessary for adolescents who were identified in equalities impact assessment but concluded that none were necessary.

The Committee discussed whether further evidence would reduce the uncertainty around the results. However, as there was a sufficient body of evidence to make clinical recommendations for the use of MRI investigation in women with deep endometriosis infiltrating the bowel, bladder or ureter, it was decided that research in populations of women with less severe disease would not be useful. The Committee therefore did not prioritise this topic for further research.

9.4.7.6 Key conclusions

The Committee concluded that due to the large number of false negative results a recommendation to use MRI testing may potentially lead to many women being falsely reassured that they do not have endometriosis. MRI was therefore discounted as a first line test and recommendations regarding its use were limited to the diagnosis of deep endometriosis infiltrating the bowel, bladder or ureter in women with more advanced stages of the disease, who may require further surgery.

9.4.8 Recommendations

22. **Do not use pelvic MRI as the primary investigation to diagnose endometriosis in women with symptoms or signs suggestive of endometriosis.**

23. **Consider pelvic MRI to assess the extent of deep endometriosis involving the bowel, bladder or ureter.**

24. **Ensure that pelvic MRI scans are interpreted by a healthcare professional with specialist expertise in gynaecological imaging.**

ISBN 978-1-4731-2661-9

9.5 Surgical diagnosis with or without histological confirmation

Review question: What is the accuracy of surgery with or without histological confirmation in diagnosing endometriosis?

9.5.1 Introduction

Laparoscopy is the "gold standard" for making a diagnosis, although there is clinical disagreement about the need for a histological specimen to confirm the visual diagnosis. This is both in relation to confirming the diagnosis and to exclude any other pathology, such as malignancy. The place of the latter in relation to ovarian and extra-ovarian malignancy will be considered.

The aim of this review was to evaluate whether during a diagnostic laparoscopy a sample for histological analysis should be taken.

For full details see review protocol in Appendix D.

9.5.2 Description of clinical evidence

Seventeen studies were included in this review (Balasch 1996; Buchweitz 2006; Chatman 1987; Cornilie 1990; De Almeida Filho 2008; El Bishry 2008; Emmert 1998; Fernando 2013; Jansen 1986; Keltz 1995; Mettler 2003; Nisolle 1990; Shafik 2000; Stratton 2002; Stripling 1988; Vercellini 1991; Walter 2001). The single studies included in the review by Wykes 2004 were considered and those, that met the inclusion criteria (13 studies) according to the review protocol, were included in the current review (Table 48).

The size of the population in the studies ranged from 14 (Emmert 1998) to 976 (De Almeida Filho 2008).

Three studies (Vercellini 1991; Walter 2001; De Almeida Filho 2008) reported sensitivity and specificity, whereas the remaining 14 studies reported positive test results only (i.e. biopsy histology results from only those who were laparoscopically diagnosed with endometriosis) (Balasch 1996; Buchweitz 2003; Chatman 1987; Cornilie 1990; El Bishry 2008; Emmert 1998; Jansen 1986; Keltz 1995; Nisolle 1990; Shafik 2000; Stratton 2002; Stripling 1988; Mettler 2003; Fernando 2013).

There were 2 studies which excised cells from 'normal-looking' area from women with laparoscopically diagnosed endometriosis, in addition to endometriotic lesions (Balasch 1996, Nisolle 1990). The results for the number of 'normal-looking' areas which were diagnosed histologically as endometriosis have not been reported here as it was felt that they did not constitute a 'true' negative test (negative laparoscopy results and negative histology results).

In contrast to the other included diagnostic tests, for this specific question a separate protocol was drafted because studies with incomplete verification of the index test (surgical diagnosis) will be included if they have reported any of the diagnostic outcomes. This was due to the nature of the surgical diagnostic procedure.

Evidence from the included studies are summarised in the clinical evidence profile in Table 49 and Table 50. See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, and study evidence tables in Appendix G.

9.5.3 Summary of included studies

A summary of the studies that were included in this review are presented in Table 48.

ISBN 978-1-4731-2661-9

Table 48: Summary of included studies

Study	Index test/reference standard	Population	Outcomes*	Comments
Balasch 1996 Spain	• Laparoscopy • Histology Biopsies were placed in formalin and processed in the routine fashion for light microscopy	Women undergoing laparoscopy for infertility N=100 (consecutive)	Endometriosis (number of patients) Positive test	Biopsies of 'normal' uterosacral ligaments were obtained from all women with laparoscopically diagnosed endometriosis. Biopsies of suspected endometriosis were taken from 19 of 47 women with laparoscopically diagnosed endometriosis.
Buchweitz 2003 Germany	• Laparoscopy • Histology Not reported how the specimens were handled	Women with pain or infertility N=118 (consecutive)	Endometriosis (number of biopsies and patients) Positive test	Only AFS 1 and 2 included
Chatman 1987 USA	• Laparoscopy • Histology Pathology specimens consisting of 5-to 10-mm tissue samples were processed and stained with haematoxylin and eosin. Histologic confirmation of endometriosis was established with light microscopy only in the presence of endometrial glands with or without stroma	Women with the primary complaint of pelvic pain N=115 (consecutive)	Endometriosis (number of patients) Positive test	158 women were not biopsied because it was thought that a biopsy would be superfluous or because the endometriotic implants were found in areas deemed unsafe for biopsy
Cornillie 1990 Belgium	• Laparoscopy • Histology Biopsies were fixed in phosphate-buffered formalin, dehydrated	Women undergoing laparoscopy for infertility, pain, or both.	Endometriosis (number of patients) Positive test	Biopsies were only taken from women with laparoscopically diagnosed endometriosis

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Outcomes*	Comments
	through alcohols and embedded in paraffin.	N=179 (consecutive)		(n=142) and with endometriosis with depth greater than 3mm (n=110)
De Almeida Filho 2008 Brazil	• Laparoscopy • Histology Tissue preparations were stained with haematoxylin-eosin or, in 15 cases, with periodic acid-Schiff stain and/or silver impregnation stain.	Women undergoing laparoscopy due to pelvic pain and/or infertility N=976	Endometriosis (number of patients) Sensitivity and specificity, positive and negative test	Out of 976 patients, who underwent laparoscopy, 48% were selected for inclusion in the present study, since the presented clinical and laparoscopic profiles were suggestive of endometriosis. In the cases of a further 8 patients, a positive histopathological diagnosis was made during surgical procedures that were performed due to other causes
El Bishry 2008 UK	• Laparoscopy • Histology All specimens were put in formalin pots and sent to the laboratory on the same day	Women who had undergone laparoscopies for investigation of pelvic pain and those found to have endometriosis N=63	Endometriosis (number of biopsies and patients) Positive test	Excisions of endometriotic lesions was undertaken in 48 patients; in 6.3% cases the histology was inconclusive
Emmert 1998 Germany	• Laparoscopy/pelviscopy • Histology Not reported how the specimens were handled	Adolescent girls undergoing laparoscopy/pelviscopy for chronic or acute pelvic pain and right-sided lower abdominal pain	Endometriosis (number of patients) Positive test	14 of 37 girls with laparoscopically diagnosed endometriosis had histology samples taken. It is not clear

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Outcomes*	Comments
		N=105		why the other girls did not have biopsies taken.
Fernando 2013 Australia	• Laparoscopy • Histology All excised lesions were processed and embedded in paraffin blocks, sectioned and stained with haematoxylin-eosin.	Women with suspected endometriosis because of pain or infertility N=431	Endometriosis (number of biopsies Positive test	In 40 patients surgery was performed by training registrars or fellows and these patients were excluded because the number of procedures performed by each physician were too small to lead to meaningful conclusions
Jansen 1986 Australia	• Laparoscopy • Histology Biopsy specimens were fixed in formalin, acetic acid and alcohol, embedded in paraffin, step sectioned, and stained with haematoxylin and eosin according to standard techniques	Women who underwent laparoscopy for infertility or other indications including pelvic pain and assessment for sterilization reversal N=77	Endometriosis (number of biopsies) Positive test	
Keltz 1995 USA	• Laparoscopy • Histology All specimens were fixed in paraffin, underwent haematoxylin and eosin staining	Women undergoing laparoscopy for chronic pelvic pain N= 51	Endometriosis (number of biopsies and patients) Positive test	
Mettler 2003 Germany	• Laparoscopy • Histology Not reported how the specimens were handled	Women who underwent laparoscopy for suspected endometriosis N=164	Endometriosis (number of biopsies and patients) Positive test	

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Outcomes*	Comments
Nisolle 1990 Belgium	• Laparoscopy • Histology All biopsy specimens were fixed in formaldehyde and embedded in paraffin. 3 micrometer serial sections were stained with Gomori's Trichrome and examined, on a blind basis, with a Leitz Orthoplan microscope	Women undergoing laparoscopy for infertility N=118	Endometriosis (number of biopsies and patients) Positive test	Samples were taken from women with laparoscopically diagnosed endometriosis from both suspected 'endometriotic tissue' and some 'normal' looking peritoneum. The results for histologically 'normal' tissue samples were not presented.
Shafik 2000 UK	• Laparoscopy • Histology Biopsies were fixed in neutral, buffered 4% formal saline and examined with the light microscope after staining with haematoxylin and eosin	Women with chronic pelvic pain N=59	Endometriosis (number of biopsies and patients) Positive test	Biopsies from 3 women were unsuitable for histological evaluation and were excluded
Stratton 2002 USA	• Laparoscopy • Histology Specimen were fixed in formalin and embedded in paraffin, and stained with haematoxylin and eosin	Women with chronic pelvic pain though to be due to endometriosis N=77	Endometriosis (number of biopsies and patients) Positive test	
Stripling 1988 USA	• Laparotomy/ +/- laparoscopy • Histology Standard haematoxylin and eosin stains were performed on all specimens	Postoperative diagnosis of endometriosis N=109	Endometriosis (number of biopsies and patients) Positive test	
Vercellini 1991 Italy	• Laparotomy • Histology At least 10 serial sections were made	Women who underwent laparotomy for ovarian cysts	Endometrioma (number of biopsies	

ISBN 978-1-4731-2661-9

Study	Index test/reference standard	Population	Outcomes*	Comments
	for each specimen, stained with haematoxylin and eosin and examined at the light microscope at 10x and 40x magnifications	N=245	and patients) Sensitivity and specificity, positive and negative test	
Walter 2001 USA	• Laparoscopy • Histology The specimens were fixed in formalin and embedded in paraffin, and 3-to4-µm sections were obtained every 50 to 60 µm. The sections were stained with haematoxylin and eosin. Four to 6 sections per specimen were evaluated by means of light microscopy	Women with chronic pelvic pain or known endometriosis (diagnosed histologically or by visualization) N=44 (consecutive)	Endometriosis (number of biopsies) Sensitivity and specificity, positive and negative test	

*N: number of participants in study; AFS: American Fertility Society classification; * positive test: number (%) of positive histologic findings of endometriosis/endometrioma among the positive visual findings; negative test: number (%) of normal histologic findings among the negative visual findings*

9.5.4 Clinical evidence profile

The clinical evidence profile for this review question is presented in Table 49 and Table 50. Results from the Diagnostic Accuracy Studies 2 (QUADAS 2) checklist ranged from very high to moderate risk.

The evidence could not be pooled, due to the differences in study design and how results were reported. Therefore the results are reported by study. Please see Table 49 and Table 50 below for endometriosis and endometrioma, respectively.

Table 49: Summary clinical evidence profile for diagnosis of endometriosis

Study	Sensitivity (95%CI)	**Specificity (95% CI)**	**Endometriosis -number of biopsies*** **+test** **-test**	**Endometriosis - number of patients*** **+test** **-test**	**Risk of bias**
Balasch 1996	-	-	-	17/19 (89%) -	Very high[1]

ISBN 978-1-4731-2661-9

Study	Sensitivity (95%CI)	**Specificity (95% CI)**	**Endometriosis -number of biopsies*** **+test** **-test**	**Endometriosis - number of patients*** **+test** **-test**	**Risk of bias**
Buchweitz 2003	-	-	77/137 (56%) -	49/69 (71%) -	High[2]
Chatman 1987	-	-	-	74/115 (64%) -	Very high[3]
Cornillie 1990	-	-	-	84/110 (76%) -	Very high[4]
De Almeida Filho 2008	98% (95% to 99%)	79% (76% to 82%)	-	337/468 (72%) 500/508 (98%)	High[2]
El Bishry 2008	-	-	104/132 (79%) -	36/48 (75%) -	Very high[5]
Emmert 1998	-	-	-	6/14 (43%) -	Very high[6]
Fernando 2013	-	-	1082/1439 (75%) -	-	Very high[5]
Jansen 1986	-	-	73/137 (53%) -	-	Very high[5]
Keltz 1995	-	-	21/37 (57%) -	21/37 (57%) -	Very high[7]
Mettler 2003	-	-	142/264 (54%) -	138/164 (84%) -	Very high[5]
Nisolle 1990	-	-	80/86 (93%) -	80/86 (93%) -	Very high[8]
Shafik 2000	-	-	85/150 (57%) -	43/59 (73%) -	Very high[9]
Stratton 2002	-	-	189/314 (60%) -	57/65 (88%) -	Very high[5]

ISBN 978-1-4731-2661-9

Study	Sensitivity (95%CI)	**Specificity (95% CI)**	**Endometriosis -number of biopsies*** **+test** **-test**	**Endometriosis - number of patients*** **+test** **-test**	**Risk of bias**
Stripling 1988	-	-	148/164 (90%) -	106/109 (97%) -	High[2]
Walter 2001	97% (90% to 100%)	77% (72% to 82%)	67/138 (49%) 240/242 (99%)	-	Moderate[10]

*CI: confidence interval; * +test: number (%) of positive histologic findings of endometriosis among the positive visual findings; -test: number (%) of normal histologic findings among the negative visual findings*
1 not all laparoscopically diagnosed patients had a biopsy taken. It is unclear how the patients were selected for biopsy and whether this could have influenced the results; unclear whether the study was blinded.
2 unclear whether the study was blinded
3 not all patients were include in the analysis; unclear whether the study was blinded
4 only lesions with depth greater than 3mm were excised; unclear whether lesions of lower depth would have the same results. Unlikely that the study was blinded.
5 unclear whether a consecutive or random sample was used and whether the study was blinded
6 not all laparoscopically diagnosed patients had a biopsy taken. It is unclear how the patients were selected for biopsy and whether this could have influenced the results; unclear whether the study was blinded.
7 lack of information about methods included in the study; unclear whether the study was blinded.
8 unclear whether a consecutive or random sample was used.
9 not all patients were included in the analysis; unclear whether a consecutive or random sample was used and whether the study was blinded
10 unclear whether the index test results were interpreted without knowledge of the results of the reference standard

Table 50: Summary clinical evidence profile for diagnosis of endometrioma

Study	**Sensitivity (95%CI)**	**Specificity (95% CI)**	**Endometrioma** **(number of ovarian cysts)*** **+test** **-test**	**Risk of bias**
Vercellini 1991	97% (94 to 99)	95% (90 to 99)	213/218 (98%) 106/113 (94%)	Very high[1]

*CI: confidence interval: * +test: number (%) of positive histologic findings of endometrioma among the positive visual findings; -test: number (%) of normal histologic findings among the negative visual findings*
1 unclear whether a consecutive or random sample was used and whether the study was blinded

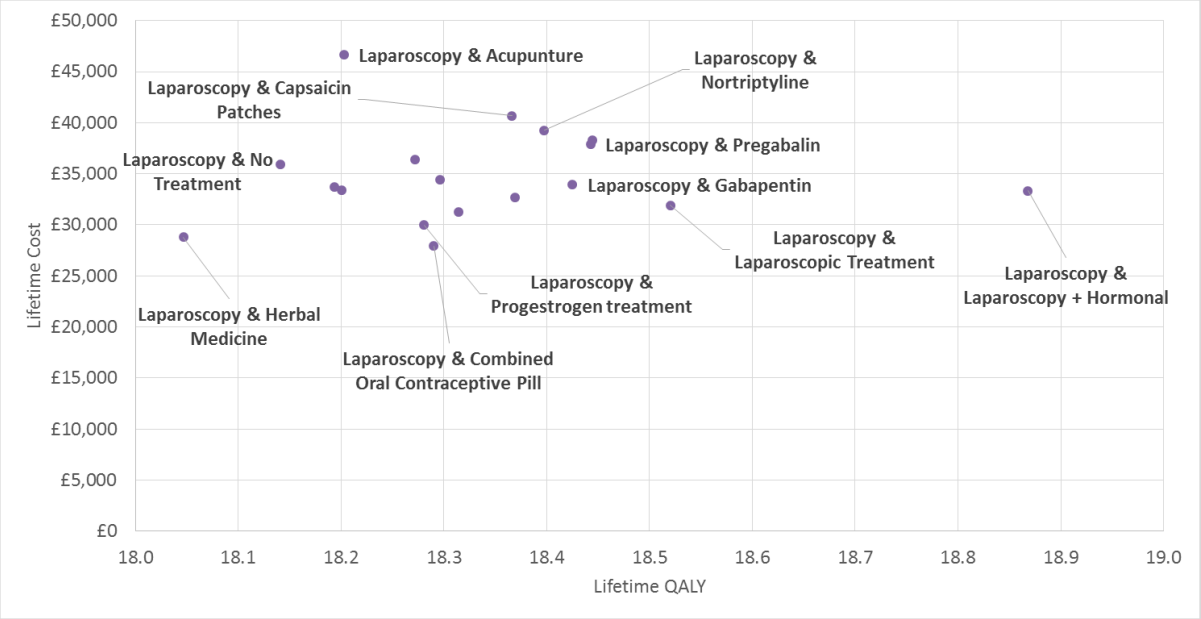
9.5.5 Economic evidence

A significant source of dissatisfaction with the current treatment pathway for endometriosis relates to the slow diagnosis and treatment of the condition. Consequently a de novo economic model was constructed to consider the optimal diagnosis and treatment strategies to attempt to increase the speed of accurate diagnosis in a cost-effective way. However, as the choice of diagnostic test depends in part on the choice of treatment (which is itself influenced by the availability of other diagnostic tests) it does not make sense to consider the 'cost-effectiveness' of one particular diagnostic strategy as though this were independent from the cost-effectiveness of other such strategies.

ISBN 978-1-4731-2661-9

No health economic evidence was found on the cost-effectiveness of surgical diagnosis. As a modelling assumption, supported by the Committee, surgical diagnosis was assumed to be the reference standard. This assumption was relaxed in sensitivity analysis. Figure 12 demonstrates how surgery interacts with various treatment options and Table 51 tabulates the same data. The findings show that in general the most cost-effective way to use the expensive surgical diagnosis is to offer it with more expensive and effective treatments. Consequently the incremental benefit of surgery is highest for the most expensive treatment, laparoscopic surgery and adjunct hormonal treatment. This has an extremely low ICER of £3,700. The ICER increases substantially when other diagnostic strategies are considered, most notably empirical diagnosis in combination with cheap treatments such as hormonal contraceptives.

Figure 12: Costs and Lifetime QALYs of offering various treatment options in combination with surgery



Source: Economic model

Table 51: Costs and Lifetime QALYs of offering various treatment options in combination with surgery (showing only non-dominated strategies)

Treatment	Cost	QALY	ICER	Probability cost-effective vs no treatment (£20,000 / QALY)	Probability cost-effective vs no treatment (£30,000 / QALY)
Empirical Diagnosis & No Treatment	£22,752.60	18.120	Base Case	N/A	N/A
Laparoscopy & Combined Oral Contraceptive Pill	£27,924.59	18.290	Extendedly Dominated	96.7%	96.7%
Laparoscopy & Danazol	£31,292.18	18.315	Extendedly Dominated	90.1%	93.4%
Laparoscopy &	£31,899.07	18.520	Extendedly Dominated	92.3%	94.5%

ISBN 978-1-4731-2661-9

Treatment	Cost	QALY	ICER	Probability cost-effective vs no treatment (£20,000 / QALY)	Probability cost-effective vs no treatment (£30,000 / QALY)
Laparoscopic Treatment					
Laparoscopy & Laparoscopy + Hormonal	£33,344.74	18.868	£3,709.17	97.8%	100.0%

9.5.6 Clinical evidence statements

9.5.6.1 Endometriosis

Two moderate and high risk of bias studies reported similar findings regarding sensitivity and specificity: 97% (90% to 100%) and 98% (95% to 99%), and 77% (95%CI: 72% to 82%) and 79% (95%CI: 76% to 82%), respectively.

Biopsies

In studies with very high to high risk of bias, where no sensitivity and specificity were reported, the papers only reported positive test results, i.e. where results of histology matched the positive surgical diagnosis. The results were highly variable. The positive test result ranged from 53% to 93% (based on the number of biopsies). The median of visual diagnosis confirmed histologically was 58.5% based on biopsies (n=11 studies).

Number of patients

In studies, where positive test values were presented based on the number of patients, the positive test range was between 42% and 97%. The median of visual diagnosis confirmed histologically was 75.5% based on the number of patients (n=13 studies).

9.5.6.2 Endometrioma

A very high risk of bias study reported a sensitivity of 97% (94% to 99%) and a specificity of 95% (90% to 99%) (based on the number of ovarian cysts). The positive and negative test results, i.e. where results of histology matched the positive or negative surgical diagnosis, were 98% and 94%, respectively.

9.5.7 Evidence to recommendations

9.5.7.1 Relative value placed on the outcomes considered

As sensitivity and specificity as a proxy for patient level outcomes, these were prioritised as critical outcomes for this review. No test and treat randomised controlled trials which would directly report patient level outcomes (such as health related quality of life) were identified. Inconclusive results and test complications were also considered by the Committee.

9.5.7.2 Consideration of clinical benefits and harms

The diagnosis of endometriosis is made on the basis of visualisation during laparoscopy. Biopsies can also be taken to confirm the visual diagnosis by histology. The Committee

ISBN 978-1-4731-2661-9

discussed whether it is practical to perform histology to diagnose endometriosis and concluded that histology may be important in order to diagnose other conditions and/or malignancies.

In terms of endometrioma, the Committee considered histology would be performed when undergoing treatment by fenestration or fenestration plus ablation of capsule to ensure histological evidence was available during a therapeutic laparoscopy. The Committee also agreed that surgical treatment of endometrioma should include histology to rule out an alternative diagnosis of ovarian lesions and to exclude malignancies, and that it is a good practice, when undertaking laparoscopic excision, to send excised tissue for histology.

The Committee recognised that diagnosis would be dependent on the individual histopathologist in terms of how detailed their examination of the sample was to identify endometriosis – the greater the scrutiny the more likely it would be found. The Committee were aware that subtle differences might be hard to detect although acknowledged that identification of stromal endometriosis is becoming more standard.

Although there is no specific guideline for histopathology, there is literature that shows that further histology results in better identification of endometriosis. Therefore, they suggested that in order to perform a thorough histology, additional tissue samples may be required and that the histological examiner should be trained to look for endometriotic tissue.

9.5.7.3 Consideration of economic benefit and harms

The health economic model considers laparoscopy to be the ‘gold standard’ of diagnosing endometriosis. Consequently even though the cost of obtaining a sample of the endometrium for biopsy is relatively cheap, the use of such a confirmatory test could never be cost-effective as it costs more money than a test which is ‘perfect’. This is at odds with clinical reality, where it is obvious that histopathology would not be undertaken if it did not add value. Because the evidence underpinning the economic model makes the assumption that laparoscopy is the ‘gold standard’, the economic model must do this too. However the assumption that laparoscopy is perfect is varied in sensitivity analysis.

Given the assumption described above, laparoscopy is a highly effective form of diagnosis, especially given the Committee believed most of the time a diagnostic laparoscopy is conducted it would have some clinical benefit. The model finds surgical diagnosis to be cost-effective even at quite low cost per QALY thresholds when considered in isolation. However when taken as one of many possible diagnosis/treatment strategies, surgical diagnosis is not preferred to empirical diagnosis and cheap treatment.

The economic model considers histology alone as a possible diagnostic strategy. Further details are given in the Health Economic Appendix K.

The Committee discussed how histology could be used to diagnose or exclude other conditions. While this would be outside the scope of the guideline, the cheap cost of histopathology taken at the time of unrelated surgery and the possibility for reducing diagnostic delays indicate that this suggestion is likely to be both cost-saving and improve quality of life by achieving accurate diagnosis.

9.5.7.4 Quality of evidence

The risk of bias was very high to moderate according to QUADAS 2 criteria. Main reasons leading to downgrading of evidence shared by the majority of studies were no information on blinding and it was unclear whether patients were selected consecutively or randomly.

The Committee noted that, in terms of the histologic diagnosis of endometriosis, the harder that it is looked for, the more likely it will be found. The Committee noted that it is highly likely that in some papers included in the review, where the visual surgical diagnosis of

ISBN 978-1-4731-2661-9

endometriosis was often not confirmed by histology, the researchers did not look hard enough to find the condition. They also believed that if a woman had a visual diagnosis of endometriosis, it would not be always be confirmed by histology. On this basis, it was agreed that having a histology report is very useful for the patient as it may offer her more reassurance.

9.5.7.5 Other considerations

The Committee were aware that laparoscopies are sometimes performed with inadequate examination of the pelvis resulting in false negative results, for example, where the bowel is only visualised without being moved, and the Committee agreed that there should be a systematic examination of the pelvis.

It was recommended that this systematic inspection should be carried out by a gynaecologist with training and skills in laparoscopic surgery because it is possible to miss significant endometriosis (e.g. inspections need to include draining fluid to enable visualisation of hidden deposits). The importance of recognised training standards, such as those set by the Royal College of Obstetricians and Gynaecologists, was highlighted in this context. Good documentation was also agreed to be very important (e.g. photo-documentation / video-documentation of findings during surgery). This could ensure that patients referred on for specialist care have adequate documentation of their disease if further surgery is required.

It was recognised that when women were suspected of having endometriosis involving the bowel, bladder or ureters that imaging prior to the procedure may be helpful to identify key areas for further inspection during laparoscopy. Diagnostic laparoscopy should also investigate for signs of non-pelvic endometriosis.

9.5.7.6 Key conclusions

The Committee concluded that laparoscopy should be considered in women with symptoms of endometriosis even when imaging has given a normal result. A negative finding following a thorough laparoscopic visualisation is highly specific and women can be reassured that they do not have endometriosis. Histological examination of biopsied tissue is considered to be a gold standard test and helpful to confirm the visual diagnosis; it is also required to exclude malignancy if ovarian endometriosis (endometrioma) is fenestrated and ablated.

9.5.8 Recommendations

Also refer to section 11.3.5 on surgical management, and section 12.3.4 on surgical management if fertility is a priority.

25. **Consider laparoscopy to diagnose endometriosis in women with suspected endometriosis, even if the ultrasound was normal.**

26. **For women with suspected deep endometriosis involving the bowel, bladder or ureter, consider a pelvic ultrasound or MRI before an operative laparoscopy.**

27. **During a diagnostic laparoscopy, a gynaecologist with training and skills in laparoscopic surgery for endometriosis should perform a systematic inspection of the pelvis.**

28. **During a diagnostic laparoscopy, consider taking a biopsy of suspected endometriosis:**
    - to confirm the diagnosis of endometriosis (be aware that a negative histological result does not exclude endometriosis)

ISBN 978-1-4731-2661-9

- to exclude malignancy if an endometrioma is treated but not excised.

29. If a full, systematic laparoscopy is performed and is normal, explain to the woman that she does not have endometriosis, and offer alternative management.

ISBN 978-1-4731-2661-9

10 Staging systems

Review question: What is the effectiveness of using endometriosis-staging systems to guide treatment of endometriosis?

10.1 Introduction

Women with endometriosis would benefit from the adoption of a robust classification or staging system that allows immediate description of the severity of the condition, correlates with symptoms, a tool to guide treatment, reliable assessment of therapeutic outcomes and is a useful tool in clinical trials. Due to the complex nature of the condition and women's wide variability of clinical presentations and outcome needs, such as fertility preservation and pain relief, a single staging system that fits all presents a challenge.

A number of classification systems have been developed for staging endometriosis and are in use. They are usually based upon the anatomic location, severity and depth of disease. For example, a widely used system is the American Society for Reproductive Medicine revised classification which uses 4 stages based on description of lesions at laparoscopy. It is unclear whether there is an accepted classification system that can allow assessment of superficial versus deeply infiltrating disease as well as the structures affected, and correlate findings to surgical complexity and outcomes in terms of guiding treatment to improve pain or other symptoms, and reduce recurrence and complication rates by stage.

The aim of this chapter is to review the literature to assess what is the effectiveness of using endometrial-staging systems to guide treatment of endometriosis. The specific treatment of women with infertility associated with endometriosis was outside the scope of this guideline.

10.2 Description of clinical evidence

The objective of this review was to determine if it is clinically useful to formally classify the stages of endometriosis with a view to guiding management decisions and improving patient outcomes. For full details, see the review protocol in Appendix D.

No relevant study was identified that compared the use of any staging system with other staging systems or with not using it. See also the study selection flow chart in Appendix F and study exclusion list in Appendix H. Summary of included studies.

No study was included in this systematic review.

10.3 Clinical evidence profile

Not applicable.

10.4 Economic evidence

No health economic studies were found relevant to this question, and therefore no health economic modelling was conducted for this question.

10.5 Clinical evidence statements

No relevant study addressing the question of this systematic review was identified.

ISBN 978-1-4731-2661-9

10.6 Evidence to recommendations

10.6.1 Relative value placed on the outcomes considered

The aim of the review was to determine if it is clinically useful to formally classify the stages of endometriosis with the aim of guiding management decisions and improving patient outcomes. Therefore, the Committee considered that it was not the outcome of the staging that was critical, but the management decisions based on different staging systems and the outcomes following particular treatments based on those systems. The priority treatment-based outcomes were pain, quality of life and fertility.

10.6.2 Consideration of clinical benefits and harms

No study was included in this review. There is not enough evidence to show the effectiveness of using staging systems to guide treatment of pain associated with endometriosis.

The Committee agreed that, in their clinical experience, the correlation between the severity of a woman's symptoms and the extent of endometriosis was not good. The Committee also noted that from experience, good visual documentation during laparoscopy and written description can provide more useful information to guide treatment than a staging system categorisation which might be more appropriate for quantifying endometriosis in research situations. Therefore the Committee concluded that treatment for endometriosis should be based on the women's symptoms and not only on the stage of the endometriosis. There are many staging systems in use for endometriosis but these are often related to infertility with few validated staging systems related to other symptoms such as pain. However, the stage of endometriosis does not always correlate with the presence of other symptoms. For example, a woman classified as having 'stage I or II' endometriosis may present with severe pain and a woman classified as 'stage III or IV' may present with minimal or mild pain. As there is no evidence to show the benefit of using staging systems to guide treatment, the expert opinion of the Committee was that the decision for treatment should be based on the woman's symptoms and not the endometriosis stage.

10.6.3 Consideration of economic benefits and harms

The use of a staging system by itself does not invoke an opportunity cost, but the use of a staging system that guides treatment decisions will invoke an opportunity cost – treatment undertaken under staging system A that would not have been undertaken under staging system B. However, with no evidence comparing staging system A vs. staging system B, it is not possible to estimate the size of this opportunity cost, nor to determine the most cost-effective staging system. It is possible that there is a minor direct saving if some staging systems are proprietary, but the most common systems are not.

As the Committee chose not to make recommendations that differed from current practice in an economically significant way, the recommendations do not carry a high resource impact.

10.6.4 Quality of evidence

No study was identified to address the review question.

10.6.5 Other considerations

The Committee emphasised the importance of careful visualisation of the entire pelvis during laparoscopy and of documenting the appearance and site of all endometriotic lesions, including any vaginal lesions. The Committee was of the opinion that there is no classification systems with the specific aim of guiding treatment of endometriosis. The

ISBN 978-1-4731-2661-9

Committee agreed that staging systems for endometriosis were designed to guide treatment for fertility with a correlation between staging and likelihood of achieving a spontaneous pregnancy. The assessment of fertility in women with endometriosis and therefore the staging in relation to fertility is outside the scope of this guideline. For the assessment of fertility related to endometriosis see NICE guideline on fertility (CG156).

The Committee agreed that the treatment of patients with endometriosis should be based on symptoms rather than staging. The Committee noted that staging systems do not accurately correspond to a level of pain and complications that the women are experiencing. Some women have symptoms in excess of the stage of the disease, whereas other women have a higher stage but fewer symptoms.

The Committee agreed that current commissioning of endometriosis services are related to the staging systems with funding allocated for the treatment of women assessed as having stage III or IV endometriosis. The Committee agreed that this was not necessarily appropriate since women with a lower stage could have severe symptoms requiring intervention and vice versa. They therefore noted that the current commissioning of endometriosis services need to take the symptoms rather than the stage into consideration.

The Committee discussed possible options for further research on this clinical issue, but decided not to propose a research recommendation. They agreed that it would always be difficult to have an agreed system that would classify women with endometriosis to 1 particular treatment choice. The treatment strategy would always need to be tailored to the individual women and her priorities and preferences rather than to a particular stage of the condition.

10.6.6 Key conclusions

The Committee concluded that current staging systems cannot guide decisions about treatments because there is no clear correlation between stage and severity of symptoms (for example, severe pain and low stage). The Committee agreed treatment decisions need to be based on the symptoms and be tailored to individual needs, preferences and priorities in terms of pain and fertility preservation. The Committee agreed that it was important to note that the assessment of fertility was outside the scope and that fertility related staging is therefore not covered. Healthcare professionals and women with endometriosis should therefore consult the NICE guideline on fertility (CG156).

10.7 Recommendations

30. **Offer endometriosis treatment according to the woman's symptoms, preferences and priorities, rather than the stage of the endometriosis.**

31. **When endometriosis is diagnosed, the gynaecologist should document a detailed description of the appearance and site of endometriosis.**

ISBN 978-1-4731-2661-9

11 Management strategies

11.1 Pharmacological management

11.1.1 Analgesics

Review question: What is the effectiveness of analgesics for reducing pain in women with endometriosis, including recurrent and asymptomatic endometriosis?

11.1.1.1 Introduction

Pain is the most debilitating and common symptom of endometriosis. Endometriosis may cause cyclical pelvic pain, typically during menstruation, and often starting a few days before a woman's period. Referred pain to the back and legs is common. Apart from acute pain during menstruation, women may also experience non-cyclical pain, deep pain during sexual intercourse, and pain associated with bowel and bladder functions. For many women, pain becomes persistent or chronic.

Most women who experience menstrual pain and who would like pharmacological analgesia will buy over-the-counter medications or be prescribed simple analgesics such as paracetamol and non-steroidal anti-inflammatory drugs (NSAIDs), for example, ibuprofen, naproxen or aspirin. Mefanamic acid, another NSAID, is also commonly chosen for menstrual pain. For moderate to severe pain, weak opioids such as codeine are often used but the side effects of these are often limiting; constipation in particular may aggravate endometriosis symptoms. Stronger medication such as morphine is also prescribed if the pain is severe and does not respond to other treatments.

Symptomatic management of pain using analgesics is thus very important for women with endometriosis. Because of disease recurrence and potential chronicity of pain, women need access to analgesics throughout a lifetime living with endometriosis.

11.1.1.2 Description of clinical evidence

The objective of this review is to determine the clinical and cost effectiveness of analgesics in reducing pain in women with endometriosis.

For full details, see review protocol in Appendix D.

One study was included (Kauppila 1985) that used a crossover design to evaluate the effect of non-steroidal anti-inflammatory drugs (NSAIDs) compared with placebo in 24 women with 'moderate' to 'very severe' painful menstrual periods secondary to endometriosis. Endometriosis was diagnosed by pelvic examination, or by visualisation (for example, laparoscopy or laparotomy). One group of women received naproxen tablets for 2 menstrual cycles and then crossed over to placebo for 2 further menstrual cycles. The second group received placebo for the first and second menstrual cycles, then crossed over to naproxen sodium for 2 further menstrual cycles. Both groups received 275 mg naproxen tablets (1 or 2 tablets 4 times a day).

Results are presented from the first treatment period for 20 women who used a questionnaire immediately after each menstrual cycle to self-record outcomes of pain severity, use of supplementary analgesia and unintended effects from treatment. For severity of pain a score (range 1–3) was used where 'mild improvement' was scored as 1, 'moderate improvement' was scored 2 and 'excellent relief' was scored 3. It is not clear how the questionnaire was developed or validated.

ISBN 978-1-4731-2661-9

No evidence was identified for the critical outcome of quality of life or for the important outcomes of effect on daily activities, absence from work or school, number of women requiring more invasive treatment and participant satisfaction with treatment.

Evidence is summarised in the clinical GRADE evidence profile below Table 53. See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G.
Summary of included studies

A summary of the studies that were included in this review are presented in Table 52.

Table 52: Summary of included studies

Study	Intervention/ comparison	Population	Outcomes	Comments
Kauppila 1985 Finland	NSAIDs (Naproxen Sodium)/placebo	20 women with endometriosis classified using American Fertility Society (AFS) criteria	• overall pain relief • supplementary analgesia needed • unintended effects from treatment	Crossover trial Study funded by pharmaceutical company

NSAIDs: non-steroidal anti-inflammatory drugs; RCT: randomised controlled trial

11.1.1.3 Clinical evidence profile

The clinical evidence profile for this review question (NSAIDs for treatment of endometriosis) is presented in Table 53.

Table 53: Summary clinical evidence profile: analgesics versus placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	Abso-lute effect	No. of partici-pants (studies)	Evidence quality (GRADE)	Comments
	Assumed risk	Corres-ponding risk					
	Placebo	Interven-tion (analge-sics)					
Overall pain relief (self-measured by question-naire)	625 per 1,000	906 per 1000 (512-1,000)	RR 1.45 (0.82 to 2.57)	281 more per 1,000 (from 113 fewer-981 more)	19 (1 study)	⊕⊖⊖⊖ Very low[1,2,3,4]	Measured with 3 point scale question-naire
Unintended effects from treatments (hypo-menorrhea, diarrhoea. increased diuresis,	778 per 1,000	366 per 1,000	RR 0.47 (0.2 to 1.1)	412 fewer per 1,000 (from 622 fewer-	20 (1 study)	⊕⊖⊖⊖ Very low[1,3,4]	

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	Abso-lute effect	No. of partici-pants (studies)	Evidence quality (GRADE)	Comments
	Assumed risk	Corres-ponding risk					
headache, epigastric pain nausea, tremor and dizziness)				78 more)			
Supplement ary analgesia needed	222 per 1,000	91 per 1,000 (9-849)	RR 0.36 (0.04 to 3.35)	160 fewer per 1,000 (from 240 fewer-587 more)	19 (1 study)	⊕⊖⊖⊖ Very low[1,3,5]	Additional medication needed

CI: confidence interval; RR: risk ratio
1 Unclear allocation concealment, sequence generation and selective reporting
2 Unvalidated tool used for pain assessment
3 n=24 randomised, n=20 analysed (19 for overall pain relief and supplementary analgesia needed), no clear exclusion criteria hence high risk of selection bias
4 Wide confidence interval
5 Very wide confidence interval

11.1.1.4 Economic evidence

No economic evidence was found on the use of analgesics in women with endometriosis.

Consequently, data from NICE CG173 (neuropathic pain) was used to inform an economic model that is described in more detail in Appendix K.

The economic cost of analgesics is very difficult to quantify. Although the drugs and the dosing regimen are normally very well understood, compliance and indirect costs (such as additional GP visits) can create uncertainty over the 'true' cost of prescribing one drug over another. In addition, many patients will self-medicate with over-the-counter analgesics, meaning that the cost to the NHS of recommending over-the-counter medicines such as paracetamol is only a fraction of the cost of recommending prescription-only medicines such as codeine (moreover, over-the-counter medicines tend to be less expensive to begin with).

Table 54 gives the direct cost of the 3 analgesics considered in the economic model for endometriosis (selected because of the availability of evidence on their cost and effectiveness). Table 55 gives indicative costs of all other analgesics specified in the protocol. The true economic cost of prescribing one over the other depends on factors not included in this table, including side effects, compliance and indirect costs.

The cost of 'Generic' analgesia is given as the cost of aspirin. Aspirin has a slightly higher cost than some other NSAIDs according to the electronic drug tariff; for example, Ibuprofen costs £0.86 for 24 400g tabs giving an annual cost of £40.05 and Naproxen costs £0.93 for 28 250g tabs giving an annual cost of £36.37. Nevertheless, it was thought appropriate to use the cost of aspirin as it is probably the most commonly prescribed NSAID, and the slightly higher cost is expected to offset indirect costs from drug prescription, such as side-effects, which are not included in Electronic Drug Tariff prices.

ISBN 978-1-4731-2661-9

Table 54: Estimated annual direct cost of analgesics included in economic model

Treatment	Cost	Source
Codeine	£563.42	NICE CG 173
Tramadol	£542.13	NICE CG 173
'Generic' analgesia[a]	£93.15	Electronic Drug Tariff, retrieved 14/12/16

(a) There is a lack of clarity in the evidence regarding exactly which analgesic was given to patients in a handful of trials – it appears to be simple NSAIDs, but to avoid confusion it is labelled in the model as 'generic' treatment

Table 55: Estimated annual direct costs of analgesics specified in protocol

Compound	Cost per annum (min)[b]	Source
Paracetamol	£87.60	Electronic Drug Tariff, retrieved 14/12/16
Diclofenac	£29.20	Electronic Drug Tariff, retrieved 14/12/16
Ibuprofen	£36.50	Electronic Drug Tariff, retrieved 14/12/16
Naproxen	£54.75	Electronic Drug Tariff, retrieved 14/12/16
Celecoxib	£394.20	Electronic Drug Tariff, retrieved 14/12/16
Mefenamic acid (tabs)	£219.00	Electronic Drug Tariff, retrieved 14/12/16
Mefenamic acid (caps)	£186.15	Electronic Drug Tariff, retrieved 14/12/16
Etoricoxib	£299.30	Electronic Drug Tariff, retrieved 14/12/16
Indomethacin	£52.93	Electronic Drug Tariff, retrieved 14/12/16
Tolfenamic acid (as Clotam Rapid)	£698.98	Electronic Drug Tariff, retrieved 14/12/16
High-dose aspirin	£536.55	Electronic Drug Tariff, retrieved 14/12/16
Co-codamol	£169.73	Electronic Drug Tariff, retrieved 14/12/16
Co-codaprin	£2,642.60	Electronic Drug Tariff, retrieved 14/12/16
Co-dydramol	£114.98	Electronic Drug Tariff, retrieved 14/12/16
Dyhydrocodeine[a]	£1,053.03	https://www.ukmeds.co.uk/treatments/pain-relief/dihydrocodeine-30mg-tablets/
Buprenorphine (as Temgesic)	£202.58	Electronic Drug Tariff, retrieved 14/12/16

(a) Whereas all other costings taken from Electronic Drug Tariff, dyhydrocodeine costs were not available and were estimated from online pharmacy costs.

*(b) The cost is given by taking the average of the minimum and maximum daily dose multiplied by 365. For example, if the recommendation was to take 1–2 capsules of a drug 4–6 times a day, we assume the average daily dose is 1.5*5=7.5 capsules per day.*

The economic model suggests that no analgesic is likely to be better than hormonal treatment; hormonal treatment is likely to be both more effective and cheaper than the best analgesics. These results are demonstrated in Table 56. The table shows that Tramadol likely dominates no treatment – being both cheaper and more effective – but that the next most effective set of analgesics are outside the range which would normally be considered for the NICE cost-effectiveness threshold of around £20,000

NSAIDs were excluded from most runs of the model; the evidence for their effectiveness was weak and contradictory (and the evidence upon which this was based was not clear in specifying which exact analgesic was used; NSAIDs were inferred from a description of the side effects). If the results for NSAIDs are accepted at face value, they would be more effective than hormonal treatment at a slightly higher cost, which would nonetheless be cost-effective at £20,000/quality adjusted life year (QALY) threshold. The Committee discussed how this could well be important evidence highlighting the effectiveness of NSAIDs versus other analgesics.

ISBN 978-1-4731-2661-9

Table 56: Cost and effectiveness of all treatment strategies containing an analgesic

Treatment	Cost	QALY	ICER	Pr. cost-effective vs. no treatment (£20k/QALY)	Pr. cost-effective vs. no treatment (£30k/QALY)
Empirical Diagnosis & No Treatment	£22,752.60	18.120	Base Case	100.0%	100.0%
Empirical Diagnosis & Tramadol	£21,875.58	18.174	-£16,159.27	85.7%	86.8%
Empirical Diagnosis & Codeine (as Morphine)	£22,776.51	18.180	£161,978.83	86.8%	87.9%
Laparoscopy & Codeine (as Morphine)	£33,431.95	18.200	£518,261.74	75.8%	80.2%

(c) QALY: quality adjusted life year; ICER: incremental cost-effectiveness ratio; Pr.: probability NICE CG 173 does not have QALY data on codeine, so it is assumed the opioid codeine behaves (as morphine), a different opioid, for the purpose of determining between-class performance

11.1.1.5 Clinical evidence statements

Very low quality evidence from 1 crossover RCT (n=20) showed that there was no difference in overall pain relief, unintended effects or need for supplementary analgesia when women with endometriosis received naproxen sodium compared to placebo for 2 menstrual cycles, although there was uncertainty around the estimate.

11.1.1.6 Evidence to recommendations

11.1.1.6.1 Relative value placed on the outcomes considered

The Committee prioritised pain relief, health-related quality or life and adverse events from analgesics (particularly those leading to withdrawal from treatment) as critical outcomes.

The Committee also discussed the need to take further supplementary analgesia, which was another outcome that was reported. No evidence was identified that reported on health-related quality of life.

11.1.1.6.2 Consideration of clinical benefits and harms

Pain is a common symptom of endometriosis and, when severe and/or persistent, can be completely debilitating, affecting one's ability to perform routine daily activities, greatly limiting lifestyle and quality of life.

The Committee acknowledged that analgesia would only provide symptomatic relief of pain, rather than addressing any underlying pathology, but that effective pain relief can provide an alternative to more invasive treatment. The Committee noted that hormonal therapies used to treat endometriosis may take at least 1 menstrual cycle to become effective. For this reason, pain relief medication may be used until the long-term treatment begins to work.

The Committee noted that some women might tolerate significant harms associated with side effects of analgesics in order to have respite from their pain and that this trade off was variable depending on the severity of the woman's symptoms and her individual circumstances.

ISBN 978-1-4731-2661-9

11.1.1.6.3 Consideration of economic benefits and harms

The Committee acknowledged that hormonal treatment was likely to be more cost-effective than the best analgesics but reflected that this did not exclude giving an analgesic with another kind of treatment as, in general, analgesics were not thought to interact with other forms of treatment. The Committee also noted that analgesics might be considered cost-effective in the absence of other treatments. However, as there was no direct evidence on the effectiveness of analgesics in combination with other treatments for endometriosis the Committee made it clear that clinical judgement would be required if considering analgesics in combination with other treatments (e.g. hormonal or surgical treatments).

Although there are no results for the impact of analgesics on fertility (as this was not modelled), the Committee considered that the presence or absence of analgesics would be unlikely to alter a woman's fertility and have a relatively smaller impact on fertility than other treatment options considered in this guideline. The Committee acknowledged some limited evidence that NSAIDs might inhibit ovulation if taken continuously during the cycle, (making conception less likely), but noted that if taken during the period, would not have an effect on ovulation. Members further pointed out that severe pain might reduce the likelihood of intercourse and hence analgesics might improve the chance of conception. Overall the Committee concluded that the impact of analgesics on fertility (especially NSAIDs) was not sufficiently researched to underpin a recommendation.

Estimating the resource impact of analgesics is difficult as many women will chose to self-medicate if prescribed over-the-counter analgesia (as this can often work out cheaper for both the woman and the NHS). The Committee described how the general principle of their recommendations – trialling cheap medication and considering more expensive analgesia if this failed – was current NHS practice, and so the recommendations are unlikely to represent a significant resource impact.

11.1.1.6.4 Quality of evidence

The available evidence was drawn from a single small trial conducted in 1985 and was of very low quality. A self-reported questionnaire to assess pain was used, although the validity of the pain scoring system was unclear. While the study indicated that 24 women were randomised, the results for only 20 women were reported for unintended effects of treatment and 19 for overall pain relief and for supplementary analgesia needed. There were other methodological flaws such as unclear allocation concealment and unclear reporting of exclusion criteria. The direction of the effect for overall pain relief, unintended effects and need for supplementary analgesia outcomes was in favour of naproxen sodium but, due to the small sample size, the study was underpowered and outcome effects had wide confidence intervals (CIs). No evidence was available for the other outcomes prioritised and no other relevant evidence assessing the effectiveness of any other type of analgesic for endometriosis-related pain was available.

The Committee considered that the small number of women included in the study and its short duration made it difficult to draw any valid conclusions. The Committee agreed that although there is no good evidence for use of analgesics in management of acute pain specific to endometriosis, there is robust evidence of effectiveness of analgesics for pain management in other areas and hence gave little weight to the limited evidence.

11.1.1.6.5 Other considerations

Due to the poor quality and limited evidence available, the Committee based their decisions on consensus and the experience and expertise of its members.

The Committee discussed the Pain Ladder developed by the World Health Organization (WHO) for analgesia for cancer-related pain but which has since been adopted for acute and chronic non-malignant pain relief. This describes a 3-step progressive approach to use of

ISBN 978-1-4731-2661-9

pharmacologic agents proportional to the level of pain reported. The initial step uses oral administration of non-opioids such as paracetamol or NSAIDs. If pain is not controlled, then mild opioids such as codeine are tried and, as a last step, strong opioids such as morphine are used until the patient's pain is alleviated. One benefit of the stepped approach is that adverse events can be discovered throughout the process.

The Committee discussed whether the addition of an opioid analgesic could be considered if pain was not adequately controlled after a trial period. However, the potential adverse effects of opioid analgesia, such as dependency, were recognised, given the chronic nature of endometriosis-related pain and, particularly, constipation. Therefore, the Committee concluded that a referral for diagnosis might be more appropriate and that there were other treatment options available.

The Committee also considered whether a research recommendation should be drafted for this topic. They agreed that research into analgesia in the management of pain related to endometriosis is not a priority for this guideline because there is sufficient indirect evidence from other conditions available to draw upon.

The Committee considered whether any different recommendations were necessary for adolescent women but concluded that none were required.

11.1.1.6.6 Key conclusions

The Committee concluded that a short trial of analgesics for first line management of pain in women with endometriosis-related pain is appropriate.

11.1.1.7 Recommendations

32. For women with endometriosis-related pain, discuss the benefits and risks of analgesics, taking into account any comorbidities and the woman's preferences.

33. Consider a short trial (for example, 3 months) of paracetamol or a non-steroidal anti-inflammatory drug (NSAID) alone or in combination for first-line management of endometriosis-related pain.

34. If a trial of paracetamol or an NSAID (alone or in combination) does not provide adequate pain relief, consider other forms of pain management and referral for further assessment.

11.1.2 Neuromodulators (neuropathic pain treatment)

Review question: What is the effectiveness of neuromodulators for treating endometriosis, including recurrent and asymptomatic endometriosis?

11.1.2.1 Introduction

Neuromodulators, otherwise known as neuropathic analgesics, are used mainly by pain specialists and general practitioners (GPs) in the management of chronic, also known as, persistent pain. Neuromodulators differ from conventional analgesics such as NSAIDs in that they primarily affect the central nervous system's modulation of pain, rather than peripheral meditators of inflammation. An overactive and hypersensitive nervous system contributes to the development and maintenance of chronic pain. Neuromodulators exert their effects via their modulation of this overactive and hypersensitive nervous system.

Many neuromodulators were originally developed with different aims, for example, as antidepressants or anticonvulsants. The main classes of neuromodulators are: the tricyclic antidepressants, for example, amitriptyline and nortriptyline; the selective serotonin re-uptake

ISBN 978-1-4731-2661-9

inhibitors such as duloxetine; and the gabapentinoids (gabapentin and pregabalin). Under this heading we also considered capsaicin, ketamine, local anaesthetics (lidocaine) and nerve blocks. Certain opioid medications, such as tramadol and tapentadol, also have neuromodulating properties.

These medicines may also have important other effects, depending on their dose, on other related conditions that may be concurrently present, such as anxiety and/or depression. NICE already recommends a choice of amitriptyline, duloxetine, gabapentin or pregabalin as the initial treatment for neuropathic pain (CG 173).

Understanding the effectiveness of neuromodulators for women with endometriosis is important as, if useful, they might reduce the burden of pain and/or side effects from other medications, or offer an alternative to other types of treatment such as hormonal. If effective, they might reduce the need for surgery and prevent or reduce the chronicity of pain with its far-reaching consequences.

11.1.2.2 Description of clinical evidence

The objective of this review is to determine the clinical and cost effectiveness of neuromodulators to improve outcomes in women with endometriosis.

For full details, see review protocol in Appendix D.

We looked for systematic reviews, randomised and comparative observational studies assessing the effectiveness of neuromodulators in the management of endometriosis of any stage or severity. These may also include suspected diagnoses as described in detail in the protocol.

Two trials were identified that used local anaesthetics with a procedure called perturbation, which involves the insertion of a thin plastic catheter in the cervical canal. This catheter is then used to infuse the local anaesthetic through the uterine cavity and is then pertubated into the peritoneal cavity.

One trial was conducted in Sweden (Wickström 2013) with a number of associated published abstracts and 1 further full article are both associated with this particular trial (Edelstam 2012, Wickström 2012a, 2012b, 2012c). The local anaesthetic used in this trial was lidocaine. The second trial was conducted in Egypt, using the same procedure but with a different local anaesthetic bupivacaine (Shokeir 2015). In both trials the inclusion criteria included the requirement that endometriosis had been confirmed by laparoscopy.

Both trials reported pain as an outcome (as indicated on the visual analogue scale [VAS]). One of them also reported the rate of women who were overall satisfied with the procedure. The other trial also reported health-related quality of life as measured by the Endometriosis Health Profile-30 (EHP-30) as well as recurrence and need for other therapies. Fertility outcomes cannot be assessed because both studies excluded women who intended to become pregnant within the forthcoming year.

No further evidence was identified for any other type of neuromodulator or neuropathic analgesia.

Evidence for the outcomes from these trials is summarised in the clinical GRADE evidence profile below (Table 58). See also the study selection flow chart in Appendix F study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G. Summary of included studies

A brief summary of the studies that were included in this review is presented in Table 57.

ISBN 978-1-4731-2661-9

Table 57: Summary of included studies

Study	Intervention/ Comparison	Population	Outcomes	Comments
Shokeir 2015 Egypt	Pertubal 10ml diluted bupivacaine infusion (0.25%) versus placebo infusion (sterile water) N=62	Women with chronic pelvic pain for at least 6 months who had a pain score of at least 5 (on a VAS ranging from 0 to 10 cm) and had laparoscopically confirmed endometriosis of any stage	• Pain as measured on a VAS rating measured at 1, 2 and 3 months • Overall level of satisfaction at 3 months	• Small sample size • Short follow-up length
Wickström 2012, 2013 Sweden	Pertubation of 10 ml lidocaine/lignocaine versus placebo N=42	Women with chronic pelvic pain for at least 6 months who had a pain score of at least 5 (on a VAS ranging from 0 to 10 cm) and had laparoscopically confirmed endometriosis of any stage	• VAS of pain (at 3, 6, 9 and 12 months) – categorised as a VAS score that is improved by ≥50% • EHP-30 (health-related quality of life score specific to endometriosis measured at 6 and 12 months • Recurrence at 12 months • Escalating pain with need for other therapies at 12 months	• Small sample size • Flow of participants not easy to follow (2 different types of analyses do not match) • Minimally important difference is set very high and does not correspond to the continuously analysed pain score • Large loss to follow-up at 12 months

N: number of participants in study; VAS: Visual analogue scale

11.1.2.3 Clinical evidence profile

The clinical evidence profile for this review question is presented in Table 58.

Table 58: Summary clinical evidence profile: Local anaesthetic (pertubation) versus placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Placebo	Local anaesthetic			
Pain score – VAS >50% improved – at 3 months	56 per 1,000	375 per 1,000 (52 to 1,000)	RR 6.75 (0.94 to 48.57)	42 (1 study)	⊕⊕⊖⊖ Low[1,2]
Pain score – VAS >50% improved – at 6 months	56 per 1,000	167 per 1,000 (21 to 1,000)	RR 3 (0.37 to 24.61)	42 (1 study)	⊕⊖⊖⊖ Very low[1,3]

ISBN 978-1-4731-2661-9

Pain score–- VAS >50% improved – at 9 months	0 per 1,000	0 per 1,000 (0 to 0)[6]	Peto OR 6.01 (0.35 to 102.4)	42 (1 study)	⊕⊖⊖⊖ Very low[1,3]
Pain score – VAS >50% improved – at 12 months	0 per 1,000	0 per 1,000 (0 to 0)[6]	Peto OR 6.81 (0.84 to 51.68)	42 (1 study)	⊕⊖⊖⊖ Very low[1,3]
Pain – VAS continuous – at 1 month	-	The mean pain - vas continuous - at 1 month in the intervention groups was 1.3 lower (2.18 to 0.42 lower)	MD -1.3 (-2.18 to -0.42)	60 (1 study)	⊕⊕⊖⊖ Low[4,5]
Pain – VAS continuous – at 2 months	-	The mean pain - vas continuous - at 2 months in the intervention groups was 1.9 lower (2.92 to 0.88 lower)	MD -1.9 (-2.92 to -0.88)	60 (1 study)	⊕⊕⊕⊖ Moderate[4]
Pain – VAS continuous – at 3 months	-	The mean pain - vas continuous - at 3 months in the intervention groups was 2.3 lower (3.46 to 1.14 lower)	MD -2.3 (-3.46 to -1.14)	60 (1 study)	⊕⊕⊕⊖ Moderate[4]
Rate of satisfaction with treatment at 3 months	67 per 1,000	733 per 1,000 (189 to 1,000)	RR 11 (2.83 to 42.7)	60 (1 study)	⊕⊕⊕⊖ Moderate[4]
Rate of recurrence at 12 months	0 per 1,000	0 per 1,000 (0 to 0)[6]	Peto OR 6.01 (0.34 to 102.42)	42 (1 study)	⊕⊖⊖⊖ Very low[3,6]
Escalating pain with a need for other therapies at 12 months	167 per 1,000	42 per 1,000 (5 to 368)	RR 0.25 (0.03 to 2.21)	42 (1 study)	⊕⊖⊖⊖ Very low[1,3]

CI: confidence interval; RR: risk ratio; Pero OR: Peto odds ratio; MD: mean difference; VAS: visual analogue scale
1 The patient flow is a little unclear and there is a difference in results using 2 types of analyses. The categorisation of the pain scale favours the treatment group and there are conflicting results with another pain outcome used in the same trial.
2 The CI is large ranging from no effect to effect favouring the treatment.
3 The CI for this outcome ranges from an affect favouring placebo to an effect favouring the treatment. There is therefore too much uncertainty around this effect.
4 Some of the reported CIs seem to be incorrectly reported.
5 The CI ranged from a high effect to no appreciable benefit.
6 Due to zero events in the control group Peto OR were used rather than Risk Ratios because this method performs well when events are very rare (Bradburn 2007). This means that the risk difference is reported with confidence intervals.

Other reported findings – EHP-30 (endometriosis-related quality of life)

Quality of life scores were reported as medians with interquartile ranges and therefore could not be graphically presented as forest plots. They are presented in Table 59 below.

ISBN 978-1-4731-2661-9

Table 59: Clinical evidence table: local anaesthetic (pertubation) versus placebo - endometriosis health related quality of life scores

	Change after 6 months			Change after 12 months		
EHP-30 dimension	**Lidocaine median IQR**	**Placebo median IQR**	**Mann-Whitney U-test p-value**	**Lidocaine median IQR**	**Placebo median IQR**	**Mann-Whitney U-test p-value**
Pain	-13.6 (-27.3 to 2.3)	-11.4 (-22.7 to 2.3)	0.99	-8 (-29.5 to 2.3)	-11.4 (-20.5 to 4.5)	0.69
Control and powerless-ness	-8.3 (-33.3 to 2.1)	-6.3 (-35.4 to 2.1)	0.84	-12.5 (-37.5 to -8.3)	-20.8 (-41.7 to 0)	0.74
Emotional wellbeing	-4.2 (-37.5 to -4.17)	-12.5 (-20.8 to -6.25)	0.99	-20.8 (-37.5 to 0)	-12.5 (-25 to 4.17)	0.63
Social support	-18.8 (-31.25 to 0)	-6.3 (-12.5 to -6.25)	0.034	-12.5 (-37.5 to 0)	-6.3 (-31.25 to 12.5)	0.50
Self-image	-8.3 (-16.7 to 0)	0 (-16.67 to 8.33)	0.24	-8.3 (-16.7 to 0)	0 (-16.7 to -0)	0.57
Sexual intercourse	-10 (-25 to 10)	-5 (-10 to 5)	0.24	-7.5 (-15 to 5)	-7.5 (-20 to 7.5)	0.97

(a) EHP-30: Endometriosis Health Profile-30; IQR: interquartile range

11.1.2.4 Economic evidence

No economic evidence was found on the use of neuromodulators in women with endometriosis.

As no evidence was found on the use of neuromodulators in women with endometriosis, the effectiveness of these treatments was calculated from. Consequently, not all treatments listed in the protocol could be included in the economic model.

Table 60: Annual cost of neuromodulator treatments included in the model

Treatment	Cost per year	Source
Amitriptyline	£227.25	CG173
Nortriptyline	£1,086.43	CG173
Duloxetine	£871.00	CG173
Venlafaxine	£383.44	CG173
Capsaicin patches	£1,210.80	CG173
Gabapentin	£365.62	CG173
Pregabalin	£1,000.77	CG173
Topiramate	£63.07	CG173

(a) CG96: Neuropathic pain in adults: pharmacological management in non-specialist settings

Table 61 demonstrates which neuromodulators might be selected as a cost-effective treatment on average. Both amitriptyline and gabapentin perform well relative to an incremental cost-utility ratio (ICER) of £20,000/QALY and are cheap enough that a diagnostic strategy of 'empirical diagnosis' – treating based on symptoms rather than a definitive diagnosis - can be pursued. However, this is only with reference to the class of neuromodulators; the main economic model indicates that neuromodulators are neither cheap enough to be considered in preference to hormonal treatment nor effective enough to be considered in preference to surgery. Given that there are some women who cannot tolerate hormonal therapy (usually because they are seeking a pregnancy, which is

ISBN 978-1-4731-2661-9

discussed below) these results might be important, as it is possible neuromodulators will be cost-effective in these women. This is relevant as, if a woman cannot have hormonal therapy but responds to neuromodulators, then it is unlikely surgery will be cost-effective for this woman.

Table 61: Cost and effectiveness of all non-dominated treatment strategies containing a neuromodulator treatment

Treatment	Cost	QALY	ICER	Pr. cost-effective vs. no treatment (£20k/QALY)	Pr. cost-effective vs. no treatment (£30k/QALY)
Empirical Diagnosis & No Treatment	£22,752.60	18.120	Base Case	N/A	N/A
Empirical Diagnosis & Amitriptyline	£21,702.24	18.340	-£4,774.17	92.3%	95.6%
Empirical Diagnosis & Gabapentin	£22,734.50	18.399	£17,458.76	94.5%	95.6%
Peritoneal biopsy & Gabapentin	£25,400.16	18.401	Extendedly Dominated	86.8%	89.0%
Empirical Diagnosis & Pregabalin	£27,488.25	18.448	£96,666.23	85.7%	94.5%

(a) Note: ICER: incremental cost-effectiveness ratio; QALY: quality adjusted life year

It was thought that neuromodulators would not have a positive effect on women seeking to conceive and some neuromodulators would be harmful to a developing fetus. For these reasons, neuromodulators were not considered in an analysis of women where infertility was the main reason for their seeking treatment.

11.1.2.5 Clinical evidence statements

No evidence was identified that addressed the effectiveness of commonly used neuropathic analgesics.

11.1.2.5.1 Pertubation of lidocaine vs. placebo

Pain up to 12 months

Very low to low quality evidence from 1 randomised controlled trial (RCT) with 42 women with endometriosis suggested higher rates of women who reported a significant improvement in pain associated with pertubation of lidocaine compared to placebo at 3, 6, 9 and 12 months. However the uncertainty around this improvement was too large to draw clear conclusions about its clinical effectiveness.

EHP-30

Very low quality evidence from 1 RCT with 42 women with endometriosis reported no clear differences between women treated with lidocaine compared to placebo at 6 and 12 months for the subscales pain, control and powerlessness, emotional well-being, self-image and sexual intercourse. A small difference on the social support subscale was reported at 6 but not 12 months (Table 59).

ISBN 978-1-4731-2661-9

Recurrence at 12 months

Very low quality evidence from 1 RCT (N=42) suggested a higher rate of recurrence in those receiving lidocaine compared to those in the placebo group. However, the uncertainty around this effect was too large to draw clear conclusions about this finding.

Escalating levels of pain with a need for other therapies at 12 months

Very low quality evidence from 1 RCT (N=42) suggested that there were fewer women needing other treatments in the lidocaine group compared to the control group. However, there was too much uncertainty around this effect to draw clear conclusions from these findings

11.1.2.5.2 Pertubation of bipuvacaine vs. Placebo

Pain up to 3 months

Moderate to high quality evidence from 1 randomised controlled trial (RCT) conducted with 60 women who have endometriosis reported improvements in pain at 1, 2 and 3 months associated with bipuvacaine pertubation. However, the uncertainty around this effect make it difficult to draw conclusions about the clinical significance of this finding.

Satisfaction with treatment at 3 months

High quality evidence from 1 RCT conducted with 60 women who have endometriosis showed a higher rate of satisfaction with bipuvacaine treatment compared to placebo.

11.1.2.6 Evidence to recommendations

11.1.2.6.1 Relative value placed on the outcomes considered

All reported outcomes (pain, endometriosis health profile, recurrence, satisfaction and need for further therapies) are critical for decision-making. However, the Committee did not place trust in the evidence for these outcomes since pertubation with local anaesthetic is not used in current practice.

11.1.2.6.2 Consideration of clinical benefits and harms

The Committee agreed that it was disappointing that there was no clinical evidence for the effectiveness of commonly used neuromodulators.

They recognised that there was much useful guidance in the NICE guidance Neuropathic pain in adults: pharmacological management in non-specialist settings (Clinical Guideline 96). The Committee discussed how this guidance could be useful for professionals looking to manage pain in certain settings as it was unlikely to interact with surgical or hormonal treatments, which would be the main alternative pharmacological management strategies. Therefore a neuromodulator for pain management in addition to first line treatment might help reduce pain further. The Committee was made aware that because of the well-established value of neuromodulators in pain management the evidence for these treatments for endometriosis specifically was almost entirely lacking and consequently an expert consensus was reached that there was no feature of endometriosis that would specifically indicate that neuromodulators would behave systematically differently in endometriosis than other long-term conditions, and therefore that the findings of CG173 would be appropriate to rely on. The Committee discussed the risks of extrapolating the CG79 guidance which focuses on neuropathic pain. Endometriosis could be considered to have similar pathophysiological processes via central sensitisation to neuropathic pain conditions but the CG79 guidance which may mean that it may be questionable whether it is directly translatable.

ISBN 978-1-4731-2661-9

Even though the trials indicated that there might be benefits of the pertubation method for the administration of local anaesthesia, the Committee considered the invasive nature of this. They agreed that this is a procedure that is not currently used in the NHS and that the evidence is not convincing to warrant a change in practice. The Committee raised concerns that the discomfort and possible side effects from the intervention would outweigh the possible benefits.

The Committee was of the opinion that the nature of this treatment make it unlikely to be adopted because it would require repeated monthly administrations (to co-occur with the menstrual cycle).

11.1.2.6.3 Consideration of economic benefits and harms

Based on NICE guidance CG173, both amitriptyline and gabapentin perform well relative to an ICER of £20,000/QALY and are inexpensive enough that a diagnostic strategy of 'empirical diagnosis' – treating all those with symptoms of endometriosis without a confirmatory diagnostic test - can be pursued. However, this is only with reference to the class of neuromodulators. The Committee discussed comparative economic considerations indicating that neuromodulators are neither inexpensive enough to be considered in preference to hormonal treatment nor effective enough to be considered in preference to surgery. There are also some women who cannot tolerate or do not want to take hormonal therapy (usually because they are seeking to conceive, at which time neuromodulators would not be the appropriate option). In other cases where a woman cannot have hormonal therapy, does not consider pregnancy but responds to neuromodulators, then it is unlikely surgery will be cost-effective for this woman.

In the very specific case of a woman who cannot have hormonal therapy, is not considering pregnancy and yet responds to neuromodulators, then the economic model indicates that neuromodulators should be trialled as a first line treatment (before considering surgery). It is difficult to imagine the personal circumstances of such a woman, and so it may be that in most cases where neuromodulators are recommended by the economic model as a first line treatment that the economic model does not accurately capture these specific circumstances.

As the Committee is only recommending neuromodulators in line with the NICE Guideline on the topic, there will be no significant resource impact.

11.1.2.6.4 Quality of evidence

The evidence was of very low to moderate quality, according to GRADE criteria. Even though the methodology of the trials was well described, there were inconsistencies in the results reported (for instance, differences in results when pain was reported as a categorical or continuous measure). There were also a number of outcomes that were only reported as medians, for which it is difficult to estimate the confidence in the effect size.

The Committee therefore had little confidence in the findings of the trials.

11.1.2.6.5 Other considerations

The Committee noted that there is a substantial amount of evidence for nerve ablation, specifically in the form of Laparoscopic uterine nerve ablation (LUNA). However LUNA has been covered by a NICE Interventional Procedure Guideline (IPG234) and so was outside the scope of this Guideline. The IPG concluded that the evidence on laparoscopic uterine nerve ablation for chronic pelvic pain suggests that it is not efficacious and therefore should not be used.

ISBN 978-1-4731-2661-9

11.1.2.6.6 Key conclusions

The Committee concluded that there was currently insufficient evidence for the effectiveness of neuromodulators in managing pain of women with endometriosis. Little confidence was placed in the evidence for a method of administering local anaesthetics, which is not currently used in the NHS and hence the Committee decided not to make any recommendation regarding this technique. They agreed that the recommendations set out in NICE guidance CG173 would be generalizable to women with endometriosis and therefore cross-referenced to this guidance.

11.1.2.7 Recommendations

35. For recommendations on using neuromodulators to treat neuropathic pain, see the NICE guideline on neuropathic pain.

11.1.3 Hormonal medical treatments

Review question: What is the effectiveness of hormonal medical treatments for treating endometriosis compared to placebo, other hormonal medical treatments, usual care, surgery, or surgery in combination with hormonal treatment?

11.1.3.1 Introduction

Endometriosis is considered a predominantly oestrogen-dependent condition. Thus, ovarian suppression with hormones is currently offered as an alternative to surgical excision to treat the disease and its symptoms. However, clinical practice with regards to hormonal treatment varies widely, because of the implications of each option. None of the hormones used to manage endometriosis (or, in fact, any drug) are free of side effects, but the severity and tolerability of the side effects can vary quite significantly. Many of the hormones used to manage endometriosis-associated pain will also reduce menstrual bleeding and this may be advantageous. Similarly, the contraceptive properties of the hormones may be welcome if the woman does not wish to become pregnant at this moment in time, or unwanted if fertility is an issue. All these factors should be taken into consideration when prescribing hormones to women for the treatment of endometriosis. The effects of hormonal contraceptives, progestogens, anti-progestogens, gonadotrophin releasing hormone agonists (GnRH agonists) and aromatase inhibitors on endometriosis symptoms are discussed below.

The principal aim of this review is to determine the clinical and cost effectiveness of hormonal medical treatments in reducing pain in women with endometriosis.

For full details, see the review protocols in Appendix D.

11.1.3.2 Network Meta-analysis

11.1.3.2.1 Methods

The results of conventional pairwise comparison (and meta-analyses) of direct evidence alone do not help to fully inform which intervention is most effective in the treatment of endometriosis. The challenge of interpretation arises for 2 main reasons:

- In isolation, each pairwise comparison does not fully inform the choice between the different treatments and having a series of discrete pairwise comparisons can be disjointed and difficult to interpret.
- RCT evidence is not available that directly compares treatments of clinical interest are not fully available, for example, comparison between certain types of hormonal therapy. This makes choice difficult unless based on patient preference or cost.

ISBN 978-1-4731-2661-9

To overcome these issues, a hierarchical Bayesian network meta-analysis (NMA) was performed in addition to a pairwise comparison of hormonal treatments. Advantages of performing this type of analysis are:

- It allows the synthesis of data from direct and indirect comparisons without breaking randomisation, to produce measures of treatment effect and ranking of different interventions. If treatment A has never been compared against treatment B head to head, but these 2 interventions have been compared to a common comparator directly, then an indirect treatment comparison can use the relative effects of the 2 treatments versus the common comparator. Indirect estimates can be calculated whenever there is a path linking 2 treatments through a set of common comparators. All the randomised evidence is considered simultaneously within the same model.
- For every intervention in a connected network, a relative effect estimate (with its 95% credible intervals) can be estimated versus any other intervention. These estimates provide a useful clinical summary of the results and facilitate the formation of recommendations based on all of the best available evidence, while appropriately accounting for uncertainty.
- Estimates from the NMA can be used to directly parameterise treatment effectiveness in cost-effectiveness modelling of multiple treatments.

The terms indirect treatment comparisons, mixed treatment comparisons and network meta-analysis are used interchangeably, though we use the term NMA throughout the guideline.

Study selection and data collection

For full details, see review and analysis protocols in Appendices K and L.

11.1.3.2.2 Outcome measures for NMA

For assessing the effectiveness of treatments, the Committee identified pain relief, health-related quality of life (QoL) and adverse events as critical outcomes for which NMA could be used to aid decision-making. NMAs were performed on these outcomes where evidence was available.

Pain relief

For pain relief, the visual analogue scale (VAS) was considered by the Committee to be the most widely used useful pain scale for which data would be available. A series of subscales first reported by Biberoglu and Behrman (1981) were also frequently used in studies of hormonal treatments and NMAs of these subscales were also performed to provide additional information on pain relief. There was sufficient evidence available for NMA for dysmenorrhoea, dyspareunia and non-menstrual pelvic pain subscales, though not for induration and pelvic tenderness subscales. Therefore induration and pelvic tenderness were analysed within a separate pairwise comparison analysis. Dysmenorrhoea and non-menstrual pelvic pain were used in a multivariate analysis to inform the VAS scale, so their results are not presented separately here.

Health-related QoL

For health-related QoL, the Short Form 36 Health Survey (SF-36) was determined by the Committee to the most useful scale that was widely used in the literature. However, there were not a sufficient number of studies available from the systematic review to allow for NMA. Therefore these studies were analysed within the separate pairwise comparison analysis where appropriate.

ISBN 978-1-4731-2661-9

Adverse events

As adverse events varied substantially depending on the treatment in question, the Committee felt that the number of women discontinuing treatments due to adverse events was a more generalizable and useful outcome, as this also accounted for how severe women felt an adverse event to be (i.e. it had to be sufficiently severe for them to discontinue treatment).

11.1.3.2.3 Statistical methodology

Due to difficulty in obtaining stable estimates from the model, NMAs were conducted separately for hormonal and non-pharmacological therapies, and for surgery and surgery plus hormonal treatment. The Committee felt that the difficulties in model estimation were likely to be because the populations may not have been sufficiently homogeneous, as patients receiving surgical treatment were likely to have failed on hormonal treatments, thus violating the assumption of transitivity.

Data were available for a number of treatments and routes of administration. Due to the sparseness of the networks, it was necessary to group treatments within different classes and assume a common class effect (Table 62). The common class effects were assessed to identify if it was reasonable to assume similarity of treatment effects within classes. Though data were often too limited to be able to closely examine within-class variation there was no evidence to suggest that treatment effects differed substantially within classes. Multi-level NMA models with treatments nested within classes were also examined, though this added complexity did not improve model fit for any of the analyses. Therefore common class effects were assumed throughout the analyses.

There are 3 key assumptions behind an NMA: similarity, transitivity and consistency.

Similarity across trials is the critical rationale for the consistency assumption to be valid as, by ensuring the clinical characteristics of the trials are similar, we ensure consistency in the data analysis.

More specifically, randomisation holds only within individual trials, not across the trials. Therefore, if the trials differ in terms of patient characteristics, measurement and/or definition of outcome, length of follow-up across the direct comparisons, the similarity assumption is violated and this can bias the analysis. Potential sources of heterogeneity arising from trials of interventions for endometriosis and attempts made to identify and account for heterogeneity are:

- Different population: for example, mixed populations of women with and without endometrioma.
    - Sensitivity analyses were performed to test the validity of the assumption of similarity of effect for treatments for women with and without endometrioma.
- Different duration of treatment or study follow-up:
    - Although data were limited to reliably assess the effect of study duration, relative treatment effects appeared to be similar across studies of different duration that fitted the inclusion criteria specified in the analysis protocol.
    - Sensitivity analyses were conducted to assess the impact of removing studies of short duration.
- Different dosages of pharmacological treatments:
    - These typically showed little variation and were within the dose ranges specified by the British National Formulary (BNF).

Transitivity is the assumption that an intervention (A) will have the same efficacy in a study comparing A vs. B as it will in a study comparing A vs. C. Another way of looking at it, in terms of the study participants, is that we assume that it is equally likely that any patient in

ISBN 978-1-4731-2661-9

the network could have been given any of the treatments in the network and would have responded to the treatments in the same way (depending on how efficacious the treatments are).

This assumption is closely related to similarity in that if participants in a study comparing A vs. B are not the same as those in a study comparing A vs. C. For example, if those in a comparison of A vs. B were women seeking treatment to improve fertility and those in A vs. C were women whose primary concern was pain relief, then both the similarity and transitivity assumptions would be violated, hence the importance in our analysis of keeping these populations distinct.

The final assumption is consistency/coherence of the network. It is important that for a network that contains closed loops of treatments (e.g. with studies comparing A vs. B, B vs. C and A vs. C), the indirect comparisons are consistent with the direct comparisons. Discrepancies between direct and indirect estimates of effect may result from several possible causes. One possible cause is 'chance' and if this is the case then the NMA results are likely to be more precise as they pool together more data than conventional meta-analysis estimates alone. However, a second possible cause could be due to differences between the trials included in terms of their clinical or methodological characteristics, which would therefore raise concerns about the validity of the network.

Table 62: Dose ranges of treatments in different classes of interventions, with abbreviations used in tables and figures within this chapter

Class	Treatment	Abbreviation
Placebo/no treatment	Placebo No treatment/waiting list	-
Danazol/gestrinone	Danazol (100-800 mg/d) Gestrinone	Dan/gest
Oestrogens (oral)	Oestradiol (1–2 mg/d) Conjugated equine oestrogens (0.3–1.25 mg/d)	Oest(o)
Progestogens (oral)	Norethisterone (2.5 mg/d) Medroxyprogesterone (15–30 mg/d) Levonorgestrel (30 micrograms/d) Desogestrel (75 micrograms/d) Dienogest (2 mg/d)	Prog(o)
Progestogens (depot)	Medroxyprogesterone (150 mg/3 months) Gestodene (5–10 mg)	Prog(i.m.)
Progestogens (subcutaneous)	Medroxyprogesterone (104 mg/3 months) Promegestone	Prog(s.c.)
Progestogens (intrauterine)	Levonorgestrel (20 micrograms/day)	Prog(i.u.)
GnRH agonists (depot)	Leuprorelide (3.75 mg/m) Triptorelin (3 mg/m)	GnRHa(i.m.)
GnRH agonists (subcutaneous)	Goserelin (3.6 mg/m)	GnRHa(s.c.)
GnRH agonists (nasal spray)	Nafarelin (200 micrograms b.d.) Buserelin (300 micrograms t.d.)	GnRHa(i.n.)
GnRH antagonists	Elagolix	GnRHant
Aromatase inhibitors	Anastrozole (1 mg/d) Letrozole (2.5 mg/d)	AromaInhib
Anti-androgens	Cyproterone acetate (only in combination as combined oral contraceptive)	Anti-And
Selective oestrogen receptor modulators	Raloxifene (60 mg/d)	SERM

ISBN 978-1-4731-2661-9

Class	Treatment	Abbreviation
Tibolone	Tibolone (2.5 mg/d)	-
Nutritional supplements	Calcium Vitamin D	Supp
Chinese herbal medicine	Nei yi pills Dan'e mixture	CHM
Dietary interventions	Dietary intervention	Diet

(b) Table only includes treatments in full-text studies assessed for inclusion/exclusion. Treatments only in studies that were not included in the NMA could not be included in the network.

There were no studies that fitted the NMA inclusion criteria for the following treatments in Table 62: anti-androgens, selective oestrogen receptor modulators, tibolone, nutritional supplements, Chinese herbal medicine, and dietary interventions. As no studies investigating non-pharmacological treatments fitted the inclusion criteria for the NMA, the analyses presented are only of hormonal treatments.

11.1.3.2.4 Summary of included studies

Studies included in the NMA

All studies included women with laparoscopic confirmation of endometriosis.

Table 63: Characteristics of included studies

First author	Pub date	rAFS	Surgery type	Endome-triomas included	Risk of bias[a]	Outcomes reported in study (1=reported, 0=not reported)				
						Disc	VAS	Dysmen	Dyspar	Pelv pain
Acs	2015	NR	None	NR	Mod	1	0	0	0	0
Agarwal	1997	I–II	None	NR	Low	1	0	1	1	1
Bergqvist	1997	I–II	None	NR	Low	1	0	0	0	0
Bergqvist	1998	I–II	None	NR	Mod	1	0	0	0	0
Bergqvist	2000	I-IV	None	NR	Mod	1	0	0	0	0
Burry	1989	I–IV	None	None	Mod	1	0	0	0	0
Burry	1992	NR	None	NR	Mod	1	0	0	0	0
Carr	2014	I–IV	None	NR	Low	1	0	0	0	0
Crosig-nani	2006	NR	None	NR	Mod	1	0	0	0	0
Diamond	2014	I–IV	None	NR	Low	1	0	0	0	0
Dlugi	1990	I–IV	None	Some	Mod	0	0	1	1	1
Dmowski	1989	I–IV	None	NR	High	0	0	1	1	1
Fedele	1989	I–IV	None	None	Mod	1	0	0	0	0
Fernandez	2004	III–IV	None	NR	Low	0	0	0	0	1
Ferreira	2010	NR	None	None	Mod	0	1	0	0	0
Franke	2000	III–IV	None	NR	Mod	1	0	0	0	0
GISG (Verce-llini)	1996	I–II	None	NR	Low	0	0	1	1	1
Gomes	2007	III–IV	None	NR	Mod	0	1	0	0	0
Granese	2015	III–IV	Exci-sion/ ablation	Some	High	0	1	0	0	0

ISBN 978-1-4731-2661-9

First author	Pub date	rAFS	Surgery type	Endome-triomas included	Risk of bias[a]	Outcomes reported in study (1=reported, 0=not reported)				
						Disc	VAS	Dysmen	Dyspar	Pelv pain
Guzick	2011	NR	None	NR	Mod	0	1	0	0	0
Harada	2008	NR	None	All	Low	1	1	1	0	1
Harada	2009	NR	None	Some	Low	1	0	0	0	0
Henzl	1989	I–IV	None	Some	Mod	1	0	0	0	0
Horn-stein	1998	I–II	None	NR	Low	1	0	1	0	1
Jelley	1988	I–IV	None	NR	Mod	1	0	0	0	0
Kennedy	1990	I–II	None	NR	Mod	1	0	0	0	0
Kiesel	1996	NR	None	NR	Low	1	0	0	0	0
Kiilholma	1995	III–IV	None	NR	Mod	1	0	0	0	0
Ling	1999	NR	None	NR	Mod	0	0	1	1	1
NEET	1992	I–IV	None	NR	Mod	1	0	0	0	0
Petta	2005	III–IV	None	NR	Low	0	1	0	0	0
Razzi	2007	NR	Exci-sion	All	Mod	0	1	0	0	0
Rock	1993	I–IV	None	NR	High	1	0	0	0	0
Rolland	1990	I–IV	None	NR	Mod	1	0	0	0	0
Rotondi	2002	I–IV	None	NR	Mod	1	0	0	0	0
Schlaff	2006	NR	None	NR	Low	1	0	0	0	0
Seibel	1982	NR	None	None	High	1	0	0	0	0
Shaw	1990	NR	None	NR	Mod	1	0	0	0	0
Shaw	1992	I–IV	None	Some	Mod	1	0	0	0	0
Strowitz-ki	2010	III–IV	None	NR	Mod	1	1	0	0	0
Strowitz-ki	2010 b	I–IV	None	NR	Mod	1	1	0	0	0
Sutton	1994	I–II	Ablation	None	High	0	1	0	0	0
Telimaa	1987	I–II	None	Some	High	1	0	0	0	0
Vercellini	1996	I–IV	None	NR	Low	1	0	1	0	0
Walch	2009	I–IV	None	None	High	1	0	0	0	0
Wheeler	1993	I–IV	None	NR	Mod	1	0	0	0	0
Wong	2010	III–IV	None	NR	High	1	0	0	0	0
Zhu	2014	I–II	Exci-sion/ ablation	None	Mod	0	1	0	0	0
ZOLA-DEX	1996	I–II	None	NR	Mod	1	0	0	0	0

(a) Cochrane risk of bias checklist
(b) Abbreviations - rAFS: revised American Fertility Scale; Mod: moderate; NR: not reported; Disc: discontinuation of treatment due to adverse events; Dysmen: dysmenorrhoea; Dyspar: dyspareunia; Pelv Pain: non-menstrual pelvic pain.

11.1.3.2.5 Studies excluded from the NMA

Table 64 lists the studies that were excluded from the NMA for statistical reasons.

ISBN 978-1-4731-2661-9

Table 64: Table of studies excluded from the NMA for statistical reasons

First author	Pub date	Reason for exclusion
Fernandez	2004	Study adds no information to any network
Ferrero	2011	Treatment not connected to any network
Howell	1995	Study adds no information to any network
Soysal	2004	Treatment not connected to any network

11.1.3.2.6 Clinical evidence profile

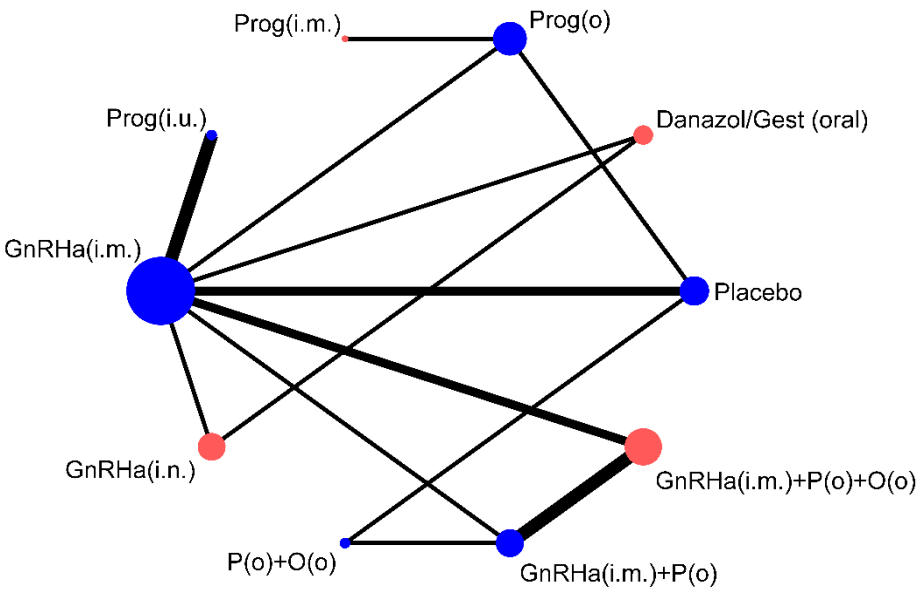
Pain relief – VAS

Due to difficulty in achieving convergence during estimation, NMAs were conducted separately for hormonal and non-pharmacological therapies, and for surgery and surgery plus hormonal treatment. The Committee felt that this was likely to be because the populations may not have been sufficiently homogeneous, as patients receiving surgical treatment were likely to have failed on hormonal treatments, thus violating the assumption of transitivity.

Hormonal treatments

Fifteen trials of 10 hormonal treatment classes were included in the network for the outcome of pain relief on the VAS, with a total sample size of 1,680 women (Figure 13). No studies reported data on non-pharmacological treatment that could be used in the network. One study was at high risk of bias, 7 were at moderate risk of bias and 7 at low risk of bias.

Figure 13: Network for hormonal therapy for pain relief (VAS)



The size of nodes is proportional to the number of women in the network who were given a particular treatment class. The thickness of connecting lines is proportional to the number of studies directly comparing 2 treatment classes. Red nodes indicate treatment classes that are informed only from Biberoglu and Behrman scales. For treatment name abbreviations, see Table 62.

Table 65 presents the results of the pairwise meta-analyses of the VAS where they were available (direct comparisons; upper right section of table) together with the results from the multivariate NMA for every possible class comparison (lower left section of table), presented

ISBN 978-1-4731-2661-9

as mean differences. A multivariate NMA was performed as this allowed for the incorporation of additional information from dysmenorrhoea and non-menstrual pelvic pain Biberoglu and Behrman subscales, allowing estimation of the efficacy of treatments not investigated using the VAS (progestogens (i.m.), danazol/gestrinone, GnRHa (i.n.) and GnRHa (i.m.) plus the pill). The VAS is a 0–100 patient-reported scale, on which a difference of 10 points has been shown to be clinically significant to patients (Gerlinger 2012).

NMA results were derived from a fixed effects multivariate model. Figure 14 graphically presents the results computed by the NMA for each treatment versus placebo.

All treatments led to a clinically significant reduction in pain on the VAS when compared to placebo. The magnitude of this treatment effect was similar for all treatments, suggesting that there was little difference between them in their capacity to reduce pain. No other significant differences were found between the hormonal treatments.

The levornorgestrel implant (progestogens (i.u.)) had the highest probability of being among the best 3 treatments (74.2%), followed by danazol/gestrinone (52.6%) and GnRHa (i.m.) plus the pill (52.5%). The results of this are described in Table 66.

Results were broadly similar from the multivariate and univariate NMA where information was available for comparison. The largest differences were for the progestogens (i.u.) (considerably more effective in the multivariate than in the univariate NMA) and GnRHa (i.m) (less effective in the multivariate than in the univariate NMA) (Appendix L).

Sufficient data to calculate standard errors (SEs) was not available in 4 of the 15 trials. However, sensitivity analyses using the upper 95% CrI of the posterior for the imputed SEs showed that estimates and their 95% CrIs were very insensitive to the imputed SEs (Appendix L).

The multivariate nature of the network did not allow for simple assessment of incoherence, though it was assessed for each of the univariate outcomes and was not found to be present in any closed loops. However there were some differences between the direct estimates on the VAS scale and those from the NMA, particularly for progestogens (oral) versus GnRHa (i.m.). These differences are due to the multivariate analysis and the inclusion of evidence from the Biberoglu and Behrman scales and therefore reflect incoherence between the outcomes rather than between the treatment comparisons. Although this appears to change the direction of effect in some comparisons, the change is very small and not clinically meaningful.

ISBN 978-1-4731-2661-9

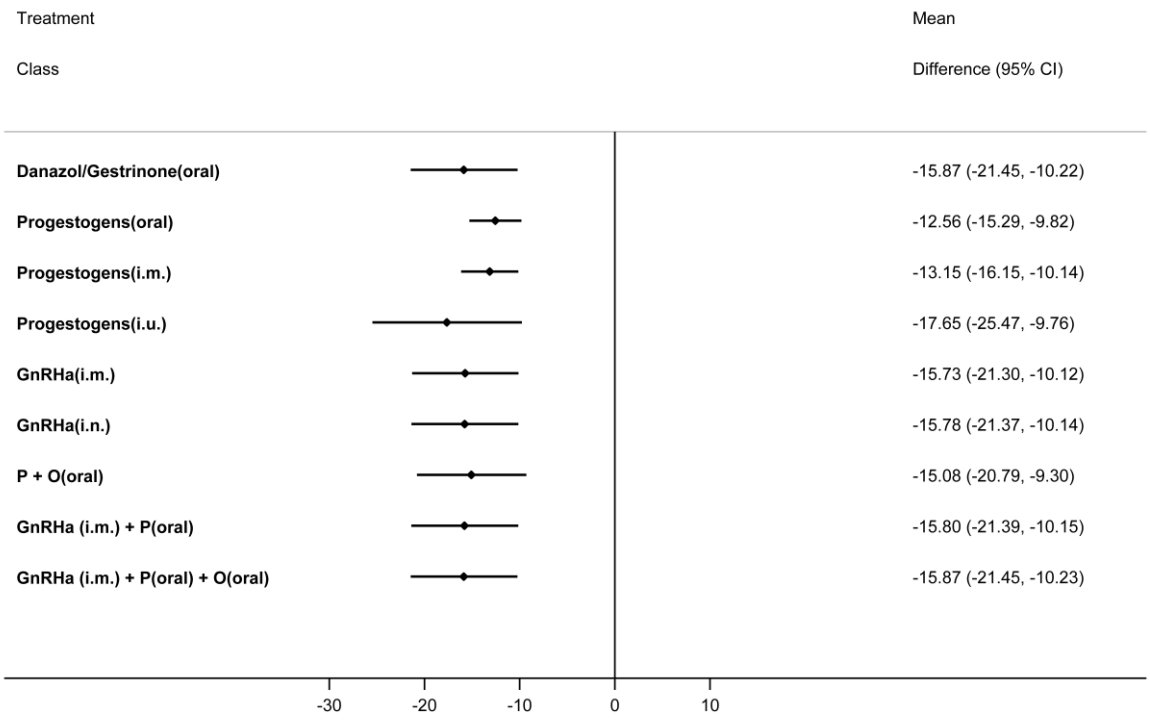
Table 65: Matrix of results for the multivariate NMA of hormonal therapy for pain relief on the VAS

Placebo/no treatment		**-12.3 (-12.7 to -11.9)**					**-18.6 (-20.4 to -16.8)**		
-15.9 (-21.5 to -10.2)	Danazol/ gestrinone (oral)								
-12.6 (-15.3 to -9.8)	3.3 (-2.1 to - 8.7)	Progesto-gens (oral)			1.5 (0.7 to 2.3)				
-13.2 (-16.2 to -10.1)	2.7 (-2.8 to 8.2)	-0.6 (-1.8 to 0.6)	Progesto-gens (i.m.)						
-17.7 (-25.5 to -9.8)	-1.8 (-7.2 to 3.6)	-5.1 (-12.8 to 2.7)	-4.5 (-12.4 to 3.4)	Progesto-gens (i.u.)	-1.4 (-2.8 to 0.1)				
-15.7 (-21.3 to -10.1)	0.1 (-0.5 to 0.8)	-3.2 (-8.5 to 2.2)	-2.6 (-8.1 to 2.9)	1.9 (-3.4 to 7.3)	GnRHa (i.m.)				
-15.8 (-21.4 to -10.1)	0.1 (-0.6 to 0.8)	-3.2 (-8.6 to 2.2)	-2.6 (-8.2 to 2.9)	1.8 (-3.5 to 7.3)	0.0 (-0.7 to 0.6)	GnRHa (i.n.)			
-15.1 (-20.8 to -9.3)	0.8 (-0.1 to 1.6)	-2.5 (-8.0 to 3.0)	-1.9 (-7.6 to 3.7)	2.5 (-2.8 to 8.1)	0.7 (-0.2 to 1.5)	0.7 (-0.2 to 1.5)	Prog(oral)+ Oest(oral)	3.5 (-1.7 to 8.7)	
-15.8 (-21.4 to -10.2)	0.1 (-0.7 to 0.8)	-3.3 (-8.6 to 2.2)	-2.7 (-8.2 to 2.9)	1.8 (-3.5 to 7.3)	-0.1 (-0.8 to 0.6)	0 (-0.8 to 0.7)	-0.7 (-1.6 to 0.2)	GnRHa(i.m.) +Prog(oral)	
-15.9 (-21.5 to -10.2)	0.0 (-0.7 to 0.7)	-3.3 (-8.7 to 2.1)	-2.7 (-8.3 to 2.8)	1.8 (-3.6 to 7.2)	-0.1 (-0.8 to 0.5)	-0.1 (-0.8 to 0.6)	-0.8 (-1.7 to 0.1)	-0.1 (-0.8 to 0.6)	GnRHa(i.m.) +Prog(oral)+ Oest(oral)

Mean differences and 95% CrIs from the multivariate NMA (bottom left diagonal) and conventional pairwise VAS meta-analyses (top right diagonal) treatment effects between the column-defined and row-defined treatments. Mean differences less than 0 favour the row-defined treatment. Numbers in bold, grey-shaded cells denote results where the 95% CrI do not include 0. Treatment effects for danazol/gestrinone (oral), progestogens (i.m.), GnRHa (i.n.), progestogen plus oestrogen (oral) and GnRHa (i.m.) plus progestogen (oral) plus oestrogen (oral) are informed from Biberoglu subscales for dysmenorrhoea and non-menstrual pelvic pain. Pairwise results for these treatments are therefore not shown here as they would be reported on different scales. For treatment name abbreviations, see Table 62.

ISBN 978-1-4731-2661-9

Figure 14: Forest plot showing mean differences (95% CrI) of multivariate NMA estimates for each treatment versus placebo/no treatment for pain relief on the VAS



For treatment name abbreviations, see Table 61

Table 66: Mean differences versus placebo from multivariate and univariate NMAs for each treatment, with probabilities of being among the best 3 treatments and the worst 3 treatments, and the rank and 95% CrI from the multivariate NMA for each treatment

Treatment class	Probability of being within the best 3 (%)	Probability of being within the worst 3 (%)	Rank (95% CrI)
Placebo/no treatment	0.00%	100.00%	10 (10 to 0)
Danazol/gestrinone (o)	57.33%	0.47%	3 (1 to 7)
Progestogens (o)	0.05%	95.70%	9 (7 to 9)
Progestogens (i.m.)	14.27%	31.98%	7 (1 to 9)
Progestogens (i.u.)	74.46%	16.01%	1 (1 to 9)
GnRHa (i.m.)	22.14%	0.80%	5 (2 to 7)
GnRHa (i.n.)	34.64%	0.87%	4 (1 to 7)
Prog (o) + Oest (o)	0.57%	52.83%	8 (6 to 9)
GnRHa (i.m.)+Prog (o)	39.65%	0.87%	4 (1 to 7)
GnRHa (i.m.) + Prog (o) + Oest (o)	56.88%	0.47%	3 (1 to 7)

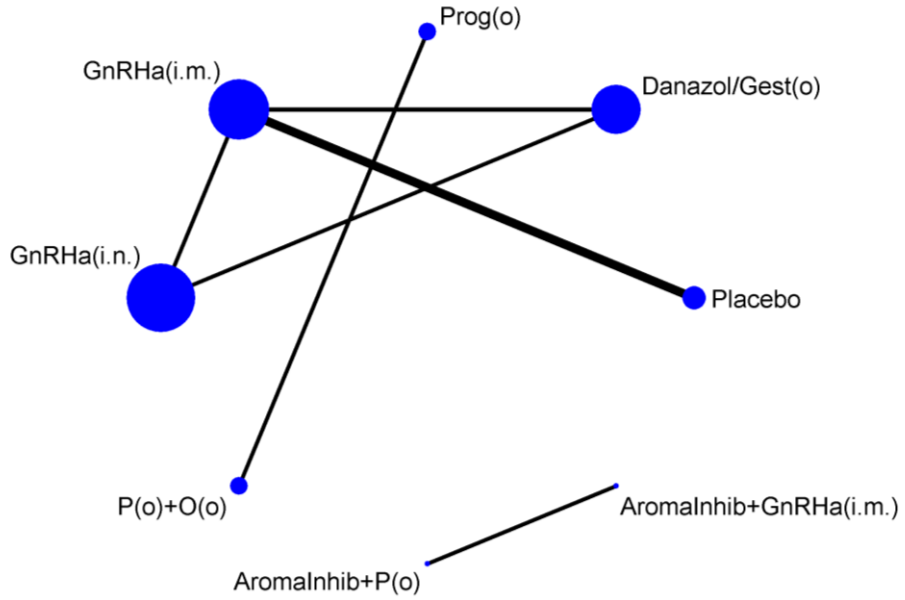
(c) For treatment name abbreviations, see Table 62

ISBN 978-1-4731-2661-9

Pain relief – dyspareunia (Biberoglu and Behrman)

Five trials of 4 treatment classes were included in the network for the outcome of dyspareunia, with a total sample size of 572 women (Figure 15). One study was at high risk of bias, 2 were at moderate risk of bias and 2 were at low risk of bias.

Figure 15: Network for treatments for relief of dyspareunia



The size of nodes is proportional to the number of women in the network who were given a particular treatment class. The thickness of connecting lines is proportional to the number of studies directly comparing 2 treatment classes. Four treatment classes were not connected and could not be compared in the NMA (progestogens (oral), progestogen + oestrogen (oral), aromatase inhibitor + progestogen (oral), aromatase inhibitor + GnRHa (i.m.)). For treatment name abbreviations, see Table 62.

Table 67 presents the results of the conventional pairwise meta-analyses (direct comparisons; upper right section of table) together with the results from the NMA for every possible class comparison (lower left section of table), presented as mean differences. Dyspareunia was assessed using a 0–3 patient-reported scale developed by Biberoglu and Behrman (1981). NMA results were derived from a fixed effects model.

All treatments were significantly better at relieving dyspareunia than placebo/no treatment, although the improvement was quite small. GnRHa (i.n.) was also found to be significantly better at relieving dyspareunia than GnRHa (i.m.), which led to it having the highest probability of being the best treatment (85.1%), followed by danazol/gestrinone (14.3%) (see Table 68). Results from this NMA should be interpreted with caution, as sufficient data to calculate SEs was only available in 2 of the 5 trials. Sensitivity analyses using the upper 95% CrI of the posterior for the imputed SEs showed that the probability of being the best treatment results were sensitive to the imputed SEs. With larger SEs, there was more uncertainty regarding whether GnRHa (i.n.) or danazol/gestrinone were the better treatment (Appendix L).

There was no clear evidence of incoherence in the closed loop of GnRHa (i.m.), danazol/gestrinone and GnRHa (i.n.). However, there was very limited statistical power to test for this and, as the direction of effect differs between 2 of the direct and indirect estimates, results of this network should be treated with caution.

- GnRHa (i.m.) vs. danazol/gestrinone (p=0.123)
  - direct MD=0.33 (95% CrI: 0.04 to 0.65)

ISBN 978-1-4731-2661-9

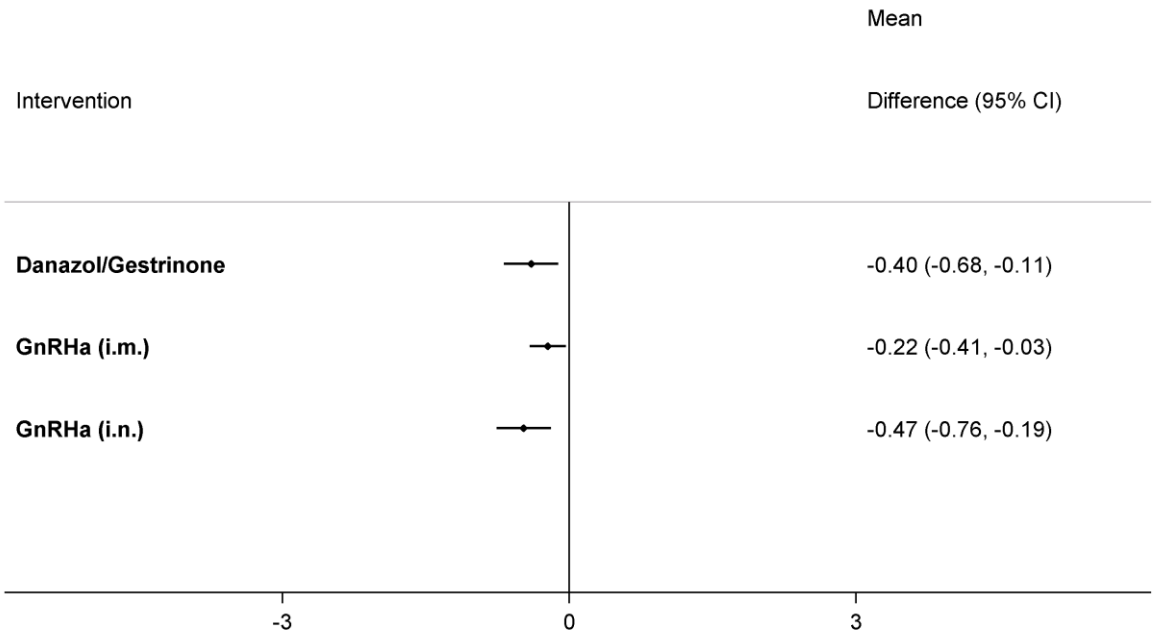
- indirect MD=-0.01 (95% CrI: -0.33 to 0.31)
- GnRHa (i.n.) vs. danazol/gestrinone (p=0.115)
  - direct MD=-0.12 (95% CrI: -0.27 to 0.03)
  - indirect MD=0.22 (95% CrI: -0.17 to 0.62)
- GnRHa (i.n.) vs. GnRHa (i.m.) (p=0.115)
  - direct MD=-0.11 (95% CrI: -0.38 to 0.17)
  - indirect MD=-0.45 (95% CrI: -0.77 to-0.13)

Table 67: Matrix of results for the NMA of dyspareunia

Placebo/no treatment		**-0.22 (-0.41 to -0.03)**	
-0.4 (-0.68 to -0.11)	Danazol/gestrinone	0.32 (0.04 to 0.61)	-0.12 (-0.27 to 0.03)
-0.22 (-0.41 to -0.03)	0.18 (-0.04 to 0.39)	GnRHa (i.m.)	-0.11 (-0.39 to 0.17)
-0.47 (-0.76 to -0.19)	-0.08 (-0.22 to 0.06)	-0.25 (-0.46 to -0.04)	GnRHa (i.n.)

Mean differences and 95% CrIs from the NMA (bottom left diagonal) and conventional meta-analyses (top right diagonal) treatment effects between the column-defined and row-defined treatments. Mean differences less than 0 favour the row-defined treatment. Numbers in bold, grey-shaded cells denote results where the 95% CrI do not include 0. For treatment name abbreviations, see Table 62.

Figure 16: Forest plot showing mean differences (95% CrI) of NMA estimates for each treatment versus placebo for the relief of dyspareunia



For treatment name abbreviations, see Table 61

Table 68: Probabilities of being the best treatment and the rank (with 95% CrI) for each treatment

Treatment class	Probability of being the best treatment (%)	Rank (95% CrI)
Placebo/no treatment	0.03%	4 (4 to 4)
Danazol/gestrinone	14.26%	2 (1 to 3)
GnRHa (i.m.)	0.67%	3 (2 to 3)
GnRHa (i.n.)	85.05%	1 (1 to 2)

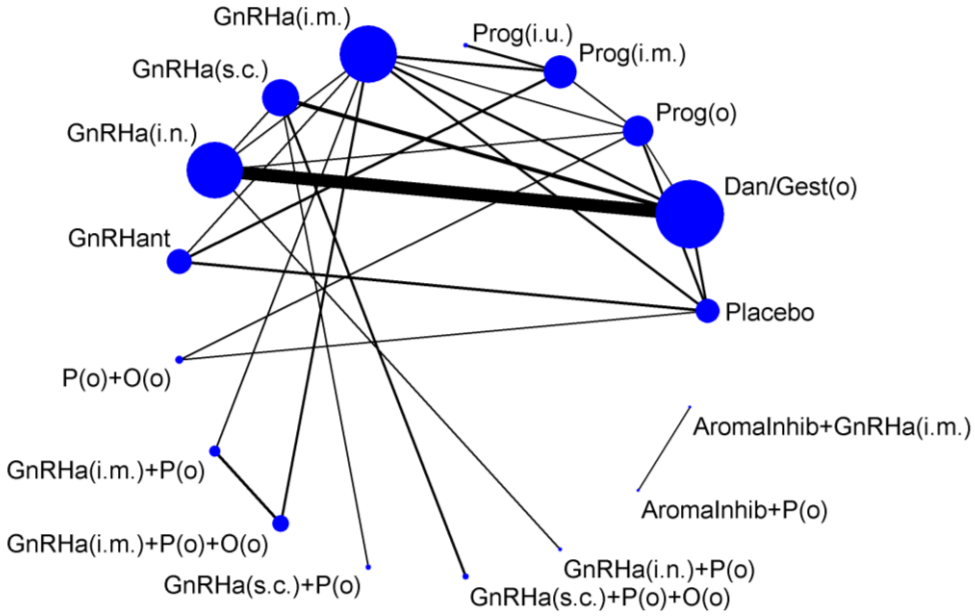
(d) For treatment name abbreviations, see Table 62

ISBN 978-1-4731-2661-9

Discontinuation of treatment due to adverse events

36 trials of 15 treatment classes were included in the network for the outcome of discontinuation of treatment due to adverse events, with a total sample size of 5,319 women (Figure 17). No studies that reported data on non-pharmacological treatments could be included in the network. Five studies were at high risk of bias, 21 studies were at moderate risk of bias and 10 studies were at low risk of bias.

Figure 17: Network for discontinuation of treatment due to adverse events



The size of nodes is proportional to the number of women in the network who were given a particular treatment class. The thickness of connecting lines is proportional to the number of studies directly comparing 2 treatment classes. Two treatment classes were not connected and could not be compared in the NMA (aromatase inhibitors + progestogens (oral) and aromatase inhibitors + GnRHa (i.m.)). For treatment name abbreviations, see Table 62.

Table 69 presents the results of the pairwise meta-analyses (direct comparisons; upper right section of table) together with the results from the NMA for every possible class comparison (lower left section of table), presented as odds ratios (ORs). These results were derived from a random effects model with very high heterogeneity (between-study SD: 0.94 (95% CrI: 0.45 to 1.69)). Accounting for severity of endometriosis (as measured by the rAFS) did not further explain the high heterogeneity.

Several treatment classes were found to result in significantly more discontinuations of treatment due to adverse events than placebo/no treatment (danazol/gestrinone, progestogens (oral), progestogens (i.m.), GnRHa (i.m.), GnRHa (i.n.) and GnRHa (i.n.) plus progestogen). The combined oral contraceptive pill (progestogen plus oestrogen (oral)), was found to lead to significantly less discontinuation than danazol/gestrinone, progestogen alone (oral), progestogen (i.m.), GnRHa (i.m.) and GnRHa (i.n.) plus progestogen. Figure 18 graphically presents the results computed by the NMA for each treatment versus placebo.

Though this outcome was taken where reported in studies as discontinuation due to adverse events, there may be some degree of reporting bias for this outcome - it is likely that women who are not finding the treatment effective or women who have difficulty with treatment compliance, may also be likely to discontinue treatment. For these women, even though they may cite adverse events as their reason for discontinuing treatment, treatment efficacy may play a part. Therefore this outcome is not independent of treatment efficacy. So because the combined oral contraceptive pill (progestogen plus oestrogen (oral)) was found to be effective, this may in part explain why it had the highest probability of being 1 of the best 3

ISBN 978-1-4731-2661-9

treatments for discontinuation due to adverse events (87.8%). Placebo/no treatment had the next highest probability (82.13%) (Table 70).

The treatments with the highest probability of being 1 of the 3 worst for discontinuation were GnRHa (i.n.) plus progestogen (oral) (78.8%), progestogen (i.m.) (39.1%), GnRHa (s.c.) plus progestogen (38.8%).

There was strong evidence of serious incoherence in the closed loop of GnRHa (s.c.), danazol/gestrinone and GnRHa (i.n.). As the direction of effect differs between direct and indirect estimates, results of this network should be treated with caution. No significant incoherence was found in any other closed loops of treatments (Appendix L).

- GnRHa (s.c.) vs. danazol/gestrinone (p=0.005)
  - direct OR = 0.10 (95% CrI: 0.03 to 0.25)
  - indirect OR = 2.25 (95% CrI: 0.41 to 12.18)
- GnRHa (i.n.) vs. danazol/gestrinone (p=0.025)
  - direct OR = 1.09 (95% CrI: 0.45 to 2.34)
  - indirect OR = 0.15 (95% CrI: 0.05 to 0.59)
- GnRHa (i.n.) vs. GnRHa (s.c.) (p=0.04)
  - direct OR = 0.42 (95% CrI: 0.09 to 1.88)
  - indirect OR = 9.03 (95% CrI: 3.00 to 33.12).

ISBN 978-1-4731-2661-9

Table 69: Matrix of results for the NMA of discontinuation of treatment due to adverse events

Placebo/ no treatment	**5.28 (0.28 to 305)**	**10.19 (0.96 to 371)**			**>999 (12.8 to >999)**			**>999 (12.2 to >999)**	**0.49 (0.07 to 3.18)**					
24.1 (2.14 to >999)	Danazol /gestrinone				0.47 (0.12 to 1.63)	0.11 (0.02 to 0.31)	1.12 (0.45 to 2.33)							
17.9 (1.76 to 676)	0.75 (0.16 to 3.27)	Progestogens (oral)	1.66 (0.15 to 21.3)		0.79 (0.12 to 4.95)		1.22 (0.21 to 7.26)							
26.8 (2.11 to 999)	1.1 (0.18 to 6.89)	1.47 (0.27 to 9.1)	Progestogens (i.m.)	0.4 (0.06 to 2.37)	0.88 (0.22 to 3.26)			0.34 (0.09 to 1.17)						
10.7 (0.35 to 811)	0.42 (0.02 to 7.11)	0.57 (0.03 to 9.32)	0.39 (0.04 to 3.23)	Progestogens (i.u.)										
17.4 (1.66 to 701)	0.73 (0.21 to 2.52)	0.98 (0.23 to 4.23)	0.66 (0.15 to 2.72)	1.71 (0.13 to 24.4)	GnRHa (i.m.)		0.48 (0.09 to 2.54)			1.04 (0.19 to 5.66)	0.8 (0.21 to 3.09)			
5.4 (0.34 to 251)	0.23 (0.05 to 0.73)	0.3 (0.04 to 1.87)	0.21 (0.02 to 1.59)	0.53 (0.02 to 10.9)	0.31 (0.05 to 1.57)	GnRHa (s.c.)	0.42 (0.08 to 2.18)					2.94 (0.07 to >999)	0.43 (0.01 to 6.83)	
14.98 (1.28 to 625)	0.63 (0.25 to 1.42)	0.84 (0.19 to 3.72)	0.57 (0.08 to 3.36)	1.48 (0.08 to 25.6)	0.86 (0.23 to 3.06)	2.76 (0.75 to 12.5)	GnRHa (i.n.)							10.31 (0.36 to >999)
11.57 (0.75 to 638)	0.47 (0.05 to 4.97)	0.63 (0.07 to 6.77)	0.43 (0.08 to 2.56)	1.11 (0.07 to 20.0)	0.65 (0.09 to 5.24)	2.08 (0.18 to 35.2)	0.75 (0.08 to 8.5)	GnRHant (oral)						
0.48 (0.04 to 5.53)	0.02 (<0.01 to 0.61)	0.03 (<0.01 to 0.77)	0.02 (<0.01 to 0.6)	0.04 (<0.01 to 2.92)	0.03 (<0.01 to 0.81)	0.09 (<0.01 to 3.57)	0.03 (<0.01 to 1.02)	0.04 (<0.01 to 1.56)	Prog(oral)+Oest(oral)					

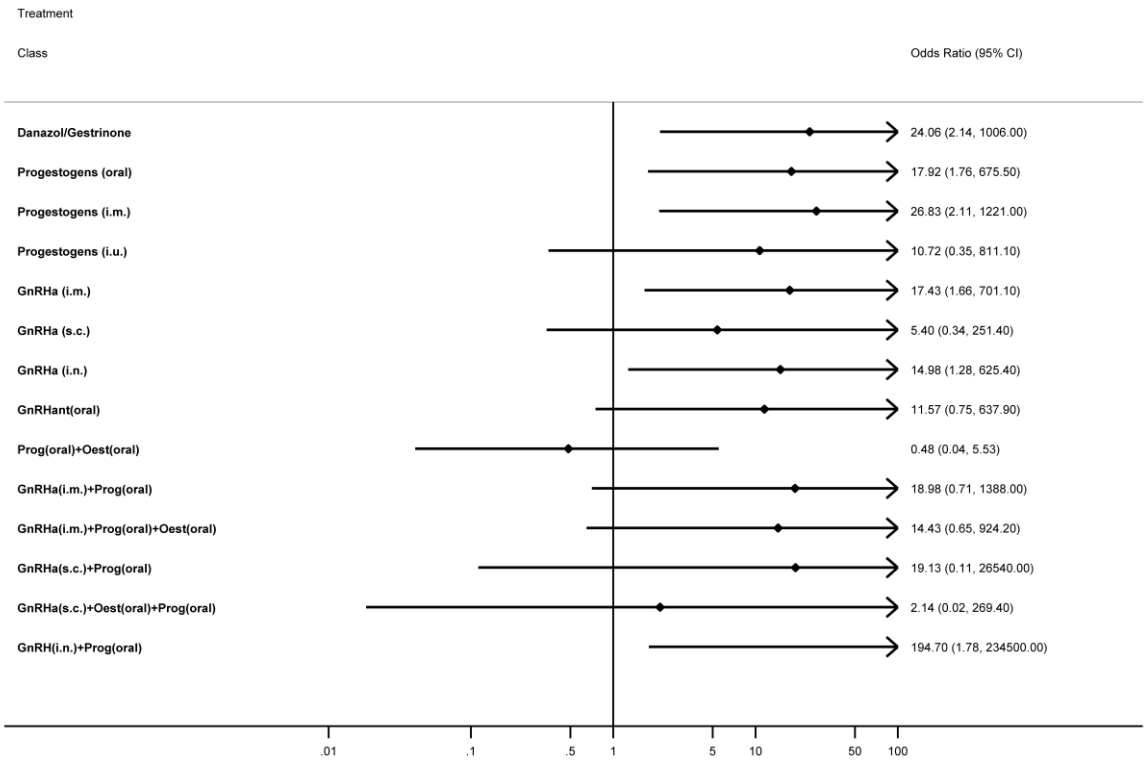
ISBN 978-1-4731-2661-9

18.98 (0.71 to 999)	0.76 (0.06 to 10.2)	1.02 (0.07 to 15.6)	0.69 (0.04 to 10.0)	1.8 (0.06 to 58.5)	1.04 (0.11 to 10.31)	3.35 (0.21 to 68.1)	1.21 (0.09 to 17.4)	1.62 (0.07 to 31.9)	39.98 (0.67 to >999)	GnRHa (i.m.)+Prog(oral)				
14.43 (0.65 to 924)	0.58 (0.05 to 6.06)	0.78 (0.07 to 9.35)	0.53 (0.04 to 6.08)	1.37 (0.05 to 36.9)	0.8 (0.11 to 5.93)	2.57 (0.2 to 41.8)	0.93 (0.09 to 10.39)	1.24 (0.07 to 19.6)	30.71 (0.59 to >999)	0.76 (0.1 to 5.67)	GnRHa (i.m.)+Prog(oral)+Oest(oral)			
19.13 (0.11 to >999)	0.66 (0.01 to 438)	0.92 (0.01 to 670)	0.62 (0.01 to 493)	1.69 (0.01 to >999)	0.93 (0.01 to 668)	2.97 (0.05 to >999)	1.07 (0.01 to 726)	1.47 (0.01 to >999)	40.65 (0.14 to >999)	0.93 (0.01 to 888)	1.2 (0.01 to >999)	GnRHa (s.c.)+Prog(oral)		
2.14 (0.02 to 269)	0.08 (<0.01 to 2.26)	0.11 (<0.01 to 4.05)	0.08 (<0.01 to 3.06)	0.2 (<0.01 to 14.6)	0.12 (<0.01 to 3.79)	0.38 (0.01 to 8.22)	0.14 (<0.01 to 3.8)	0.18 (<0.01 to 9.04)	4.42 (0.02 to 980)	0.11 (<0.01 to 6.75)	0.14 (<0.01 to 7.71)	0.11 (<0.01 to 21.4)	GnRHa (s.c.)+Oest(oral)+Prog(oral)	
196 (1.78 to >999)	6.73 (0.14 to >999)	9.18 (0.16 to >999)	6.28 (0.09 to >999)	17.12 (0.14 to >999)	9.32 (0.17 to >999)	30.95 (0.58 to >999)	10.74 (0.26 to >999)	14.87 (0.16 to >999)	**416 (2.05 to >999)**	9.37 (0.09 to >999)	12.19 (0.13 to >999)	10.51 (0.01 to >999)	94.31 (0.52 to >999)	GnRH(i.n.)+Prog(oral)

Odds ratios and 95% credible intervals (CrI) from the NMA (bottom left diagonal) and conventional meta-analyses (top right diagonal) treatment effects between the column-defined and row-defined treatments. Odds ratios less than 1 favour the row-defined treatment. Numbers in bold, grey-shaded cells denote results where the 95% CrI do not

ISBN 978-1-4731-2661-9

Figure 18: Forest plot showing odds ratios (95% CrI) of NMA estimates for each treatment versus placebo/no treatment for discontinuation due to adverse events



For treatment name abbreviations, see Table 62

Table 70: Probabilities of being among the best 3 treatments and the worst 3 treatments, and the rank and 95% Credible Interval (95%CrI) for each treatment

Treatment class	Probability of being within the best 3 (%)	Probability of being within the worst 3 (%)	Rank (95% CrI)
Placebo/no treatment	82.13%	0.05%	2 (1 to 6)
Danazol/gestrinone	0.01%	31.10%	11 (6 to 14)
Progestogens (o)	0.68%	17.46%	10 (4 to 14)
Progestogens (i.m.)	0.09%	39.11%	12 (6 to 15)
Progestogens (i.u.)	14.69%	15.91%	7 (2 to15)
GnRHa (i.m.)	0.14%	7.64%	10 (5 to 13)
GnRHa (s.c.)	16.33%	0.35%	5 (2 to 10)
GnRHa (i.n.)	0.42%	6.07%	9 (5 to 13)

ISBN 978-1-4731-2661-9

Treatment class	Probability of being within the best 3 (%)	Probability of being within the worst 3 (%)	Rank (95% CrI)
GnRHant (o)	7.43%	11.83%	7 (3 to 14)
Prog (o) + Oest (o)	87.77%	0.55%	1 (1 to 8)
GnRHa (i.m.) + Prog (o)	4.91%	29.72%	10 (3 to 15)
GnRHa (i.m.) + Prog (o) + Oest (o)	5.87%	17.80%	8 (3 to 14)
GnRHa (s.c.) + Prog (o)	19.13%	38.84%	9 (1 to 15)
GnRHa (s.c.) + Oest (o) + Prog (o)	58.78%	4.82%	3 (1 to 14)
GnRH (i.n.) + Prog (o)	1.64%	78.75%	15 (4 to 15)

For treatment name abbreviations, see Table 57

11.1.3.3 Pairwise comparison

11.1.3.3.1 Description of clinical evidence

This pairwise comparison analysis accompanies the NMA that examined pain (VAS total scores and Biblioglu and Behrman criteria and) and withdrawal due to adverse events. The potential evidence for this analysis included RCTs identified from the searches performed on the basis of the protocol (see Appendix D) as well as RCTs that were considered for the NMA.

In total, 7 studies were included in this review. Three were Cochrane systematic reviews (Brown 2010, 2012 and Davis 2007) and 4 were RCTs (Harada 2008, Ling 1999, Parazzini 2000 and Schlaff 2006). 10 RCTs from the Brown 2010 (Agarwal 1997, Bergqvist 1998, Burry 1992, Cheng 2005, Fedele 1989, Fedele 1993, Fraser 1991, NEET 1992, Petta 2005, Wheeler 1992), 2 RCTs from the Brown 2012 (Bergqvist 2001, Vercellini 1996) and 1 RCT from the Davis 2007 (Vercellini 1993) Cochrane systematic reviews were relevant.

The population of interest was women with suspected or confirmed endometriosis of any stage or severity who did not receive surgery in conjunction with the hormonal medical treatments, although who may have had surgery prior to trial recruitment. Evidence was available for comparisons of hormonal treatments with placebo or no treatment (4 RCTs), for head to head hormonal treatment comparisons with placebo (6 RCTs) or without placebo (5 RCT) use in each treatment arm and for hormonal treatment combinations compared with a single hormonal treatment (2 RCTs).

The Committee specified critical outcomes of pain (for outcomes not included in the NMA), quality of life and unintended effects from treatment. However, many reports of unintended effects were identified (type, incidence and duration of side effects), preventing their meaningful inclusion in the pairwise analysis. Therefore these were addressed as ‘withdrawal from hormonal treatment due to adverse events’ in the NMA.

Hormonal treatments compared with placebo

Evidence was available from 3 studies that compared hormonal treatments with placebo or no treatment. One was a Cochrane systematic review (Brown 2010) and 2 were RCTs (Harada 2008 and Ling 1999). Two RCTs within the Cochrane systematic review were relevant (Bergqvist 1998, Fedele 1993).

ISBN 978-1-4731-2661-9

All participants had a diagnosis or symptoms of endometriosis. One RCT examined a comparison of a GnRH agonist (buserelin intranasal (IN)) to expectant management in a population of women whose main symptom was infertility and who had undergone diagnostic laparoscopy combined with dilation and curettage (D&C) (Fedele 1993).

Two RCTs examined comparisons of GnRH agonists to placebo (triptorelin IM (intramuscular) depot and leuprolide IM depot) (Bergqvist 1998 and Ling 1999, respectively). One RCT compared a combined oral contraceptive pill to placebo (Harada 2008).

Evidence was only available for the critical outcome of pain (outcomes not included in the NMAs). There was no evidence available for any other critical or important outcomes.

Hormonal treatment compared with another hormonal treatment

Evidence was available from 2 studies comparing a hormonal treatment to another hormonal treatment. One was a Cochrane systematic review (Brown 2010) and one was a RCT (Schlaff 2006). Four RCTs within the Cochrane systematic review were relevant (Burry 1992, Cheng 2005, Fedele 1989, Petta 2005).

Three RCTs examined a comparison of a GnRH agonist (nafarelin IN or buserelin IN) to danazol (Burry 1992, Cheng 2005, Fedele 1989). One RCT compared leuprolide IM to a levonorgestrel-releasing intrauterine system (LNG-IUS) to (Petta 2010) and 1 RCT compared leuprolide to depot medroxyprogesterone acetate (DMPA) subcutaneous (SC) injections (Schlaff 2006). All participants had laparoscopically confirmed endometriosis. One trial (Fedele 1989) included infertile women only.

Evidence was available for the critical outcomes of pain (outcomes not included in the NMA) and quality of life, and for the important outcomes of patients requiring surgery because of reappearance of symptoms and the effect on daily activities. There was no evidence available for any other important outcomes.

Hormonal treatment with placebo compared with another hormonal treatment with placebo

Evidence was available from 2 Cochrane systematic reviews (Brown 2010; Brown 2012) comparing a GnRH agonist to another hormonal treatment with use of placebos in each trial arm to blind for route of administration. Five RCTs were relevant in total: 4 RCTs from the Brown 2010 Cochrane systematic review (Agarwal 1997, Fraser 1991, NEET 1992, Wheeler 1992); and 1 RCT from the Brown 2012 Cochrane systematic review (Bergqvist 2001).

Four trials examined intranasal nafarelin (Agarwal 1997, Bergqvist 2001, Fraser 1991, NEET 1992) and 1 trial examined the use of depot leuprolide (Wheeler 1992).

One RCT examined a comparison of nafarelin IN and placebo IM injections to leuprolide acetate depot intramuscular (IM) injections and placebo IN (Agarwal 1997). One RCT compared nafarelin IN plus placebo tablets twice daily to MPA tablets and placebo IN (Bergqvist 2001).

Three trials compared the use of a GnRH agonist to danazol (Fraser 1991, NEET 1992, Wheeler 1992). Two trials compared the use of nafarelin IN to danazol with placebo in both treatment arms (Fraser 1991, NEET 1992). The first RCT compared of nafarelin IN and oral placebo to oral danazol and placebo IN over 6 months (Fraser 1991). The second RCT compared nafarelin IN and oral placebo capsules to oral danazol capsules and IN placebo (NEET 1992).

The final RCT compared a form of leuprolide depot injections and oral placebo to danazol and placebo IM injections (Wheeler 1992).

ISBN 978-1-4731-2661-9

Evidence was available for the critical outcomes of pain relief (those outcomes not included in the NMA) and quality of life and for the important outcome of effects on daily activities. There was no evidence available for any important outcomes.

Hormonal treatment compared with combined oral contraceptive pill

Three studies comparing hormonal treatment to combined oral contraceptive pill (cOCP) were included in this review. Evidence was available from 2 Cochrane systematic reviews (Davis 2007, Brown 2012) and 1 RCT (Parazzini 2000). One RCT within each Cochrane systematic review was relevant (Vercellini 1993 and 1996, respectively).

All participants had laparoscopically confirmed endometriosis. One RCT examined a comparison of a GnRH agonist (triptorelin slow release for 4 months) followed by treatment with gestodene and ethinylestradiol (E/P pill) to E/P pill alone (Parazzini 2000). One RCT compared a GnRH agonist (goserelin subcutaneous depot) to cOCP (ethinylestradiol and desogestrel) (Vercellini 1993) and 1 RCT compared depot medroxyprogesterone acetate to cOCP (ethinylestradiol and desogestrel) plus danazol (Vercellini 2012). In 1 study (Parazzini 2000) additional treatment for relief of pain with naproxen sodium as first-line treatment was allowed.

Evidence was available for the critical outcome of pain (outcomes not included in the NMA) and for an important outcome of patient satisfaction. There was no evidence available for any other critical or important outcomes.

Studies are summarised in the tables below Table 71 and the available evidence is presented by comparison in the clinical GRADE evidence profiles below (Table 72 to Table 85). See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G. Summary of included studies

A summary of the studies that were included in this review are presented in Table 71.

Table 71: Summary of included studies

Study	Intervention/Comparison	Population	Outcomes	Comments
Agarwal 1997 (Brown 2010 CSR) USA	Nafarelin 200mcg BDS IN + placebo every 4 weeks IM for 6 months (n=105) versus LA Depot 3.75mg every 4 weeks IM + placebo BDS IN for 6 months (n=103)	n=208* Inclusion criteria: Laparoscopically diagnosed endometriosis within 18 months prior to study, 19–44 years old, patients demonstrating clinical symptoms and signs, bone mineral density within normal age range	Assessed at 6 months after the end of treatment period Pelvic tenderness Pelvic induration Measured using a 4-point numerical scale: 0=none; 1=mild; 2=moderate; 3=severe	*No information on stages of endometriosis provided
Bergqvist 1998 (Brown 2010 CSR) Sweden	Triptorelin 3.75mg IM depot every 4 weeks for 24 weeks (n=24) vs. Placebo IM every 4 weeks for 24 weeks (n=25)	Inclusion criteria: Menstruating regularly 3 months before study Clinical symptoms of endometriosis Not taken oral contraceptive or	PAIN: VAS (0 to 10) and direct questions about pelvic pain, dysmenorrhoea and dyspareunia (none, mild, moderate, severe)	

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes	Comments
		oral steroid therapy for 3 months Not taken long-acting depot gestagens or GnRH agonists within past 6 months Not pregnant in prior 3 months Not breastfeeding No history of osteoporosis or coagulation disorders		
Bergqvist 2001 (Brown 2012) Sweden	Nafarelin 200 µg BDS IN and placebo MPA tablets (n=23) for 6 months versus MPA 15mg PO BDS and placebo nafarelin IN (n=25) for 6 months	n=48* Inclusion criteria: Diagnosis of endometriosis by laparoscopy or laparotomy, within 3 months regular menstruating and complaining of dysmenorrhoea, dyspareunia and/or pelvic pain	Assessed at 6 months (at the end of treatment) and 12 months (6 months after the end of the treatment period) QoL (Goldberg's general health and Nottingham Health Profile Questionnaire) Effect on daily activities (coping wheel, Inventory of Social Support and Interaction – (ISSI) and demands, control and support questionnaires)	*No information on stages of endometriosis provided
Burry 1992 (Brown 2010 CSR) USA	Nafarelin 400mcg daily IN for 6 months (n=111) versus Danazol 600mg daily PO for 6 months (n=58)	N=169 Inclusion criteria: Laparoscopically diagnosed endometriosis	Assessed at the end of the 6 month treatment period QoL (PGWBI plus a modification of Part II of the Nottingham Health Profile)	
Cheng 2005 (Brown 2010 CSR) Taiwan	Nafarelin acetate 200mcg BDS (400mcg/day) IN for 180 days (n=29) versus Danazol 200mg TDS (600mg/day) PO for 180 days (n=30)	N=59 Inclusion criteria: Laparoscopically diagnosed endometriosis within 3 months prior to study, age 18–48 years, barrier contraception	Assessed at 3 months (during treatment period) and at the end of the 6 month treatment period Pelvic tenderness Pelvic induration (TSSS, scale not defined)	

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes	Comments
Fedele 1989 (Brown 2010 CSR) Italy	Buserelin 400mcg TDS IN for 6 months (n=30) versus Danazol 200mg TDS PO for 6 months (n=32)	N=62 Inclusion criteria: Laparoscopically diagnosed endometriosis within 3 months prior to study and no therapeutic intervention	Assessed at 12 months post-treatment Patients requiring surgery because of reappearance of symptoms and positive findings at pelvic examination	Infertile women included
Fedele 1993 (Brown 2010 CSR) Italy	Buserelin acetate 1200mcg daily IN for 6 months (n=19) versus Expectant management (n=16) Treatment group followed up for 18 months and expectant management group for 12 months	Inclusion criteria: Laparoscopically diagnosed endometriosis One or more of dysmenorrhoea, pelvic pain and deep dyspareunia	Pain: dysmenorrhoea and pelvic pain measured by VAS (0 to 10): 0 (no pain); 1 to 4 (mild;, 5 to 7 (moderate); 8 to 10 (severe); deep dyspareunia	Population of women whose main symptom is infertility and who may not have had pain as a symptom at baseline
Fraser 1991 (Brown 2010 CSR) Australia/ New Zealand	Nafarelin 200mcg BDS (400mcg/d) IN + placebo PO for 6 months (n=33) versus Danazol 200mg TDS (600mg/d) PO + placebo IN for 6 months (n=16)	n=49* n=40 patients with stage I–II n=9 patients with stage III Inclusion criteria: Laparoscopically diagnosed endometriosis, symptomatic, regular menstrual cycle 24–36 days, not pregnant, negative pap smear, barrier contraception	Assessed at 6 months after the end of the treatment period Pelvic tenderness Pelvic induration Measured using a 4-point numerical scale: 0=none; 1=mild; 2=moderate; 3=severe	* American Fertility Society classification. 18 women dropped out of the study – no reasons were provided. However, 17 of these women responded to the psychosocial questions. Anxiety-depression was significantly more common among women who dropped out compared to the 30 women analysed (p=0.04).
Harada 2008 Japan	Monophasic oral contraceptive pill (ethinylestradiol 0.035mg plus norethisterone 1mg) for 21 days plus 7 days of placebo for 3 cycles (n=49) versus Placebo for 28 days for 3 cycles (n=47)	Inclusion criteria: Aged 18 or over Regular menstrual cycles (28+/- 2 days) Symptomatic endometriosis (diagnosed by laparoscopy or laparotomy) or ovarian	PAIN: dysmenorrhoea and non-menstrual pelvic pain scores Pelvic induration (hardening of soft tissues): physical examination	VAS not defined

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes	Comments
		endometrioma (diagnosed by ultrasound or MRI) Normal cervical and endometrial smear cytology Moderate or severe dysmenorrhoea (evaluated by a modified pain scale) No medical or surgical treatment for endometriosis within 8 weeks before entry into the study		
Ling 1999 USA	Leuprolide acetate 3.75mg IM depot every 4 weeks on day 0, week 4 and week 8 (n=49) versus Placebo IM every 4 weeks on day 0, week 4 and week 8 (n=46)	Inclusion criteria: 18–45 years of age Moderate to severe chronic pelvic pain for at least 6 months, unrelated to menstruation and incompletely resolved with NSAIDs Physician-assessed pain severity (B&B) Regular menstrual bleeding and menstrual cycles for 3 months prior to enrolment	PAIN: dysmenorrhoea, pelvic pain, dyspareunia, based on an 11-point VAS (0 to 10)	
NEET 1992 (Brown 2010 CSR) Europe	Nafarelin 200mcg BDS IN + placebo PO for 6 months (n=206) versus Danazol 200mg TDS PO + placebo IN for 6 months (n=101)	n=263*: n=160 patients with stage I–II n=103 patients with stage III–IV Inclusion criteria: Laparoscopically diagnosed endometriosis, 18–45 years old, not pregnant, pap smear negative for malignancy, normal menstrual cycle 21–36 days for previous 4 months, weight between 45 and 110 kg	Assessed at 12 months after the end of the treatment period** Pelvic tenderness Pelvic induration	* American Fertility Society classification ** after 12 months follow-up only 96 out of the 263 included women were analysed, main reason for dropping out the study were: 1) pregnancy 2) further medical therapy for endometriosis 3) hormonal therapy for other medical conditions

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes	Comments
				4) loss to follow-up
Parazzini 2000 Italy	GnRH agonist (triptorelin 3.75 mg) slow release every 28 days for 4 months followed by gestodene 0.75 mg/ ethinylestradiol 0.03 mg (E/P pill) for 8 months (n=55) versus gestodene 0.75 mg/ ethinylestradiol 0.03 mg (E/P pill) for 12 months (n=47)	N=102 Inclusion criteria: Laparoscopically diagnosed endometriosis and pelvic pain lasting 3–12 months after diagnosis. Additionally, only women who reported a score of >=3 for the multidimensional scale and/or >=5 for the analogue scale for dysmenorrhoea and/or non-menstrual pelvic pain were eligible 51.9% in the GnRH agonist+ E/P group and 57.8% in the E/P group had stage I–II endometriosis	Assessed at 8 months during treatment period and at the end of the treatment period (12 months) Dysmenorrhoea Non-menstrual pelvic pain (a 10-point VAS scale where 0 = the absence of pain, 10 = unbearable pain)	Additional treatment for relief of pain with naproxen sodium as first-line treatment was allowed
Petta 2005 (Brown 2010 CSR) Brazilian	Lupron 3.75mg every 28 days IM for 6 months (n=43) versus LNG-IUS (Mirena) 20mcg/day 5 years IU for 6 months (n=40)	N=83 Inclusion criteria: Laparoscopically and histologically confirmed endometriosis within 3 to 24 months prior to study enrolment, 18–40 years old, complaints of cyclic chronic pelvic pain with or without dysmenorrhoea, VAS pain score of greater or equal to 3 during the pre-treatment cycle, regular menstrual cycle of 25–35 days for at least 3 months prior to study, not used hormone treatment for at least 3 months prior to study, not taken any long-acting progestogens or GnRH agonist within 9 months	Assessed at the end of the 6 month treatment period QoL – psychological wellbeing (PGWBI)	

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes	Comments
		prior to study, no osteoporosis, coagulation disorders or contra-indications Stage: I to IV		
Schlaff 2006 USA	Leuprolide (11.25 mg given by IM injection) versus DMPA-SC 104 (104 mg/0.65 mL given by SC injection)	N=274 Inclusion criteria: Premenopausal women who ranged in age from 18 to 49 years, with persistent symptoms of pain caused by endometriosis (surgically diagnosed within the previous 42 months). A patient's pain must have returned to its previous level within 30 days after a diagnostic laparoscopy or within 3 months after laparoscopy or laparotomy with surgical treatment, and it must have persisted for a minimum of 3 months.	Assessed at the end of the 6 month treatment period and 18 months (12 months post-treatment) Effect on daily activities: Total hours of productivity lost at employment Total hours of productivity lost at housework (Endometriosis-impact diary)	
Vercellini 1993 (Davis 2007 CSR) Italy	GnRH agonist (goserelin 3.6 mg) subcutaneous depot formulation monthly for 6 months (n=29) versus a low-dose cyclic monophasic contraceptive pill, containing 0.02 mg ethinylestradiol and 0.15 mg desogestrel (n=28) for 6 months In the goserelin group the follow-up period was considered to start 4 weeks after the last injection	N=57 Inclusion criteria: Laparoscopically diagnosed endometriosis and no attempts at endometriosis reduction other than biopsy within 3 months of study entry 76% and 24% in the GnRH agonist group has stage I–II endometriosis; 82% and 28% in the cOCP group had stage III–IV endometriosis	Assessed at the end of the treatment period (6 months) and 6 months after the treatment period Dysmenorrhoea Dyspareunia Non-menstrual pelvic pain (a 10-point VAS scale where 0 = the absence of pain, 1–5 = mild pain, 6–7 = moderate pain, 8–10 = unbearable pain)	In the cOCP group, if spotting or breakthrough bleeding occurred, women could switch to a contraceptive with EE2, 0.03 mg and desogestrel 0.15 mg per pill
Vercellini 1996 (Brown 2012 CSR)	Depot medroxyprogesterone acetate 150 mg every 90 days (n=40) versus	N=80 Inclusion criteria: Laparoscopically diagnosed endometriosis with	Assessed at 6 months during the treatment period and at the end of the	

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes	Comments
Italy	Ethinylestradiol 0.02 mg + desogestrel 0.15mg plus 50 mg danazol daily for 21 days out of 28-day cycle (n=40) for 12 months	attempt at implant reduction other than biopsy in the previous 3 months, pelvic pain of greater than 6 months duration. Additionally, only women who had at least 1 moderate or severe symptom on a verbal rating scale modified from the one devised by Biberoglu and Behrman and on a 10 point visual analogue scale were eligible 55% and 45% women in both groups had stage I–II or III-IV endometriosis	treatment period (12 months) Dysmenorrhoea Dyspareunia Non-menstrual pelvic pain (a 10-cm VAS where 0 = absence of pain, >0–5 = mild pain, >5–8 = moderate pain, >8–10 = unbearable pain) Patient satisfaction (no particular scale defined: very satisfied; satisfied; uncertain; dissatisfied; very dissatisfied)	
Wheeler 1992 (Brown 2010 CSR) USA	Leuprolide 3.75mg monthly IM + placebo OD PO for 24 weeks (n=134) versus Danazol 800mg OD PO + placebo monthly IM for 24 weeks (n=136)	n=270* Inclusion criteria: Laparoscopically diagnosed endometriosis within 4 months prior to study, over 18 years of age, no surgical treatment at time of laparoscopy, premenopausal, not pregnant or lactating, never previously taken GnRH agonist, any other treatment completed at least 3 months prior to study	Assessed at 6 months after the end of the treatment period Pelvic tenderness	*No information on stages of endometriosis provided

BDS: twice per day; cOCP: combined oral contraceptive; CSR: Cochrane systematic review; DMPA-SC: depot medroxyprogesterone acetate; GnRH: gonadotrophin releasing hormone; IM: intramuscular; IN: intranasal; LA: leuprolide acetate; LNG-IUS: levonorgestrel-releasing intrauterine system; MPA: medroxyprogesterone acetate; NSAID: non-steroidal anti-inflammatory drug; OD: once per day; PGWBI: psychological well-being index questionnaire (scale 0–110); PO: by mouth; QoL: quality of life; SF-36: Short form questionnaire (36 items); TDS: 3 times per day; TSSS: total symptom severity score; 4-point scale: each symptom or sign was scored on a 4-point system at each visit from the cards prospectively recorded by each patient or at the vaginal examination (0=none, 1=mild, 2=moderate, 3 severe); VAS: Visual Analogue Scale

11.1.3.3.2 Clinical evidence profile

The clinical evidence profiles for this review question are presented in Table 72 to Table 85.

ISBN 978-1-4731-2661-9

Table 72: Summary clinical evidence profile, comparison 1: GnRH agonist versus no treatment

Outcomes	Illustrative comparative risks (95% CI)		Rela-tive effect (95% CI)	No of Partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	No treatment	GnRH agonist			
Dysmenorrhoea relief at 12 months scale: 0 (no pain); 1 to 4 (mild); 5 to 7 (moderate); 8 to 10 (severe)	188 per 1,000	579 per 1,000 (195 to 1,000)	RR 3.09 (1.04 to 9.18)	35 (1 study)	⊕⊖⊖⊖ Very low1,2

CI: confidence interval; GnRH: gonadotrophin releasing hormone; RR: risk ratio
1 The main symptom of the study population was not pain (infertility)
2 CI crosses 1 threshold

Table 73: Summary clinical evidence profile, comparison 2: GnRH agonist versus placebo

Outcomes	Illustrative comparative risks (95% CI)		Rela-tive effect (95% CI)	**No of Partici-pants (studies)**	**Quality of the evidence (GRADE)**
	Assumed risk	**Corresponding risk**			
	Placebo	**GnRH agonist**			
Mean dysmenorrhoea at week 12 (11-point VAS)	-	The mean dysmenorrhoea at week 12 in the intervention groups was 6.30 lower (9.93 to 2.67 lower)	MD -6.30 (-9.93 to -2.67)	88 (1 study)	⊕⊕⊕⊖ Moderate1
Mean pelvic pain at week 12 (11-point VAS)	-	The mean pelvic pain at week 12 in the intervention groups was 4.4 lower (6.93 to 1.87 lower)	MD -4.4 (6.93 to -1.87)	88 (1 study)	⊕⊕⊕⊖ Moderate1
Mean deep dyspareunia at week 12 (11-point VAS)	-	The mean deep dyspareunia at week 12 in the intervention groups was 3.1 lower (4.85 to 1.35 lower)	MD -3.1 (-4.85 to -1.35)	61 (1 study)	⊕⊕⊕⊖ Moderate1
Pelvic tenderness cessation at 6 months	174 per 1,000	696 per 1,000 (275 to 1,000)	RR 4 (1.58 to 10.15)	46 (1 study)	⊕⊕⊕⊖ Moderate2
Dyspareunia cessation at 6 months	391 per 1,000	434 per 1,000 (219 to 869)	RR 1.11 (0.56 to 2.22)	46 (1 study)	⊕⊖⊖⊖ Very low2,3

CI: confidence interval; CSR: Cochrane systematic review; GnRH: gonadotrophin releasing hormone; RR: risk ratio; MD: mean difference; VAS: visual analogue scale
1 Outcomes measured immediately after treatment period are of less clinical relevance than sustained post-treatment effects
2 No details provided regarding sequence generation and allocation concealment (unclear risk)
3 CIs for estimate are very wide crossing 2 thresholds

ISBN 978-1-4731-2661-9

Table 74: Summary clinical evidence profile, comparison 3: Combined oral contraceptive pill versus placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Placebo	Oral contraceptive			
Dysmenorrhoea (VAS not defined, reported on a scale 0 to 100)	-	The mean dysmenorrhoea in the intervention groups was 21.5 lower (28.14 to 14.86 lower)	MD -21.5 (-28.14 to 14.86)	96 (1 study)	⊕⊕⊕⊖ Moderate1
Non-menstrual pelvic pain score (VAS not defined, reported on a scale 0 to 100)	-	The mean non-menstrual pelvic pain score(VAS) in the intervention groups was 6.6 lower (14.27 lower to 1.07 higher)	-	96 (1 study)	⊕⊕⊖⊖ Low1,2
Induration	404 per 1,000	226 per 1,000 (121 to 420)	RR 0.56 (0.3 to 1.04)	96 (1 study)	⊕⊕⊖⊖ Low[1,2]

CI: confidence interval; RR: risk ratio; MD: mean difference; VAS: visual analogue scale
1 Short duration of treatment is of limited relevance to clinical practice
2 CI crosses 1 threshold

Table 75: Summary clinical evidence profile, comparison 4: GnRH agonist versus danazol

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Control	GnRH agonist versus danazol			
Pelvic tenderness at 3 months (TSSS, scale not defined)	-	The mean pelvic tenderness at 3 months in the intervention groups was 0.2 lower (0.78 lower to 0.38 higher)	MD -0.2 (-0.78 to -0.38)	41 (1 study)	⊕⊕⊕⊖ Moderate[1]
Pelvic tenderness at 6 months (TSSS, scale not defined)	-	The mean pelvic tenderness at 6 months in the intervention groups was 0.2 lower (0.75 lower to 0.35 higher)	MD -0.2 (-0.75 to 0.35)	41 (1 study)	⊕⊕⊕⊖ Moderate[1]
Pelvic induration at 3 months (TSSS, scale not defined)	-	The mean pelvic induration at 3 months in the intervention groups was	MD -0.1 (-0.59 to 0.39)	41 (1 study)	⊕⊕⊖⊖ Low[2]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
		0.1 lower (0.59 lower to 0.39 higher)			
Pelvic induration- at 6 months (TSSS, scale not defined)	-	The mean pelvic induration at 6 months in the intervention groups was 0.2 higher (0.29 lower to 0.69 higher)	MD 0.2 (-0.29 to 0.69)	41 (1 study)	⊕⊕⊕⊖ Moderate[1]
Patients requiring surgery because of reappearance of symptoms and positive findings at pelvic examination at 6 months	357 per 1,000	364 per 1,000 (129 to 1,000)	RR 1.02 (0.36 to 2.91)	25 (1 study)	⊕⊕⊕⊖ Moderate[1]
QoL at 6 months (PGWBI plus a modification of Part II of the Nottingham Health Profile	-	No statistically significant difference between the 2 intervention groups	Not estimable	169 (1 study)	⊕⊕⊖⊖ Low[3]

CI: confidence interval; GnRH: Gonadotrophin Releasing Hormone; RR: relative risk; MD: mean difference; PGWBI: Psychological General Well-Being Index; TSSS: Total Symptom Severity Score; 4-point scale: each symptom or sign was scored on a 4-point system at each visit from the cards prospectively recorded by each patient or at the vaginal examination (0=none, 1=mild, 2=moderate, 3 severe); QoL: quality of life
1 CI crosses 1 threshold
2 CI crosses 2 thresholds
3 Reporting bias, i.e. not possible to access imprecision as only descriptive data reported

Table 76: Summary clinical evidence profile, comparison 5: GnRH agonist versus levonorgestrel-releasing intrauterine system

Outcomes	Illustrative comparative risks (95% CI)		Rela-tive effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Control	GnRH agonist versus levonorgestrel-releasing intrauterine system			
QoL at 6 months (PGWBI, 0–110 scale)	-	The mean QoL (PGWBI) at 6 months in the intervention groups was 1.2 lower (7.79 lower to 5.39 higher)	MD -1.2 (-7.79 to 5.39)	72 (1 study)	⊕⊕⊕⊖ Moderate[1]

CI: confidence interval; GnRH: gonadotrophin releasing hormone; QoL: quality of life; PGWBI: Psychological Well-Being Index questionnaire
1 CI crosses 1 threshold

ISBN 978-1-4731-2661-9

Table 77: Summary clinical evidence profile, comparison 6: GnRH agonist versus DMPA-SC

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Control	GnRH agonist versus DMPA-SC			
Effect on daily activities from baseline to 6 month follow-up (Endometriosis-impact diary)	-	The mean number of hours of productivity lost at employment at 6 months in the intervention groups was 6.15 higher(2.17 lower to 14.47 higher)	MD 6.15 (-2.17 to 14.47)	190 (1 study)	⊕⊕⊕⊕ High
Effect on daily activities from baseline to 18 month follow-up (Endometriosis-impact diary)	-	The mean number of hours of productivity lost at employment at 18 months in the intervention groups was 6.38 higher (1.94 lower to 14.7 higher)	MD 6.38 (-1.94 to 14.7)	190 (1 study)	⊕⊕⊕⊕ High[1]
Effect on daily activities from baseline to 6 month follow-up (Endometriosis-impact diary)	-	The mean number of hours of productivity lost at housework at 6 months in the intervention groups was 7.35 lower (16.63 lower to 1.93 higher)	MD -7.35 (-16.63 to 1.93)	81 (1 study)	⊕⊕⊕⊖ Moderate[1]
Effect on daily activities from baseline to 18 month follow-up (Endometriosis-impact diary)	-	The mean number of hours of productivity lost at housework at 18 months in the intervention groups was 3.64 lower (12.92 lower to 5.64 higher)	MD -3.64 (-12.92 to 5.64)	81 (1 study)	⊕⊕⊕⊖ Moderate[1]

CI: confidence interval; GnRH: gonadotrophin releasing hormone; DMPA-SC: depot medroxyprogesterone acetate
1 CI crosses 1 threshold

Table 78: Summary clinical evidence profile, comparison 7: GnRH agonist 1 + placebo versus GnRH agonist 2 + placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	GnRH agonist (LA depot IM) + placebo IN	GnRH agonist (nafarelin IN) + placebo IM			
Relief of painful symptoms – Pelvic tenderness Follow-up: 6 months[1]	624 per 1,000	536 per 1,000 (418 to 680)	RR 0.86 (0.67 to 1.09)	192 (1 study)	⊕⊕⊖⊖ Low[2]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
Relief of painful symptoms–Pelvic induration Follow-up: 6 months[1]	813 per 1,000	740 per 1,000 (634 to 862)	RR 0.91 (0.78 to 1.06)	190 (1 study)	⊕⊕⊖⊖ Low[2]

CI: confidence interval; GnRH: gonadotrophin releasing hormone; IM: intramuscular; IN: intranasal; RR: risk ratio
1 Assessed after the end of the treatment period
2 Quality of evidence was downgraded by 2 points owing to very serious imprecision: CI crosses 2 default thresholds

Table 79: Summary clinical evidence profile, comparison 8: GnRH agonist + placebo versus progestin + placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	MPA and placebo nasal spray	GnRH agonist (nafarelin) IN + placebo tablets				
Paid working life Nottingham Health Profile Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate an improvement in the nafarelin group, but not in the MPA group (p=0.06)
Household work Nottingham Health Profile Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in household work score (data not shown)
Vacation life Nottingham Health Profile Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in vacation life score (p=0.72)
Leisure Nottingham Health Profile Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
						leisure score (p=0.93)
Sexual life Nottingham Health Profile Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in sexual life score (p=0.90)
Disturbed sleep Goldberg's General Health Q Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in sleep disturbance (difficulties of falling asleep, early wakening and nightmares) score (p=0.19)
Anxiety-depression Goldberg's General Health Q Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in anxiety-depression score (p=0.20)
Motivation coping wheel, ISSI and demands, control and support Q Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in motivation score (p=0.41)
Emotional balance Coping wheel, ISSI and demands, control and support Q Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in emotional balance score (p=0.44)
Structure Coping wheel, ISSI and demands,	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
control and support Q Follow-up[3]						groups in structure score (p=0.41)
Coping Coping wheel, ISSI and demands, control and support Q Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in coping score (p=0.39)
Psychological work demands Coping wheel, ISSI and demands, control and support Q Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in 'psychological work demands' score (p=0.51)
Intellectual discretion at work Coping wheel, ISSI and demands, control and support Q Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in 'intellectual discretion at work' score (p=0.95)
Authority over decisions at work Coping wheel, ISSI and demands, control and support Q Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in 'authority over decisions at work' score (p=0.39)
Social support at work Coping wheel, ISSI and demands, control and support Q Follow-up[3]	See comment	See comment	Not estimable	30 (1 study)	⊕⊖⊖⊖ Very low[1,2]	The results indicate no significant difference between groups in 'social support at work' score (p=0.68)

CI: confidence interval; GnRH: gonadotrophin releasing hormone; ISSI: Inventory of Social Support and Interaction; Q: questionnaire
1 The quality of the evidence was downgraded of 2 points because of the high risk of reporting bias (i.e. not Possible to access imprecision as only descriptive data with p values reported) and the potential risk of detection Bias (no details were given about randomisation and allocation concealment methods).

ISBN 978-1-4731-2661-9

2 Only descriptive data reported, insufficient details given to assess the minimally important difference threshold and the imprecision
3 Follow-up at 6 months (at the end of the treatment period) and 12 months (6 months after the end of the treatment period) using analysis of variance (ANOVA) for repeated measures (mixed model)

Table 80: Summary clinical evidence profile Comparison 9: GnRH agonist + placebo versus danazol + placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Oral danazol + IN placebo	**GnRH agonist (nafarelin) + oral placebo**			
Relief of painful symptoms – Pelvic tenderness 4-point numerical scale Follow-up: 6 months[1]	-	The mean relief of painful symptoms – - pelvic tenderness in the intervention groups was 0.1 lower (0.38 lower to 0.18 higher)	Not estimable	49 (1 study)	⊕⊖⊖⊖ Very low[2,3]
Relief of painful symptoms - Pelvic induration 4-point numerical scale. Follow-up: 6 months[1]	-	The mean relief of painful symptoms – pelvic induration in the intervention groups was 0 higher (0.28 lower to 0.28 higher)	Not estimable	49 (1 study)	⊕⊖⊖⊖ Very low[2,3]

CI: confidence interval; GnRH: gonadotrophin releasing hormone; IN: intranasal; 4-point scale: each symptom or sign was scored on a 4-point system at each visit from the cards prospectively recorded by each patient or at the vaginal examination (0=none, 1=mild, 2=moderate, 3 severe)
1 Assessed after the end of the treatment period
2 Quality of evidence was downgraded by 1 point owing to unclear risk of selection bias (no details given about allocation concealment methods)
3 Quality of evidence was further downgraded by 2 points owing to very serious imprecision: CI crosses 2 default thresholds

Table 81: Summary clinical evidence profile Comparison 9: GnRH agonist + placebo versus danazol + placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Danazol + placebo nasal spray	**GnRH agonist (nafarelin IN) + oral placebo TDS**			
Relief of painful symptoms – Pelvic tenderness	742 per 1,000	772 per 1,000 (601 to 987)	RR 1.04 (0.81 to 1.33)	96 (1 study)	⊕⊖⊖⊖ very Low[2,3]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
Follow-up: 12 months[1]					
Relief of painful symptoms – Pelvic induration Follow-up: 12 months[1]	871 per 1,000	906 per 1,000 (775 to 1,000)	RR 1.04 (0.89 to 1.22)	96 (1 study)	⊕⊖⊖⊖ Very low[2,3]

CI: confidence interval; GnRH: gonadotrophin releasing hormone; IN: intranasal; RR: risk ratio; TDS: 3 times per day
1 Assessed after the end of the treatment period
2 Quality of evidence was downgraded by 1 point owing to unclear risk of selection bias (no details about allocation concealment method and unclear description of the allocation concealment procedure)
3 Quality of evidence was downgraded by 2 points owing to very serious imprecision: CI crosses 2 default thresholds

Table 82: Summary clinical evidence profile, comparison 9: GnRH agonist + placebo versus danazol + placebo

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Danazol OD PO + placebo IM	GnRH agonist (leuprolide IM) + placebo OD PO			
Relief of painful symptoms – Pelvic tenderness Follow-up: 6 months[1]	760 per 1,000	730 per 1,000 (631 to 844)	RR 0.96 (0.83 to 1.11)	253 (1 study)	⊕⊕⊖⊖ Low[2,3]

CI: confidence interval; GnRH: gonadotrophin releasing hormone; IM: intramuscular; OD: once per day; PO: by mouth; RR: risk ratio;
1 Assessed after the end of the treatment period
2 Quality of evidence was downgraded by 1 point owing to unclear risk of detection bias (no details were given about randomisation and allocation concealment methods)
3 Quality of evidence was further downgraded by 1 point owing to serious imprecision: CI crosses 1 default threshold and p is higher than 0.1

Table 83: Summary clinical evidence profile, comparison 10: Depot medroxyprogesterone acetate versus cOCP + danazol

Outcomes	Illustrative comparative risks (95% CI)		Rela-tive effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	cOCP + desogestrel	Depot medroxyprogesterone			
Pain at 6 months during treatment period – Dysmenorrhoea	-	The mean pain at 6 months during treatment period – dysmenorrhoea in the intervention groups was 1.84 lower (2.23 to 1.45 lower)	MD -1.84 (-2.23 to -1.45)	68 (1 study)	⊕⊕⊕⊖ Moderate[1]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Rela-tive effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
(Scale: 10 cm VAS)[4]					
Pain at 6 months during treatment period – Dyspareunia (Scale: 10 cm VAS) [4]	-	The mean pain at 6 months during treatment period - dyspareunia in the intervention groups was 0.3 lower (1.18 lower to 0.58 higher)	MD -0.3 (-1.18 to 0.58)	59 (1 study)	⊕⊕⊖⊖ Low[1,2]
Pain at 6 months during treatment period – Non-menstrual pelvic pain (Scale: 10 cm VAS) [4]	-	The mean pain at 6 months during treatment period – non-menstrual pelvic pain in the intervention groups was 0.6 higher (0.09 lower to 1.29 higher)	MD 0.6 (-0.09 to 1.29)	68 (1 study)	⊕⊖⊖⊖ Very low[1,3]
Pain at the end of treatment period (12 months) – Dysmenorrhoea (Scale: 10 cm VAS) [4]	-	The mean pain at the end of treatment period (12 months) – dysmenorrhoea in the intervention groups was 1.3 lower (1.79 to 0.81 lower)	MD -1.3 (-1.79 to -0.81)	68 (1 study)	⊕⊕⊕⊖ Moderate[1]
Pain at the end of treatment period (12 months) – Dyspareunia (Scale: 10 cm VAS) [4]	-	The mean pain at the end of treatment period (12 months) – dyspareunia in the intervention groups was 0.3 lower (1.41 lower to 0.81 higher)	MD -0.3 (-1.41 to 0.81)	59 (1 study)	⊕⊕⊖⊖ Low[1,2]
Pain at the end of treatment period (12 months) – Non-menstrual pelvic pain (Scale: 10 cm VAS) [4]	-	The mean pain at the end of treatment period (12 months) – non-menstrual pelvic pain in the intervention groups was 0.4 higher (0.35 lower to 1.15 higher)	MD 0.4 (-0.35 to 1.15)	68 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Patient satisfaction (very satisfied/ satisfied) with treatment at the end of treatment period (12 months)	575 per 1,000	724 per 1,000 (523 to 1,000)	RR 1.26 (0.91 to 1.75)	80 (1 study)	⊕⊕⊖⊖ Low[1,2]

CI: confidence interval; cOCP: combined oral contraceptive pill; RR: relative risk; MD: mean difference; VAS: visual analogue scale;
1 'Open label' study, subjects not blinded
2 CI crosses 1 default threshold
3 CI crosses 2 default thresholds
4 VAS Scale: 10cm scale where 0 = absence of pain, >0–5 = mild pain, >5–8 = moderate pain, >8–10 = unbearable pain

ISBN 978-1-4731-2661-9

Table 84: Summary clinical evidence profile, comparison 11: GnRH agonist + E/P pill versus E/P pill

Outcomes	Illustrative comparative risks* (95% CI)		Rela-tive effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	E/P pill	GnRH agonist + E/P pill			
Pain at 8 months during treatment period – Dysmenorrhoea (Scale: 10-point VAS)[3]	-	The mean pain at 8 months during treatment period – dysmenorrhoea in the intervention groups was 1.9 lower (2.54 to 1.26 lower)	MD -1.9 (-2.54 to -1.26)	101 (1 study)	⊕⊕⊕⊖ Moderate[1]
Pain at 8 months during treatment period – Non-menstrual pelvic pain (Scale: 10-point VAS)[3]	-	The mean pain at 8 months during treatment period – non-menstrual pelvic pain in the intervention groups was 2.5 lower (3 to 2 lower)	MD -2.5 (-3 to -2)	101 (1 study)	⊕⊕⊕⊖ Moderate[1]
Pain at the end of treatment period (12 months) – Dysmenorrhoea (Scale: 10-point VAS)[3]	-	The mean pain at the end of treatment period (12 months) – dysmenorrhoea in the intervention groups was 2.7 lower (3.34 to 2.06 lower)	MD -2.7 (-3.34 to -2.06)	95 (1 study)	⊕⊕⊕⊖ Moderate[1]
Pain at the end of treatment period (12 months) – Non-menstrual pelvic pain (Scale: 10-point VAS)[3]	-	The mean pain at the end of treatment period (12 months) – non-menstrual pelvic pain in the intervention groups was 0.8 higher (0.33 to 1.27 higher)	MD 0.8 (0.33 to 1.27)	95 (1 study)	⊕⊕⊖⊖ Low[1,2]

CI: confidence interval; GnRH: gonadotrophin releasing hormone; E/P: ethinylestradiol pill; MD: mean difference; VAS: visual analogue scale
1 No blinding of study participants, investigators or assessors reported
2 CI crosses 1 default threshold
3. VAS scale: 0 = the absence of pain, 10 = unbearable pain

ISBN 978-1-4731-2661-9

Table 85: Summary clinical evidence profile, comparison 12: GnRH agonist versus cOCP

Outcomes	Illustrative comparative risks (95% CI)		Rela-tive effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	cOCP	Goserelin			
Pain at the end of treatment period (6 months) – Dyspareunia (Scale: 10-point VAS)[4]	-	The mean pain at the end of treatment period (6 months) – dyspareunia in the intervention groups was 1.8 lower (3.4 to 0.2 lower)	MD -1.8 (-3.4 to -0.2)	44 (1 study)	⊕⊕⊖⊖ Low[1,2]
Pain at the end of treatment period (6 months) – Non-menstrual pelvic pain (Scale: 10-point VAS)[4]	-	The mean pain at the end of treatment period (6 months)–- non-menstrual pelvic pain in the intervention groups was 0.2 higher (1.11 lower to 1.51 higher)	MD 0.2 (-1.11 to 1.51)	50 (1 study)	⊕⊕⊖⊖ Low1,2
Pain at 6 months after treatment period – Dysmenorrhoea (Scale: 10-point VAS)[4]	-	The mean pain at 6 months after treatment period – dysmenorrhoea in the intervention groups was 0.1 higher (1.08 lower to 1.28 higher)	MD 0.1 (-1.08 to 1.28)	50 (1 study)	⊕⊖⊖⊖ Very low[1,3]
Pain at 6 months after treatment period – Dyspareunia (Scale: 10-point VAS)[4]	-	The mean pain at 6 months after treatment period - dyspareunia in the intervention groups was 0.4 lower (2.1 lower to 1.3 higher)	MD -0.4 (-2.1 to 1.3)	43 (1 study)	⊕⊕⊖⊖ Low[1,2]
Pain at 6 months after treatment period – Non-menstrual pelvic pain (Scale: 10-point VAS)[4]	-	The mean pain at 6 months after treatment period – non-menstrual pelvic pain in the intervention groups was 0.3 higher (1.25 lower to 1.85 higher)	MD -0.3 (-1.25 to 1.85)	50 (1 study)	⊕⊕⊖⊖ Low[1,2]

CI: confidence interval; cOCP: combined oral contraceptive pill; MD: mean difference; VAS: visual analogue scale
1 No blinding of participants, investigators or assessors reported
2 CI crosses 1 default threshold
3 CI crosses 2 default thresholds

ISBN 978-1-4731-2661-9

4 VAS scale: 0 = the absence of pain, 1–5 = mild pain, 6–7 = moderate pain, 8–10 = unbearable pain

11.1.3.3.3 Economic evidence

Cost effectiveness papers

Three studies were identified concerned with the cost-effectiveness of hormonal therapy in the treatment of endometriosis.

Lukac (2011a)

This paper refers to an analysis of the Slovakian AU19 trial on endometriosis-associated pelvic pain. It compares dienogest with Gonadotrophin Releasing Hormone agonists (GnRHa) over a period of 2 years. The source for costing data are "published price lists, clinical guidelines, product labels and expert opinion" and the source for QALY data is the SF-36 QoL instrument. The paper describes a Markov Chain model with a discount rate of 5% although it reports some data on the direct costs of these treatments with and without diagnostic laparoscopy.

The paper finds dienogest saves €506 and gains 0.002 QALYs relative to GnRHas. This indicates dienogest dominates GnRHa and would be considered cost-effective in any system. The authors include a cost-effectiveness acceptability curve (CEAC) implying that dienogest is cost-effective at a threshold of €18,000 / QALY (the Slovakian threshold, equivalent to around £15,000 / QALY) in 69% of cases

Lukac (2011b)

This paper appears to be a re-analysis of Lukac (2011a) as it refers to the same AU19 trial and finds similar results. The difference appears to be that this paper looks at a 5-year time horizon whereas the first paper looks at a 2-year time horizon. This paper finds a cost saving of €426 and a QALY gain of 0.069 QALYs, again indicating dienogest dominates GnRHas.

Bodner, Vale, Ratcliffe & Farrar (1996)

This paper refers to a subpopulation of 60 women with infertility taken from a full cohort of 273 enrolled in the Gynaecology Audit Project in Scotland (GAPS). It was intended principally to demonstrate a methodological point around using medical audit data to underpin economic evaluation, but was still considered relevant to include in this review as part of the audit data considered were costs and health outcomes. 35 women were treated with ‘expectant management’, 21 treated medically and 2 treated surgically (the remaining 2 women were on a surgical waiting list – it is not clear why these women were not included in the expectant management group).

The main outcome measure considered was fertility rates, but participants also completed an SF-36 QoL questionnaire. The source of cost data was NHS Reference Costs and estimates obtained by interviews with clinical managers. The time horizon was 6 months and the discount rate 6%.

The cost per patient alternative were £387.29 for expectant management, £645.02 for medical management and £1594.06 for surgical management. The SF-36 general health scores (and SDs) were an improvement of 61.0 (21.1) to 61.4 (29.9) in the medical group and a deterioration of 76.4 (18.2) to 75.3 (22.0) in the expectant management group. There were not enough women in the surgical group to report accurate scores. Neither of these changes would be considered statistically significant by any reasonable criteria, but – if they were significant – would represent an ICER of £17,200 indicating medical management is likely to be cost-effective compared to no treatment at the standard threshold of £20,000 / QALY – although it should be cautioned that the short follow up means that the effect of the

ISBN 978-1-4731-2661-9

(contraceptive) hormonal medical management on long-term QALYs may not have been properly accounted for.

Only 2 of the 60 women became pregnant by the end of the study, which is consistent with a view where endometriosis is highly damaging to fertility but does not give much analysable information about the cost-effectiveness of strategies to treat endometriosis-related infertility.

Cost only papers

Additionally, 5 studies were identified looking only at the costs of hormonal therapy. Since none of these papers were based on a UK perspective it was thought that conventional NHS costing sources were likely to be more relevant and so the Committee did not weight their evidence strongly in making a final recommendation, but Table 86 gives a high-level summary of the relevant information.

Table 86: Summary characteristics of cost-only studies excluded from review

Lead Author	Date	Country	Comparison A	Comparison B	Outcomes
Araujo	2011	Brazil	Goserelin acetate for those with confirmed deep endometriosis	Goserelin acetate for all with pelvic pain	Treating all US$1662 cheaper
Avxentyeva	2013	Russia	Triptorelin, Leuprorelin, Buserelin, Dydrogestero ne, Dienogest	N/A	Triprorelin = €1102 Leuprorelin = €1118 Buserelin = €340 Dydrogestero ne = €369 Dienogest = €295
Romero	2012	Columbia	12 months Dienogest	6 months GnRHa	Diogenest = US$986.16 vs. GnRH US$2855.57
Zalis'ka	2014	Ukraine	Dydogesteron e, Dienogest, Triptorelin	N/A	Dydogesteron e = US$345 Dienogest = US$1347 Triptorelin = US$1347
Zhao	1998	US	Nafarelin	Leuprolide	Nafarelin = US$2261 vs. Leuprolide US$3245

11.1.3.3.4 Economic model output

The cost of hormonal treatments can vary greatly depending on the dose required to achieve amenorrhea, the route of administration and any issues relating to unwanted side effects (perhaps the most important of which is infertility). Nevertheless it is known that there are a cluster of extremely cheap hormonal treatments (including the combined oral contraceptive pill) and a cluster of extremely high-cost treatments including dienogest and GnRHas.

ISBN 978-1-4731-2661-9

Owing to a lack of evidence on a number of these treatments, only 4 were included for analysis in the final model as other treatments were not suitable for inclusion in the NMA.

Table 87: Annual cost of 4 hormonal treatments included in the model

Treatment	Cost per 3 months[a]	Source
Combined oral contraceptive pill (as ethinylestradiol / gestodene tablet)	£19.31	Electronic Drug Tariff, January 2017[b]
Progestogen treatment (as Desogestrel)	£14.35	Electronic Drug Tariff, January 2017 [b]
Danazol	£86.63	Electronic Drug Tariff, January 2017 [b]
GnRHa (as Leuprorelin)	£236	Electronic Drug Tariff, January 2017 [b]

(a) The economic model uses 3-month cycles as the standard step in its Markov Chains. As hormonal treatments are typically given cyclically (for example, 21 days on followed by 7 days off) the 3-month cost reflects an average of the cost over this time.
(b) Including placebo-arm costs from NICE CG 173 Table F16 to account for, for example, increased GP visits not accounted for in Electronic Drug Tariff.
GnRHa: Gonadotropin-releasing hormone agonist

Note that there is a significant issue with the costing of the 2 more routine contraceptives, which is that some women take these contraceptives purely to prevent pregnancy. This means that the opportunity cost of the NHS prescribing these drugs to these women is zero, which is a consideration the Committee made when discussing whether there was a case to recommend the more expensive hormonal treatments.

Table 88: Cost and effectiveness of all non-dominated treatment strategies containing a hormonal treatment

Treatment	Cost	QALY	ICER	Pr. cost-effective vs. no treatment (£20k / QALY)	Pr. cost-effective vs. no treatment (£30k / QALY)
Empirical Diagnosis & No Treatment	£22,752.60	18.120	Base Case	N/A	N/A
Empirical Diagnosis & Combined Oral Contraceptive Pill	£15,845.16	18.283	-£42,434.80	96.7%	96.7%
Empirical Diagnosis & Danazol	£19,158.84	18.316	£98,467.20	92.3%	93.4%

QALY: Quality Adjusted Life-Year; ICER: Incremental Cost-Effectiveness Ratio

Hormonal treatments are both highly cost-effective on average and highly likely to be cost-effective vs. no treatment for any individual patient. This effect explains why Empirical Diagnosis & Danazol can have such a high ICER (£98,467) but also such a high probability of being cost-effective relative to no treatment. Another important point is how little difference there is between the combined oral contraceptive pill and Progestogen treatment – Progestogen treatment is fractionally cheaper based on the economic evidence and fractionally less effective based on the NMA, but patient-level analysis suggests that at £20,000 / QALY around 45% - 50% of patients offered the one treatment would actually have done better if offered the other. This indicates that the type of contraceptive might not be as

ISBN 978-1-4731-2661-9

important as the model implies as there is so little difference between them. This does not apply to GnRHas and Danazol, which are notably more expensive and only cost-effective at cost/QALY thresholds around one hundred thousand pounds (GnRHas are dominated by Danazol in this model, but if Danazol is removed the ICER for the most cost-effective GnRHa is £173,760).

The Committee discussed how this was entirely expected; hormonal treatments are known to be effective for endometriosis and known to be cheap and safe to prescribe, with few side-effects. The Committee also discussed how empirical diagnosis followed by hormonal treatment was extremely likely to be the most cost-effective strategy; the cheaper hormonal treatments are so cheap that even if the number of women presenting with endometriosis was small (and even if hormonal treatments had no effect on superficially similar conditions like dysmenorrhoea) that the cost of prescribing these drugs to otherwise healthy women was negligible.

It was expected that hormonal treatments are harmful for fertility. In actual fact the NMA suggested that progestogen treatment might improve fertility, but this is thought to be an inconsistency with the evidence underpinning the NMA and not reflective of the actual effects of progestogen treatment on fertility. As a result of this, no analysis has been conducted on the best hormonal treatment for preserving fertility.

However, in women who have both pain and infertility as a symptom of endometriosis, the effectiveness of hormonal treatment at controlling pain coupled with its low cost meant hormonal treatment was preferred at ICERs less than £13,027 / QALY, where it is replaced with surgical treatment with adjunct hormonal therapy.

11.1.3.3.5 Clinical evidence statements

Comparison 1: GnRH agonist versus no treatment

Pain

Very low quality evidence from 1 trial (n=35) found a clinically significant beneficial effect of GnRH agonist treatment (buserelin IN) compared with expectant management for dysmenorrhoea relief (measured using VAS) at 12 weeks after starting treatment.

Comparison 2: GnRH agonist versus placebo

Dysmenorrhoea

Moderate quality evidence from 1 trial (n=88) demonstrated a clinically significant beneficial effect of GnRH agonist treatment (leuprorelin IM depot) compared with placebo in the reduction of dysmenorrhoea (measured using VAS) at 12 weeks after starting treatment.

Pelvic pain

Moderate quality evidence from 1 trial (n=88) demonstrated a clinically significant beneficial effect of GnRH agonist treatment (leuprorelin IM depot) compared with placebo in the reduction of pelvic pain (measured using VAS) at 12 weeks after starting treatment.

Moderate quality evidence from 1 trial (n=46) found a clinically significant beneficial effect of GnRH agonist treatment (triptorelin IM depot) compared with placebo in the cessation of pelvic tenderness at 6 months after starting treatment.

Dyspareunia

Moderate quality evidence from 1 trial (n=88) demonstrated a clinically significant beneficial effect of GnRH agonist treatment (leuprorelin IM depot) compared with placebo in the reduction of deep dyspareunia (measured using VAS) at 12 weeks after starting treatment.

ISBN 978-1-4731-2661-9

Very low quality evidence from 1 trial (n=46) found a clinically significant difference between GnRH agonist treatment (triptorelin IM depot) and placebo in the cessation of pelvic tenderness at 6 months after starting treatment.

Comparison 3: Combined oral contraceptive pill versus placebo

Pain

Low and moderate quality evidence from 1 trial (n=96) found a clinically significant beneficial effect of treatment with a combined oral contraceptive compared with placebo for dysmenorrhoea (measured using VAS), but no clinically significant difference between treatments for non-menstrual pelvic pain score (measured using VAS) or induration.

Comparison 4: GnRH agonist versus danazol

Pain

Moderate quality evidence from 1 RCT (n=59) found no clinically significant difference between GnRH agonist treatment (nafarelin IN) compared with danazol for pelvic tenderness and pelvic induration at 3 months (during treatment period) and at the end of the 6 month treatment period.

Patient requiring surgery because of reappearance of symptoms and positive findings at pelvic examination

Moderate quality evidence from 1 RCT (n=62) reported no clinically significant difference between GnRH agonist treatment (buserelin IN) and danazol in the number of patients requiring surgery because of reappearance of symptoms and positive findings at pelvic examination at follow-up at least 12 months after treatment ended.

Quality of life

Low quality evidence from 1 RCT (n=169) found no statistically significant difference in quality of life (PGWBI and modified Nottingham Health Profile) between GnRH agonist (nafarelin IN) and danazol at the end of the 6 month treatment period. Clinical significance was not calculable as the data reported in the paper were descriptive.

Comparison 5: GnRH agonist versus levonorgestrel-releasing intrauterine system

Quality of life

Moderate quality evidence from 1 RCT (n=83) reported no clinically significant difference between GnRH agonist treatment (leuprolide IM) and levonorgestrel-releasing intrauterine system in quality of life (PGWBI) at the end of the 6 month treatment period.

Comparison 6: GnRH agonist versus DMPA-SC

Effect on daily activities

High to moderate quality evidence from 1 RCT (n=274) found no clinically significant difference between GnRH agonist treatment (leuprolide IM) and depot MPA (given by SC injection) regarding the mean number of hours of productivity lost at employment and housework at the end of the 6 month treatment period and at 18 months (12 months post-treatment).

ISBN 978-1-4731-2661-9

Comparison 7: GnRH agonist 1 + placebo versus GnRH agonist 2 + placebo

Pain

Low quality evidence from 1 RCT (n=192) found no clinical significant differences between GnRH agonist treatments (nafarelin 200mcg twice per day (BDS) IN and IM placebo compared with leuprolide depot 3.75mg IM plus IN placebo) for pelvic tenderness and pelvic induration at 6 months after the end of the treatment period.

Comparison 8: GnRH agonist + placebo versus progestin + placebo

Quality of life

Very low quality evidence from 1 RCT (n=48) reported no clinical significant differences between treatment with a GnRH agonist (nafarelin 200 µg IN BDS) and oral placebo compared with oral medroxyprogesterone (BDS 15 mg) and IN placebo in terms of overall quality of life (measured using Goldberg's general health and Nottingham Health Profile Questionnaire) at 6 months after the end of the treatment period. Results were poorly reported.

Effect on daily activities

Very low quality evidence from 1 trial (n=48) reported no clinical significant differences between treatment with a GnRH agonist (nafarelin 200 µg IN BDS) and oral placebo compared with oral medroxyprogesterone (BDS 15 mg) and IN placebo in terms of the effects on daily activities (measured using the Coping wheel, Inventory of Social Support and Interaction – ISSI and demands, control and support questionnaires) including sleep disturbances, anxiety-depression, household work, vacation life and leisure, sexual life, motivation, emotional balance and work activities (including psychological work demands, intellectual discretion at work, authority over decisions at work and social support) at 6 months after the end of the treatment period. Results were poorly reported.

Comparison 9: GnRH agonist + placebo versus danazol + placebo

Pain

Very low quality evidence from 1 RCT (n=49) found no clinically significant difference between GnRH agonist treatment (nafarelin 200mcg BDS -400mcg/d- IN) and oral placebo compared with oral danazol (200mg 3 times per day (TDS)) plus IN placebo for pelvic tenderness and pelvic induration at 6 months after the end of the treatment period.

Very low quality evidence from 1 RCT (n=96) found no clinically significant differences between GnRH agonist treatment (nafarelin 200mcg BDS -400mcg/d- IN) and oral placebo compared with danazol (200mg TDS) plus IN placebo for pelvic tenderness and pelvic induration at 12 months after the end of the treatment period.

Low quality evidence from 1 RCT (n=253) found no clinically significant difference between GnRH agonist treatment (leuprolide 3.75mg monthly IM) and oral placebo compared with oral danazol (800mg once daily) plus IM placebo for pelvic tenderness at 6 months after the end of the treatment period.

Comparison 10: Depot medroxyprogesterone acetate versus cOCP + danazol

Pain

Moderate quality evidence from 1 RCT (n=80) found a clinically significant beneficial effect of depot medroxyprogesterone acetate treatment compared with cOCP plus danazol for dysmenorrhoea at 6 months after starting treatment and at the end of the treatment period (at 12 months). Very low- to low-quality evidence from the same study reported no clinically

ISBN 978-1-4731-2661-9

significant difference between the 2 intervention groups for dyspareunia and non-menstrual pelvic pain at 6 months after starting treatment and at the end of the treatment period (at 12 months).

Patient satisfaction

Low quality evidence from the same RCT (n=80) reported no clinically significant difference between depot medroxyprogesterone acetate treatment compared with cOCP plus danazol regarding patient satisfaction with treatment (very satisfied/satisfied) at the end of the treatment period (at 12 months).

Comparison 11: GnRH agonist (triptorelin) + E/P pill versus E/P pill

Pain

One RCT (n=102) reported a clinically significant beneficial effect of GnRH agonist (triptorelin) + E/P pill (gestodene 0.75 mg/ethinylestradiol 0.03 mg) treatment compared with E/P pill (gestodene 0.75 mg/ethinylestradiol 0.03 mg) alone for dysmenorrhoea and non-menstrual pelvic pain at 8 months during the treatment period and for dysmenorrhoea at the end of the treatment period (at 12 months). Evidence was of low to moderate quality.

Low quality evidence from the same study found no clinically significant beneficial effect of E/P pill (gestodene 0.75 mg/ethinylestradiol 0.03 mg) compared with GnRH agonist (triptorelin) + E/P pill (gestodene 0.75 mg/ethinylestradiol 0.03 mg) treatment for non-menstrual pelvic pain at the end of treatment period (at 12 months).

Comparison 12: GnRH agonist (goserelin) versus cOCP

Pain

Low quality evidence from 1 RCT (n=57) demonstrated a clinically significant beneficial effect of GnRH agonist (goserelin) treatment compared with cOCP (0.02 mg ethinylestradiol and 0.15 mg desogestrel) for dyspareunia at the end of the treatment period (at 6 months). The same study reported no clinically significant difference between the 2 study arms for non-menstrual pelvic pain and dysmenorrhoea at the end of the treatment period (at 6 months) and for dyspareunia, non-menstrual pelvic pain and dysmenorrhoea at 6 months after the treatment period. Evidence was of very low to low quality.

11.1.3.4 Evidence to recommendations

11.1.3.4.1 Relative value placed on the outcomes considered

As pain relief is the primary reason for patients seeking treatment, this was the most critical outcome for the NMA, pairwise meta-analysis and pairwise comparison within this review. Health-related quality of life was also critical as this might be considered to give a more broad reflection of patient experience than pain relief alone, but data were only available for the pairwise comparison. Withdrawal due to adverse events and adherence to treatment were also critical outcomes as these reflected specific issues relating to the use of certain treatments and were addressed within the NMA and pairwise meta-analysis.

Rate of success, satisfaction with treatment, effect on daily activities and reduction in size and extent of endometriotic cysts were considered important outcomes as they were less clear indicators of effectiveness and were addressed within the pairwise comparison.

11.1.3.4.2 Consideration of clinical benefits and harms

The evidence from the NMA supported the use of hormonal treatments for pain relief in women with endometriosis and evidence from the pairwise comparison was broadly consistent with this, therefore the Committee used the NMA for most decision-making. The

ISBN 978-1-4731-2661-9

Committee agreed with the evidence and further highlighted that the benefit from hormonal treatments was due to their efficacy in stopping or reducing periods. There was a desire from the Committee to reduce the number of repeated operations for women with endometriosis, further supporting maintenance of pain relief using hormonal treatments wherever possible.

Although they chose not to be specific about recommending a particular hormonal treatment in the recommendations, they stated that the first-line hormonal treatment would generally be the oral combined contraceptive pill or progestogens as they have good efficacy and typically have side effects that women may find more tolerable. The evidence showed that cyclic use of the combined oral contraceptive pill is effective, but the Committee were also aware that continuous and tricycling (where three packets are taken in a row, followed by a pill free interval) use of the pill are used in clinical practice, and although evidence was not available on these regimens in the literature, the Committee have found in their experience that these were also effective with limited adverse events.

The Committee recommended that if first-line hormonal treatment was contraindicated or not tolerated, then women should be referred to a gynaecologist for possible further treatment which could include other hormonal treatments (for example, with a GnRH-a) or surgery. The Committee discussed the results of clinical effectiveness of other hormonal treatments such as GnRH-as and danazol. Even though highly effective, use of GnRH-as requires guidance from a specialist as the NMA showed that they had higher risk of withdrawal due to adverse events and the Committee identified them as having more serious adverse events (e.g. bone density changes). The Committee noted that GnRH-as are only licensed for a 6 month period and therefore require special considerations to ensure that women do not stay on this treatment indefinitely. They also discussed that to negate their adverse events add-back therapy using oestrogens, progestogens or both would usually be prescribed as well. The Committee's view was that women found the androgenic adverse events related to other hormonal treatments such as danazol in particular to be very unpleasant (e.g. voice alteration, hair growth). The Committee therefore decided not to be prescriptive about which treatment path to follow when first line treatment is not effective, not tolerated or is contraindicated and that clinical judgement was required to weigh up the benefits and harms of options that could be used.

Throughout the care pathway, the Committee stressed the importance of a full discussion with women of their symptoms and priorities with respect to pain and fertility and of the importance of the woman's choice. Such a discussion should also relieve any concerns over future fertility with regards to taking hormonal treatments, as their use was not considered to have any detrimental effect on subsequent fertility.

Adverse events were very varied across different types of hormonal treatments (androgenic, etc.) but were consistent within the classes of hormonal treatments. Overall the Committee highlighted that potential adverse events should be discussed with women alongside the potential benefit for pain relief.

There was no evidence to recommend whether to use or not use aromatase inhibitors, selective oestrogen receptor modulators (SERMs) or selective progestogen receptor modulators (SPRMs).

11.1.3.4.3 Consideration of economic benefits and harms

The Committee agreed with the output of the health economic model that hormonal treatment was likely to be the most cost-effective first-line treatment for endometriosis. Hormonal treatments are so effective that they can be prescribed without any confirmatory testing, although the Committee discussed how such testing might be useful anyway for reasons unrelated to symptom control (for example, to ensure that the lesions were not adhering to the bowel wall). There was some discussion about whether the more expensive classes of hormonal treatment (for example, GnRHas) were likely to give better results than the simple oral contraceptive, but health economic modelling demonstrates the gain would

ISBN 978-1-4731-2661-9

have to be far in excess of the uncertainty intervals of the NMA model in order for the treatments to be cost-effective.

The Committee discussed how the certainty of the finding of the model was not sufficient to recommend the combined pill over progestogen treatment, although the contraceptive pill generated slightly more QALYs on average; the Committee decided it was best to offer whichever cheap oral hormonal contraceptive the woman preferred, especially with reference to any treatment she might currently be taking.

The Committee believed that the result from the model indicating that progestogen treatment was likely to improve fertility was an artefact. In general, the Committee argued that as hormonal treatments have no plausible biological pathway to improving fertility they should not be recommended to women seeking to conceive on health economic grounds.

As many women will already be taking hormonal contraception for reasons unrelated to their endometriosis, it is difficult to estimate precisely the resource impact of these recommendations. Although the contraception itself carries a small cost, it is expected to displace unnecessary prescriptions of expensive treatments such as GnRHas and therefore the overall effect is of uncertain direction. Assuming the most expensive scenario for the NHS (all women with symptomatic endometriosis are prescribed hormonal treatment they would not otherwise have been taking) the total cost to the NHS is fractionally above the NHS threshold for high resource impact, so it is assumed with the fact that there is a pre-existing base of women taking the treatment that the net resource impact is not high.

11.1.3.4.4 Quality of the evidence

The quality of the evidence used to make recommendations on hormonal treatments for pain relief was generally moderate and was drawn from the NMA. Although the majority of studies were appropriately blinded, they rarely reported appropriate allocation concealment or details of the randomisation procedure. Several did not report measures of variability or uncertainty in their estimates, which meant that statistical imputation of missing information was needed. However, a variety of sensitivity of analyses were performed to test assumptions made during modelling and the results seemed robust. Studies were relatively consistent in their inclusion and exclusion criteria, which led to low inconsistency within the evidence.

However, the quality of the evidence was poorer when making recommendations on potential adverse events. Withdrawal from studies due to adverse events was relatively rare, giving very low precision to the analyses, and for the NMA some of the direct and indirect evidence did not agree, raising concerns as to the validity of this network and its use in decision-making.

11.1.3.4.5 Other considerations

One of the key considerations throughout treatment for pain relief in endometriosis is women's fertility. Fertility may be a strongly influencing factor in many women's treatment choices and a timely discussion on how different treatments will impact this is essential. The Committee suggested that a particular point to highlight in such a discussion is that although there can be a delay in return to fertility after stopping treatment with hormones (which might be a particular consideration for perimenopausal women), spontaneous pregnancy rates are not affected. .

The different treatment options recommended here are based on RCT evidence from a number of different studies, which was in agreement with the experience of the Committee. Recommendations on information provision and the pathway of care were developed primarily from Committee experience and opinion, supported in part by the literature.

ISBN 978-1-4731-2661-9

11.1.3.4.6 Key conclusions

The Committee concluded that women should be offered the oral combined contraceptive pill or progestogens as the first-line treatment for pain relief. However, if these were contraindicated or if women did not tolerate them, or found the treatments to be ineffective, they should be referred to a gynaecologist to discuss the alternative management options of hormonal treatment or laparoscopy. Throughout the process, the Committee stressed the importance of the woman's choice and of fully informing them about their options.

11.1.3.5 Recommendations

36. Explain to women with suspected or confirmed endometriosis that hormonal treatment for endometriosis can reduce pain and has no permanent negative effect on subsequent fertility.

37. Offer hormonal treatment (for example, the combined oral contraceptive pill or a progestogen)[a] to women with suspected, confirmed or recurrent endometriosis.

38. If initial hormonal treatment for endometriosis is not effective, not tolerated or is contraindicated, refer the woman to gynaecology service, specialist endometriosis service (endometriosis centres) or paediatric and adolescent gynaecology service for investigation and treatment options.

11.2 Non-pharmacological management

Review question: What is the effectiveness of non-pharmacological therapies (for example, acupuncture) for managing pain associated with endometriosis?

11.2.1 Introduction

The symptoms associated with endometriosis differ with each woman; however, pain is almost always a factor, whether it be pelvic pain, painful periods, pain on intercourse, pain on urination or on defecation.

The level of pain experienced does not always relate to the extent of the disease and minor disease can be as or more painful than severe disease. It is often related to the location of the disease.

For many women treatment will involve a combination of therapies given over their lifetime depending on their circumstances at any given time. The aim of any management is primarily to reduce symptoms and maintain or improve quality of life.

There are many reasons why women may choose to use non-pharmacological therapies, for example, being offered counselling or acupuncture as alternatives or adjuncts to medical and surgical management.

In addition to reduction in pain, these therapies may be chosen to enable the woman to feel she is taking an active role in the treatment of her symptoms. Women who use self-management strategies may report regaining control over their lives and feel less dependent on healthcare professionals.

a At the time of publication (September 2017), not all combined oral contraceptive pills or progestogens have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Prescribing guidance: prescribing unlicensed medicines for further information.

ISBN 978-1-4731-2661-9

Some women have exhausted all possible hormonal and medical treatments and have discontinued these due to intolerable side effects or found them to be ineffective and are keen to seek further alternative or additional solutions for their pain. They may report reduction in medication use and potentially therefore in side effects.

Women who are trying to conceive may decide to postpone treatment for a certain time period with the hope of a resulting pregnancy. While trying to become pregnant she may still be experiencing painful symptoms but would be unable to use medical treatments during this time as most can be harmful to the developing fetus. Women often give up trying to conceive as their pain is intolerable therefore non-pharmacological therapies that have been shown to be safe for use in early pregnancy may be chosen to help them continue in their desire for a pregnancy.

The aim of this review is to determine the clinical and cost effectiveness of non-pharmacological therapies in reducing pain in women with endometriosis or suspected endometriosis

For full details, see the review protocol in Appendix D, the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G.

11.2.2 Description of clinical evidence

Ten studies were included in the evidence review (Chen 2012, de Sousa 2016, Flower 2011, Mira 2012, Sesti 2009, Wayne 2008; Wu 2006 (Flower 2012); Xia 2006; Xiang 2002; Zhu 2014). Nine of these were RCTs and the tenth was a Cochrane systematic review that provided data on 1 further RCT (Flower 2012) (Table 89).

Five RCTs were conducted in China (Chen 2012, Wu 2006 (Flower 2012), Xia 2006, Xiang 2002, Zhu 2014). Two RCTs were from Europe – 1 from the UK (Flower 2011) and 1 from Italy (Sesti 2009). Two RCTs were conducted in Brazil (de Sousa 2016, Mira 2015) and 1 in the USA (Wayne 2008).

Much of the evidence came from small RCTs and sample sizes ranged from 18 (Wayne 2008) to 259 (Sesti 2009).

The severity or stage of endometriosis was not described in many of the articles. However, 1 RCT specifically included women with deep endometriosis who were suffering from persisting pelvic pain and dyspareunia, despite hormonal therapy (Mira 2015). One RCT included women with subfertility and minimal/mild endometriosis, all of whom underwent operative laparoscopy at the start of the trial (Zhu 2014). A third RCT only recruited women with an endometrioma, who underwent cystectomy at the start of the trial (Sesti 2009).

The majority of RCTs focused on outcomes of pain relief and health-related quality of life. Two RCTs reported on reduction in the size or recurrence of endometriomas (Wu 2006 (Flower 2012); Sesti 2009). One reported on fertility outcomes (live birth and miscarriage rates) (Zhu 2014).

Five RCTs investigated the use of different forms of acupuncture for endometriosis. Two RCTs compared acupuncture to sham acupuncture (de Sousa 2016, Wayne 2008). One RCT compared the use of acupuncture with danazol (Chen 2012) and another compared acupuncture plus Chinese herbal medicine (CHM) to danazol (Xia 2006). One RCT compared acupuncture to CHM (Xiang 2002).

Three further RCTs considered the use of CHM. One compared the use of individualised CHM preparations to placebo (Flower 2011). The Cochrane review included 1 RCT including 3 treatment groups: Nei Yi tablets; Nei Yi tablets and enemas; and danazol (Wu 2006 (Flower 2012)). The third RCT assessed fertility rates in women given short-term CHM plus

ISBN 978-1-4731-2661-9

the combined oral contraceptive pill (cOCP) after surgery for endometriosis, compared to women given cOCP alone, or no treatment (Zhu 2014).

A single RCT compared acupuncture-like transcutaneous electrical nerve stimulation (TENS) to self-applied TENS (Mira 2015)

Finally, 1 RCT compared dietary therapy (a nutritional supplement of vitamins, minerals, fatty acids and probiotics) with placebo, GnRH analogues or cOCP in prevention of endometrioma recurrence after cystectomy (Sesti 2009).

Evidence for 2 critical outcomes was available (relief of endometriosis-related pain and health-related quality of life). Evidence for 2 important outcomes was also available (fertility and reduction in size of endometriotic cysts). Some evidence was available on activities of daily living. No evidence was available for the remaining outcomes (improvement of endometriosis symptoms other than pain, adverse events resulting from the intervention and adherence to the treatment programme).

11.2.3 Summary of included studies

A summary of the studies that were included in this review are presented in Table 89.

Table 89: Summary of included studies

Study	Intervention/comparison	Population	Outcomes	Comments
Chen 2012 China	Intervention: Abdominal acupuncture, administered prior to and during menses for 3 consecutive menstrual cycles. Acupuncture was performed approximately 7 times during each treatment cycle Comparison: Danazol, 200mg twice daily was administered (starting on day 1 of a menstrual cycle), for 3 consecutive cycles	Severity of endometriosis symptoms: • severe 30% • moderate 43% • mild 27%. Scoring was based on a variety of symptoms, including severity of pain, relief with common analgesics, associated symptoms (e.g. nausea and vomiting, sweating) N=70	Assessed at 6 months (3 months of treatment, then 3 months without treatment). Cure of symptoms – defined as complete relief of pain and other symptoms after medication and no relapse in the next 3 menstrual cycles	
de Sousa 2016 Brazil	Intervention: experimental treatment of acupuncture, 5 sessions of acupuncture in 5 weeks Comparison: sham-acupuncture, 5 sessions of acupuncture in 5 weeks	Mean age of 30.81 years (±6.38). These women were suffering from endometriosis for an average of 4.79 years (±2.48). No other information given. N=42	All outcomes measured at 2 months (follow up). Pain score (VAS of 0–10) for: • chronic pelvic pain • dyspareunia	
Flower 2011 UK	Intervention: an individualised CHM decoction was administered twice daily for 16 weeks Comparison:	15% of participants were using concomitant medical therapy (not described). N=33	Assessed at the end of treatment (16 weeks) Pain scores, measured with VAS 0–10:	

ISBN 978-1-4731-2661-9

Study	Intervention/comparison	Population	Outcomes	Comments
	a placebo decoction comprising inactive ingredients was administered twice daily for 16 weeks		• period pain • pain during intercourse • pain on bowel movement • daily pain. MYMOP scores (measured with 7-point Likert scale) to assess change in symptoms, well-being and limitation of activity. Endometriosis Health Profile-30 scores (range 0–100)	
Mira 2015 Brazil	Intervention: acupuncture-like TENS to S3-S4 region, 30 minute sessions were applied once a week for a period of 8 weeks Comparison: self-applied TENS to the S3-S4 region, 20 minute sessions were conducted twice daily for a period of 8 weeks	Women with deep endometriosis diagnosed in the cul-de-sac and intestinal loop who sustained pelvic pain and/or deep dyspareunia, despite continuous clinical medication. All women were undergoing hormone therapy with continuous progestin alone or combined oral contraceptives for at least 3 months. N=22	Assessed at the end of treatment (8 weeks). Endometriosis Health Profile-30 scores (range 0–100)	Women were undergoing hormone therapy with continuous progestin alone or combined contraceptives for at least 3 months
Sesti 2009 Italy	Intervention: Group 1: GnRH analogue 3.75mg every 28 days post-operatively Group 2: continuous low-dose cOCP for 6 months post-operatively Group 3: dietary intervention (including probiotics, vitamin, mineral and fatty acid supplementation) post-operatively Comparison: placebo	All women underwent laparoscopic cystectomy for endometrioma at the start of the trial. N=259	Assessed at 18 months follow up: recurrence of endometrioma >20mm diameter. Cysts suspected to be endometriomas with ultrasound were then confirmed laparoscopically	Women with an endometrioma, who underwent cystectomy at the start of the trial.
Wayne 2008 USA	Intervention: Japanese-style acupuncture, twice per week for 8 consecutive weeks Comparison:	All women had stage I endometriosis. Eligible participants were aged 13–22 years old. N=18	All outcomes measured at 4 weeks, 8 weeks (during treatment) and 6 months (follow up).	

ISBN 978-1-4731-2661-9

Study	Intervention/comparison	Population	Outcomes	Comments
	sham-acupuncture, twice per week for 8 consecutive weeks		Pain score (numerical analogue scale of 0–10). Endometriosis Health Profile-30 (range 0–100). Pediatric Quality of Life Inventory score (range 0–100). Activity scale (to assess activities limited by endometriosis) (range 0–10)	
Wu 2006 (Flower 2012 CSR) China	Intervention: CHM Group 1: Nei Yi pills 10g twice daily Group 2: Nei Yi pills 10g twice daily plus Nei Yi enema 70ml once daily Comparison: danazol 400mg per day	Laparoscopically confirmed endometriosis. No other details given N=58	Symptomatic relief was assessed within 3 years of stopping treatment, other outcomes – at the end of 3 months treatment. Five outcomes were assessed: • symptomatic relief (defined as disappearance of symptoms, pelvic mass or pregnancy within 3 years for those with infertility) • dysmenorrhoea score (range not reported) • lumbosacral pain relief (dichotomous outcome) • tenderness of vaginal nodules in posterior fornix (dichotomous outcome) • disappearance or shrinkage of adnexal masses (criteria not reported)	
Xia 2006 China	Intervention: acupuncture (started 9 days before menses and discontinued during menses) and CHM (Gui-Zhi-Fu-Ling-Wan) Comparison:	N=78	Assessed at the end of treatment (3 months of treatment). Dysmenorrhoea (pain scale not reported).	

ISBN 978-1-4731-2661-9

Study	Intervention/comparison	Population	Outcomes	Comments
	danazol 200mg twice daily. Treatment was continued for 3 consecutive cycles		Dichotomous outcome of 'cessation of signs and symptoms' of • lumbosacral pain • dyspareunia	
Xiang 2002 China	Intervention: Ear acupuncture therapy, beginning 5 days before menses and given fo4ur times every other day, for 3 menstrual cycles Comparison: CHM. A decoction was given 5 days before menstruation, 1 dose for 7 days, for 3 menstrual cycles	Laparoscopically confirmed endometriosis. No other information given. N = 67	Assessed at the end of treatment (3 menstrual cycles). Dysmenorrhoea score (5–15). Symptom cure (dichotomous outcome)	
Zhu 2014 China	Intervention: Group 1: cOCP (30µg ethinyloestradiol and 150µg desogestrel) administered once per day for 63 days after surgery Group 2: as group 1, but also received Dan'e CHM 30g per day for the latter 30 days of treatment. Comparison: no medical treatment after surgery	Women with minimal/mild endometriosis (wishing to conceive), who had failed to become pregnant after at least 12 months of unprotected intercourse. All women underwent surgery at the start of the trial, including ablation/excision of all visible lesions and division of adhesions to restore normal pelvic anatomy N=156	Fertility outcomes assessed at 12 months after treatment: • live birth rate • miscarriage rate	Women with subfertility and minimal/mild endometriosis, all of whom underwent operative laparoscopy at the start of the trial

N: number of participants in study; CSR: Cochrane systematic review

11.2.4 Clinical evidence profile

The clinical evidence profiles for this review question are presented in Table 90 to Table 103.

Table 90: Summary clinical evidence profile, Comparison 1: cOCP and Dan'e compared to no treatment for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Correspon-ding risk			
	No treatment	cOCP and Dan'e			
Live birth (denominator pregnancy) at 12 months after treatment completion	792 per 1,000	815 per 1,000	RR 1.03 (0.75 to 1.4)	40 (1 study)	⨁⨁⊖⊖ Low[1]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Correspon-ding risk			
		(594 to 1,000)			
Miscarriage (denominator pregnancy) at 12 months after treatment completion	125 per 1,000	188 per 1,000 (43 to 815)	RR 1.5 (0.34 to 6.52)	40 (1 study)	⊕⊕⊖⊖ Low[1]

CI: confidence interval; RR: risk ratio; cOCP: combined oral contraceptive pill
1 CI for estimate is very wide, crossing 2 thresholds

Table 91: Summary clinical evidence profile, Comparison 2: cOCP and Dan'e compared to cOCP for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	cOCP	cOCP and Dan'e			
Live birth (denominator pregnancy) at 12 months after treatment completion	700 per 1,000	812 per 1,000 (560 to 1,000)	RR 1.16 (0.8 to 1.68)	36 (1 study)	⊕⊕⊕⊖ Moderate[1]
Miscarriage (denominator pregnancy) at 12 months after treatment completion	200 per 1,000	188 per 1,000 (48 to 720)	RR 0.94 (0.24 to 3.6)	36 (1 study)	⊕⊕⊖⊖ Low[2]

CI: confidence interval; RR: risk ratio; cOCP: combined oral contraceptive pill
1 CI for estimate is very wide, crossing 1 threshold
2 CI for estimate is very wide, crossing 2 thresholds

Table 92: Summary clinical evidence profile, Comparison 3: Diet compared to placebo for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Placebo	Diet			
Endometrioma recurrence[1]	167 per 1,000	177 per 1,000 (82 to 387)	RR 1.06 (0.49 to 2.32)	122 (1 study)	⊕⊕⊖⊖ Low[2]

CI: confidence interval; RR: risk ratio
1 The recurrence of endometrioma was defined as the presence of a cyst, detected by transvaginal ultrasonography, with a pattern suggesting an endometrioma more than 20 mm in diameter.
2 CI for estimate is very wide, crossing 2 thresholds

Table 93: Summary clinical evidence profile, Comparison 4: Diet compared to GnRHa for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	GnRHa	Diet			

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
Endometrioma recurrence[1]	103 per 1,000	178 per 1,000 (70 to 449)	RR 1.72 (0.68 to 4.34)	120 (1 study)	⊕⊕⊖⊖ Low[2]

CI: confidence interval; RR: risk ratio
1 The recurrence of endometrioma was defined as the presence of a cyst, detected by transvaginal ultrasonography, with a pattern suggesting an endometrioma more than 20 mm in diameter.
2 CI for estimate is very wide, crossing 2 thresholds

Table 94: Summary clinical evidence profile, Comparison 5: Diet compared to cOCP for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	cOCP	Diet			
Endometrioma recurrence[1]	150 per 1,000	177 per 1,000 (79 to 398)	RR 1.18 (0.53 to 2.65)	122 (1 study)	⊕⊕⊖⊖ Low[2]

CI: confidence interval; RR: risk ratio; cOCP: combined oral contraceptive pill
1 The recurrence of endometrioma was defined as the presence of cyst, detected by transvaginal ultrasonography, with a pattern suggesting an endometrioma more than 20 mm in diameter.
2 Confidence interval for estimate is very wide crossing 2 thresholds

Table 95: Summary clinical evidence profile, Comparison 6: Acupuncture compared to sham acupuncture for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Sham acupuncture	Acupuncture			
Change (from baseline) in pain in last 4 weeks – at 4 weeks (ESSS)	-	The mean change (from baseline) in pain in last 4 weeks – at 4 weeks (ESSS) in the intervention groups was 3.4 lower (5.82 to 0.98 lower)	MD -3.4 (-5.82 to -0.98)	14 (1 study)	⊕⊕⊖⊖ Low[3,4]
Change (from baseline) in pain in last 4 weeks – at 8 weeks (ESSS)	-	The mean change (from baseline) in pain in last 4 weeks – at 8 weeks (ESSS) in the intervention groups was 0.5 lower (3.22 lower to 2.22 higher)	MD -0.5 (-3.22 to 2.22)	15 (1 study)	⊕⊖⊖⊖ Very low[3,4]
Change (from baseline) in pain in last 2 months–	-	The mean change (from baseline) in last 2 months – chronic pelvic pain in the	MD -3.29 (-3.97 to -2.61)	42 (1 study)	⊕⊕⊕⊖ Moderate[6]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
chronic pelvic pain		intervention groups was 3.29 lower (3.97 to 2.61 lower)			
Change (from baseline) in pain in last 2 months – dyspareunia	-	The mean change (from baseline) in last 2 months – dyspareunia in the intervention groups was 3.76 lower (4.55 to 2.97 lower)	MD -3.76 (-4.55 to -2.97)	42 (1 study)	⊕⊕⊕⊖ Moderate[6]
Change (from baseline) in pain in last 4 weeks – at 6 months (ESSS)	-	The mean change (from baseline) in pain in last 4 weeks–- at 6 months (ESSS) in the intervention groups was 0.8 lower (4.66 lower to 3.06 higher)	MD -0.8 (-4.66 to 3.06)	14 (1 study)	⊕⊖⊖⊖ Very low[3,5]
Change (from baseline) in QoL (EHP total score) - at 4 weeks (EHP)	-	The mean change (from baseline) in QoL (EHP total score)–- at 4 weeks (EHP in the intervention groups was 21.5 lower (39.27 to 3.73 lower)	MD -21.5 (-39.27 to -3.73)	14 (1 study)	⊕⊕⊖⊖ Low[3,4]
Change (from baseline) in QoL (EHP total score) - at 8 weeks (EHP)	-	The mean change (from baseline) in QoL (EHP total score) – at 8 weeks (EHP) in the intervention groups was 19.7 lower (38.7 to 0.7 lower)	MD -19.7 (-38.7 to -0.7)	15 (1 study)	⊕⊕⊖⊖ Low[3,4]
Change (from baseline) in QoL (EHP Total score) – at 6 months (EHP)	-	The mean change (from baseline) in QoL (EHP total score) - at 6 months (EHP) in the intervention groups was 20.9 lower (37.57 to 4.23 lower)	MD -20.9 (-37.57 to -4.23)	14 (1 study)	⊕⊕⊖⊖ Low[3,4]
Change (from baseline) in QoL (Paediatric QoL Inventory total score)[1] – at 4 weeks	-	The mean change (from baseline) in QoL (paediatric QoL inventory total score) – at 4 weeks in the intervention groups was 10.1 higher	MD 10.1 (-3.26 to 23.46)	14 (1 study)	⊕⊕⊖⊖ Low[3,4]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
		(3.26 lower to 23.46 higher)			
Change (from baseline) in QoL (Paediatric QoL Inventory total score)[1]– at 8 weeks	-	The mean change (from baseline) in QoL (paediatric QoL inventory total score) – at 8 weeks in the intervention groups was 14.2 higher (0.94 lower to 29.34 higher)	MD 14.2 (-0.94 to 29.34)	15 (1 study)	⊕⊕⊖⊖ Low[3,4]
Change (from baseline) in QoL (Paediatric QoL Inventory total score)[1] – at 6 months	-	The mean change (from baseline) in QoL (paediatric QoL inventory total score) – at 6 months in the intervention groups was 14.9 higher (1.18 to 28.62 higher)	MD 14.9 (1.18 to 28.62)	14 (1 study)	⊕⊕⊖⊖ Low[3,4]
Change (from baseline) in activities of daily living (3 activity score)[2] – at 4 weeks	-	The mean change (from baseline) in activities of daily living (3 activity score) – at 4 weeks in the intervention groups was 2.9 lower (4.85 to 0.95 lower)	MD -2.9 (-4.85 to -0.95)	14 (1 study)	⊕⊖⊖⊖ Very low[3,5]
Change (from baseline) in activities of daily living (3 activity score)[2] – at 8 weeks	-	The mean change (from baseline) in activities of daily living (3 activity score) – at 8 weeks in the intervention groups was 1.8 lower (4.48 lower to 0.88 higher)	MD -1.8 (-4.48 to 0.88)	14 (1 study)	⊕⊕⊖⊖ Low[3,4]
Change (from baseline) in activities of daily living (3 activity score)[2] – at 6 months	-	The mean change (from baseline) in activities of daily living (3 activity score) – at 6 months in the intervention groups was 1.7 lower (5.21 lower to 1.81 higher)	MD -1.7 (-5.21 to 1.81)	14 (1 study)	⊕⊖⊖⊖ Very low[3,5]

CI: confidence interval; MD: mean difference; ESSS: Endometriosis Symptom Severity Scale (0–10); EHP: Endometriosis Health Profile-30 (subscales range 0–100)
1 Paediatric QoL Inventory Total score (subscales range 0–100)
2 Activity scale scores range 0–10
3 Due to dropouts
4 CI for estimate is very wide, crossing 1 threshold

ISBN 978-1-4731-2661-9

5 CI for estimate is very wide, crossing 2 thresholds
6 The quality of the evidence was downgraded because of the unclear risk of attrition bias (no details provided in the text), besides the unclear risk of detection bias

Table 96: Summary clinical evidence profile, Comparison 7: Acupuncture compared to danazol for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Danazol	Acupuncture			
Cure of symptoms[1]	143 per 1,000	86 per 1,000 (23 to 331)	RR 0.6 (0.16 to 2.32)	70 (1 study)	⊕⊖⊖⊖ Very low[2,3]

CI: confidence interval; RR: risk ratio
1 Defined as complete relief of pain and other symptoms after medication and no relapse in the next 3 menstrual cycles
2 No blinding
3 CI for estimate is very wide, crossing 2 thresholds

Table 97: Summary clinical evidence profile, Comparison 8: Acupuncture compared to CHM for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	CHM	Acupuncture			
Dysmenorrhoea	-	The mean dysmenorrhoea in the intervention groups was 4.81 lower (6.25 to 3.37 lower)	MD -4.81 (-6.25 to -3.37)	67 (1 study)	⊕⊖⊖⊖ Very low[2,3]
Cure of symptoms[1]	100 per 1,000	297 per 1,000 (91 to 970)	RR 2.97 (0.91 to 9.7)	67 (1 study)	⊕⊕⊖⊖ Low[2,4]

CI: confidence interval; RR: risk ratio
1 Defined according Guideline for Clinical Research on New Chinese Drugs for Treatment of Pelvic Endometriosis
2 No blinding
3 CI for estimate is very wide, crossing 2 thresholds
4 CI for estimate is very wide crossing 1 threshold

Table 98: Summary clinical evidence profile, Comparison 9: CHM compared to placebo for endometriosis

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Placebo	CHM			
Change (from baseline) in pain (VAS) at week 16 – period pain	-	The mean change (from baseline) in pain (VAS) at week 16 – period pain in the intervention groups was 1.22 lower	MD -1.22 (-3.81 to 1.37)	12 (1 study)	⊕⊖⊖⊖ Very low[1,2]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
		(3.81 lower to 1.37 higher)			
Change (from baseline) in pain (VAS) at week 16 – pain during sex	-	The mean change (from baseline) in pain (VAS) at week 16 – pain during sex in the intervention groups was 0.76 higher (1.53 lower to 3.05 higher)	MD 0.76 (-1.53 to 3.05)	8 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Change (from baseline) in pain (VAS) at week 16– pain on bowel movement	-	The mean change (from baseline) in pain (VAS) at week 16 – pain on bowel movement in the intervention groups was 0.08 higher (2.87 lower to 3.03 higher)	MD 0.08 (-2.87 to 3.03)	12 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Change (from baseline) in pain (VAS) at week 16– daily pain	-	The mean change (from baseline) in pain (vas) at week 16 – daily pain in the intervention groups was 0.74 higher (1.81 lower to 3.29 higher)	MD 0.74 (-1.81 to 3.29)	13 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Change (from baseline) in patient assessed QoL (MYMOP) at week 16 – symptom 1	-	The mean change (from baseline) in patient assessed QoL (MYMOP) at week 16 – symptom 1 in the intervention groups was 0.58 lower (2.41 lower to 1.25 higher)	MD -0.58 (-2.41 to 1.25)	18 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Change (from baseline) in patient assessed QoL (MYMOP) at week 16 – symptom 2	-	The mean change (from baseline) in patient assessed QoL (MYMOP) at week 16–- symptom 2 in the intervention groups was 0.9 lower (2.68 lower to 0.88 higher)	MD -0.9 (-2.68 to 0.88)	18 (1 study)	⊕⊕⊖⊖ Low[1,3]
Change (from baseline) in patient assessed QoL (MYMOP) at week 16 – activity	-	The mean change (from baseline) in patient assessed QoL (MYMOP) at week 16 – activity in the	MD -0.69 (-2.31 to 0.93)	17 (1 study)	⊕⊖⊖⊖ Very low[1,2]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
		intervention groups was 0.69 lower (2.31 lower to 0.93 higher)			
Change (from baseline) in patient assessed QoL (MYMOP) at week 16 – wellbeing	-	The mean change (from baseline) in patient assessed QoL (MYMOP) at week 16– wellbeing in the intervention groups was 1.06 lower (2.95 lower to 0.83 higher)	MD -1.06 (-2.95 to 0.83)	17 (1 study)	⊕⊕⊖⊖ low[1,3]
Change (from baseline) in QoL (EHP 30) at week 16 – pain	-	The mean change (from baseline) in QoL (EHP-30) at week 16 – pain in the intervention groups was 0.32 lower (10.01 lower to 9.37 higher)	MD -0.32 (-10.01 to 9.37)	18 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Change (from baseline) in QoL (EHP 30) at week 16 – control and powerlessness	-	The mean change (from baseline) in QoL (EHP-30) at week 16–- control and powerlessness in the intervention groups was 1.73 lower (7.35 lower to 3.89 higher)	MD -1.73 (-7.35 to 3.89)	18 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Change (from baseline) in QoL (EHP 30) at week 16 – emotional wellbeing	-	The mean change (from baseline) in QoL (EHP-30) at week 16 – emotional wellbeing in the intervention groups was 0.37 lower (4.38 lower to 3.64 higher)	MD -0.37 (-4.38 to 3.64)	18 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Change (from baseline) in QoL (EHP 30) at week 16 – social support	-	The mean change (from baseline) in QoL (EHP-30) at week 16 – social support in the intervention groups was 2.71 lower (7.09 lower to 1.67 higher)	MD -2.71 (-7.09 to 1.67)	18 (1 study)	⊕⊕⊖⊖ Low[1,3]
Change (from baseline) in QoL (EHP 30) at week 16 – self-image	-	The mean change (from baseline) in QoL (EHP-30) at week 16 – self-image in the	MD 0.46 (-2.22 to 3.14)	18 (1 study)	⊕⊖⊖⊖ Very low[1,2]

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
		intervention groups was 0.46 higher (2.22 lower to 3.14 higher)			

CI: confidence interval; MD: mean difference; MYMOP: Measure Your own Medical Outcomes Profile (1–7-point Likert scale); QoL: quality of life
1 Due to drop outs
2 CI for estimate is very wide, crossing 2 thresholds
3 CI for estimate is very wide, crossing 1 threshold

Table 99: Summary clinical evidence profile, Comparison 10: CHM (oral) compared to danazol for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Danazol	CHM (oral)			
Symptomatic relief[1]	111 per 1,000	562 per 1,000 (142 to 1,000)	RR 5.06 (1.28 to 20.05)	34 (1 study)	⊕⊕⊖⊖ low[2,3]
Dysmenorrhoea score	-	The mean dysmenorrhoea score in the intervention groups was 1.01 lower (3.11 lower to 1.09 higher)	MD -1.01 (-3.11 to 1.09)	34 (1 study)	⊕⊕⊖⊖ Low[2,4]
Lumbosacral pain relief	722 per 1,000	874 per 1,000 (621 to 1,000)	RR 1.21 (0.86 to 1.7)	34 (1 study)	⊕⊕⊖⊖ Low[2,4]
Rectal irritation relief	500 per 1,000	835 per 1,000 (450 to 1,000)	RR 1.67 (0.9 to 3.1)	24 (1 study)	⊕⊕⊖⊖ Low[2,4]
Tenderness of vaginal nodules in posterior fornix	692 per 1,000	907 per 1,000 (602 to 1,000)	RR 1.31 (0.87 to 1.97)	24 (1 study)	⊕⊕⊖⊖ Low[2,4]
Adnexal masses disappearance or shrinkage	533 per 1,000	752 per 1,000 (421 to 1,000)	RR 1.41 (0.79 to 2.5)	27 (1 study)	⊕⊖⊖⊖ Very low[2,5]

CI: confidence interval; RR: risk ratio; MD: mean difference
1 Defined as a complete resolution of all symptoms and signs and included pregnancy, when desired, within 3 years of stopping treatment
2 Not clear if blinding was performed
3 Although the outcome is defined, it is wide, encompassing different symptoms and signs.
4 CI for estimate is very wide, crossing o1ne threshold
5 CI for estimate is very wide, crossing 2 thresholds

Table 100: Summary clinical evidence profile, Comparison 11: CHM (oral + enema) compared to danazol for endometriosis

Outcomes	Illustrative comparative risks (95% CI)			Quality of the

ISBN 978-1-4731-2661-9

	Assumed risk	Corresponding risk	Relative effect (95% CI)	No of participants (studies)	evidence (GRADE)
	Danazol	CHM (oral + enema)			
Symptomatic relief[1]	111 per 1,000	624 per 1,000 (163 to 1,000)	RR 5.63 (1.47 to 21.54)	42 (1 study)	⊕⊕⊖⊖ Low[2,3]
Dysmenorrhoea score	-	The mean dysmenorrhoea score in the intervention groups was 2.9 lower (4.55 to 1.25 lower)	MD -2.9 (-4.5 to -1.25)	42 (1 study)	⊕⊕⊖⊖ Low[2,4]
Lumbosacral pain relief	722 per 1,000	831 per 1,000 (592 to 1,000)	RR 1.15 (0.82 to 1.62)	42 (1 study)	⊕⊕⊖⊖ Low[2,4]
Rectal irritation relief	500 per 1,000	890 per 1,000 (495 to 1,000)	RR 1.78 (0.99 to 3.2)	30 (1 study)	⊕⊕⊖⊖ Low[2,4]
Tenderness of vaginal nodules in posterior fornix	692 per 1,000	872 per 1,000 (582 to 1,000)	RR 1.26 (0.84 to 1.9)	29 (1 study)	⊕⊕⊖⊖ Low[2,4]
Adnexal masses disappearance or shrinkage	533 per 1,000	907 per 1,000 (555 to 1,000)	RR 1.7 (1.04 to 2.78)	36 (1 study)	⊕⊕⊖⊖ Low[2,4]

CI: confidence interval; RR: risk ratio; MD: mean difference; CSR: Cochrane systematic review
1 Defined as a complete resolution of all symptoms and signs and included pregnancy, when desired, within 3 years of stopping treatment
2 Not clear if blinding was performed
3 Although the outcome is defined, it is wide, encompassing different symptoms and signs.
4 CI for estimate is very wide, crossing 1 threshold

Table 101: Summary clinical evidence profile, Comparison 12: CHM (oral+ enema) compared to CHM (oral) for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	CHM (oral)	CHM (oral+ enema)			
Symptomatic relief[1]	562 per 1,000	624 per 1,000 (366 to 1,000)	RR 1.11 (0.65 to 1.89)	40 (1 study)	⊕⊖⊖⊖ Very low[2,3,4]
Dysmenorrhoea score	-	The mean dysmenorrhoea score in the intervention	MD -1.89 (-	40 (1 study)	⊕⊕⊖⊖ Low[2,5]

ISBN 978-1-4731-2661-9

		groups was 1.89 lower (3.89 lower to 0.11 higher)	3.89 to 0.11)		
Lumbosacral pain relief	875 per 1,000	831 per 1,000 (648 to 1,000)	RR 0.95 (0.74 to 1.23)	40 (1 study)	⊕⊕⊖⊖ Low[2,5]
Rectal irritation relief	833 per 1,000	892 per 1,000 (658 to 1,000)	RR 1.07 (0.79 to 1.44)	30 (1 study)	⊕⊖⊖⊖ Low[2,5]
Tenderness of vaginal nodules in posterior fornix	909 per 1,000	873 per 1,000 (673 to 1,000)	RR 0.96 (0.74 to 1.25)	27 (1 study)	⊕⊕⊖⊖ Low[2,5]
Adnexal masses disappearance or shrinkage	750 per 1,000	908 per 1,000 (638 to 1,000)	RR 1.21 (0.85 to 1.72)	33 (1 study)	⊕⊕⊖⊖ Low[2,5]

CI: confidence interval; RR: risk ratio; MD: mean difference
1 Defined as a complete resolution of all symptoms and signs and included pregnancy, when desired, within 3 years of stopping treatment
2 Not clear if blinding was performed
3 Although the outcome is defined, it is wide, encompassing different symptoms and signs
4 CI for estimate is very wide, crossing 2 thresholds
5 CI for estimate is very wide, crossing 1 threshold

Table 102: Summary clinical evidence profile, Comparison 13: CHM and acupuncture compared to danazol for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Danazol	Chinese herbal medicine and Acupuncture			
Dysmenorrhoea (cessation)	342 per 1,000	400 per 1,000 (222 to 715)	RR 1.17 (0.65 to 2.09)	78 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Lumbosacral pain (cessation)	316 per 1,000	376 per 1,000 (202 to 695)	RR 1.19 (0.64 to 2.2)	78 (1 study)	⊕⊖⊖⊖ Very low[1,2]
Dyspareunia (cessation)	53 per 1,000	125 per 1,000 (26 to 606)	RR 2.38 (0.49 to 11.51)	78 (1 study)	⊕⊖⊖⊖ Very low[1,2]

CI: confidence interval; RR: risk ratio
1 No blinding
2 CI for estimate is very wide, crossing 2 thresholds

ISBN 978-1-4731-2661-9

Table 103: Summary clinical evidence profile, Comparison 14: Acupuncture TENS compared to Self-applied TENS for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Self-applied TENS	Acupuncture TENS				
Change (from baseline) in QoL (EHP-30 total score)	-	The mean change (from baseline) in QoL (EHP-30 total score) in the intervention groups was 1.39 lower (8.94 lower to 6.16 higher)	MD -1.39 (-8.94 to 6.16)	22 (1 study)	⨁⊖⊖⊖ Very low[1,2]	All women were undergoing hormone therapy with continuous progestin or combined oral contraceptives for at least 3 months

CI: confidence interval; MD: mean difference; EHP-30: Endometriosis Health Profile
1 No blinding
2 CI for estimate is very wide, crossing 2 thresholds

11.2.5 Economic evidence

No health economic evidence was found on the cost effectiveness of non-pharmacological interventions for the treatment of endometriosis. Consequently this issue was considered in a de novo economic model. Some of the relevant sections to this review are described in the guideline (for further information on the complete model see Appendix K).

11.2.5.1 Summary of relevant section of the health economic model

Owing to a lack of clinical evidence, only 2 non-pharmacological techniques were considered for economic analysis. These were acupuncture and a generic category of Chinese Herbal Medicine (CHM). Unlike pharmacological or surgical interventions, the cost of non-pharmacological interventions is not well fixed and can vary greatly depending on the technique and supplier. Consequently there is a considerable margin for error on these estimates. Table 104 below gives the estimated annual cost for these interventions, but their estimation is described in more detail in the subsequent paragraphs.

The cost of acupuncture was taken from NICE guidance NG23 (Menopause). This estimates £65 for an initial appointment and then £40 for 12 subsequent appointments and is based on estimates from the UK Acupuncture Clinic (retrieved 15/11/16). Committee opinion was that this likely underestimated the cost of acupuncture in the case of endometriosis, as most women would not consider a 3-month-on treatment 9-month-off treatment schedule to be acceptable to them.

TCM is not typically prescribed on the NHS and thus it is difficult to acquire costings from the BNF. Anecdotally, most users purchase their TCM from health food stores or online from sites such as Amazon.com. This difficulty is compounded by inconsistency in labelling the active ingredient; for example, Dan'e is a mixture of Radix Salviae miltiorrhizae and Rhizoma Zedoariae with no clear indication of the typical ratio between them. Estimating dosage is also difficult, as typically users are advised to vary the dose until the desired effect is achieved. A TCM advocacy group (retrieved 15/11/16) recommends a dose of between 5g and 10g of Dan'e daily, which, based on purchasing a bulk bag from Amazon.com (retrieved

ISBN 978-1-4731-2661-9

15/11/16) would require between 3 to 7 such bags a year. At the recommended maximum dose of 30g it would require 22 bags per year. Based on an average of a 7.5g daily dose, the total annual cost for the drugs would be £120.77.

Table 104: Estimated annual direct cost of non-pharmacological interventions included in economic model

Technique	Cost	Source (see above for details)
Acupuncture	£545.00	NG23
TCM	£120.77	Amazon.com

TCM: Traditional Chinese Medicine

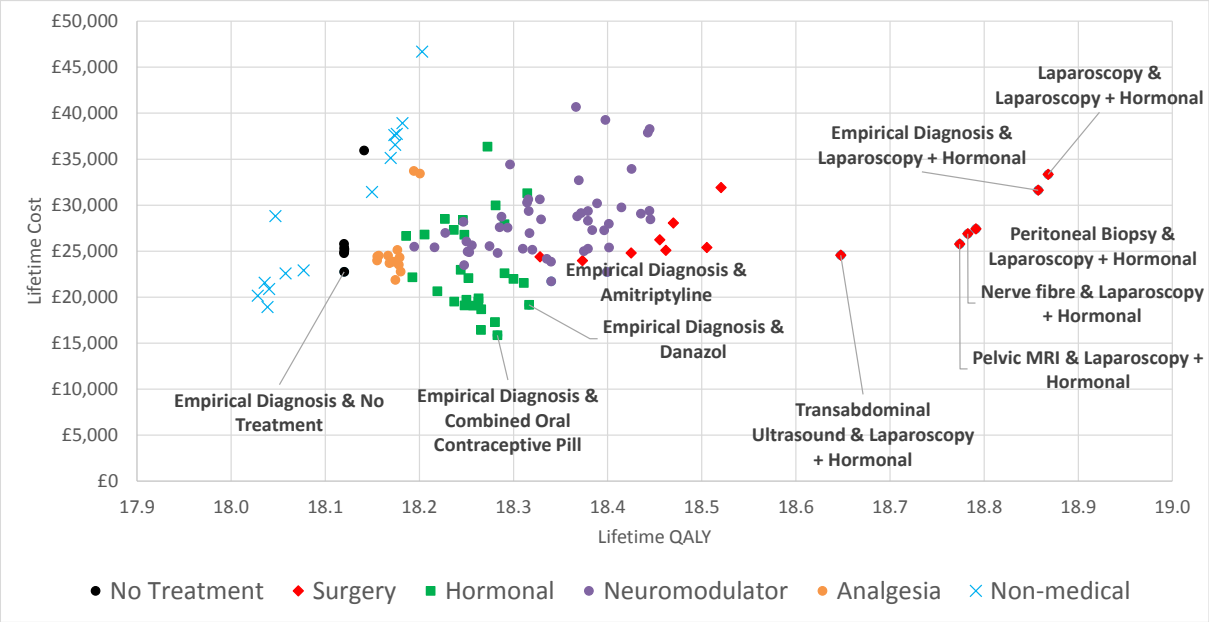
Table 105 identifies the cost and effectiveness of all non-dominated treatment strategies containing a non-pharmacological intervention (in the subset of all treatment strategies containing a non-pharmacological intervention; the table does not imply that these interventions are likely to be superior to hormonal or surgical treatment); if a test/treat dyad is not listed then it is because an alternative treatment is available at the same cost that gives more QALYs. The results demonstrate that herbal medicine is both unlikely to be cost-effective on average and unlikely to benefit any woman more than placebo. Acupuncture is likely to be cost-effective on average and moderately likely to be cost-effective for an individual patient at the upper limit of the conventional NICE threshold (£30,000/QALY). However, this is only true when looking at non-pharmacological interventions in isolation; Figure 19 demonstrates that acupuncture is dominated by pharmacological methods of pain relief including hormonal treatments and a willingness to pay for acupuncture implies a willingness to pay for surgery if these methods are inappropriate (it is extendedly dominated).

Table 105: Cost and effectiveness of all non-dominated treatment strategies containing a non-pharmacological intervention

Treatment	Cost	QALY	ICER	Pr. cost-effective vs. no treatment (£20k/QALY)	Pr. cost-effective vs. no treatment (£30k/QALY)
Empirical diagnosis and no treatment	£22,899.35	18.739	Base case	N/A	N/A
Empirical diagnosis and herbal medicine	£17,215.53	18.645	Extendedly dominated	14.29%	19.05%
Pelvic MRI and herbal medicine	£19,355.67	18.653	Extendedly dominated	14.29%	14.29%
Transabdominal ultrasound and herbal medicine	£21,234.45	18.665	Extendedly dominated	28.57%	33.33%
Empirical diagnosis and acupuncture	£22,984.66	18.799	£1,419.84	52.38%	61.90%
Laparoscopy and acupuncture	£32,746.25	18.822	£428,026.09	61.90%	66.67%

ISBN 978-1-4731-2661-9

Figure 19: Base case analysis (pain) – costs and QALYs, no outliers



Source: Economic Model

11.2.6 Clinical evidence statements

11.2.6.1 Comparison 1: Conventional oral contraceptive pill and Dan'e Chinese herbal medicine vs. no treatment

Fertility

Low and moderate quality evidence from 1 RCT (n=156) found no clinically significant difference in incidence in live birth or miscarriage at 12 months after treatment ended when use of cOCP and Dan'e CHM in combination was compared to no treatment.

11.2.6.2 Comparison 2: Conventional oral contraceptive pill and Dan'e Chinese herbal medicine vs. conventional oral contraceptive pill

Fertility

Low and moderate quality evidence from 1 RCT (n=156) found no clinically significant difference in incidence in live birth or miscarriage at 12 months after treatment ended when use of cOCP and Dan'e CHM in combination was compared to use of cOCP alone.

11.2.6.3 Comparison 3–5: Dietary supplements vs. placebo, dietary supplements vs. GnRH agonist and dietary supplements vs. conventional oral contraceptive pill

Recurrence rates

Low quality evidence from 1 RCT (n=240) found no clinically significant difference in endometrioma recurrence at 18 months after surgery when post-operative use of a 6 month course of dietary supplements (including vitamin, mineral and fatty acid supplementation) was compared to placebo, GnRH agonist (tryptorelin or leuprorelin) or a cOCP (continuous, low-dose).

ISBN 978-1-4731-2661-9

11.2.6.4 Comparison 6: Acupuncture vs. sham acupuncture

Pain

Very low and low quality evidence from 1 RCT (n=18) found a clinically significant improvement in pain reduction at 4 weeks during treatment when Japanese-style acupuncture was compared to sham acupuncture. However, there was no clinically significant difference between the 2 interventions for pain assessed at the end of 8 weeks of treatment and at 6 month follow-up.

Moderate quality evidence from 1 RCT (n=42) found a clinically significant improvement in pain reduction for chronic pelvic pain and dyspareunia at 2 months after treatment when acupuncture was compared to sham acupuncture.

Quality of life

Low quality evidence from 1 RCT (n=18) found a clinically significant improvement in quality of life (EHP total score) at 4 weeks during treatment, at the end of 8 weeks of treatment and at 6 month follow-up when Japanese-style acupuncture was compared to sham acupuncture.

Low quality evidence from 1 RCT (n=18) found no clinically significant difference in quality of life at 4 weeks during treatment (Pediatric QoL Inventory total score) when Japanese-style acupuncture was compared to sham acupuncture. There may be a clinically significant benefit of Japanese-style acupuncture compared to sham acupuncture for improvement in quality of life at the end of 8 weeks of treatment, but there is uncertainty around the estimate. However, there was a clinically significant improvement in quality of life at 6 month follow up when Japanese-style acupuncture was compared to sham acupuncture.

Activities of daily living

Very low and low quality evidence from 1 RCT (n=18) found a clinically significant benefit in improvement in activities of daily living at 4 weeks during treatment when Japanese-style acupuncture was compared to sham acupuncture. However, there was no clinically significant difference between the 2 interventions for activities of daily living assessed at the end of 8 weeks of treatment and at 6 months follow up.

11.2.6.5 Acupuncture vs. danazol

Cure of symptoms

Very low quality evidence from 1 RCT (n=70) found no clinically significant difference in cure of endometriosis symptoms at 3 months post-treatment when use of abdominal acupuncture was compared to danazol over 3 menstrual cycles.

11.2.6.6 Comparison 8: Acupuncture vs. Chinese herbal medicine

Dysmenorrhoea

Very low quality evidence from 1 RCT (n=67) found a clinically significant improvement in dysmenorrhoea at the end of 3 months treatment when use of ear acupuncture therapy was compared to oral administration of CHM.

Cure of symptoms

Low quality evidence from 1 RCT (n=67) found that there may be a clinically significant benefit at the end of 3 months treatment with ear acupuncture therapy compared to oral

ISBN 978-1-4731-2661-9

administration of CHM for cure of endometriosis symptoms, but there is uncertainty around the estimate.

11.2.6.7 Comparison 9: Chinese herbal medicine (individualised decoction) vs. placebo

Pain and quality of life

Very low and low quality evidence from 1 RCT (n=33) found no clinically significant differences in pain symptoms (VAS) or quality of life (MYMOP and EHP 30) at the end of 16 weeks treatment with an individualised CHM decoction compared to a placebo decoction.

11.2.6.8 Comparison 10: Chinese herbal medicine (Nei Yi pills) vs. danazol

Pain

Low quality evidence from 1 RCT (n=58) found clinically significant improvement in symptomatic relief within 3 years of stopping treatment. However, there was no clinically significant difference dysmenorrhoea score, lumbosacral pain relief, rectal irritation relief, tenderness of vaginal nodules in the posterior fornix at the end of 3 months treatment with CHM (Nei Yi pills) compared to danazol (low quality evidence).

Reduction in the size and extent of endometriotic cysts

Very low quality evidence from 1 RCT (n=58) found no clinically significant difference in disappearance or shrinkage of adnexal masses at the end of 3 months treatment with CHM (Nei Yi pills) compared to danazol.

11.2.6.9 Comparison 11: Chinese herbal medicine (Nei Yi pills plus Nei Yi enema) vs. danazol

Pain

Low quality evidence from 1 RCT (n=58) found clinically significant benefit in symptomatic relief (within 3 years of stopping treatment) and reduction in dysmenorrhoea score at the end of 3 months treatment with CHM (Nei Yi pills plus Nei Yi enema) compared to danazol. There may be a clinically significant benefit of CHM (Nei Yi pills plus Nei Yi enema) compared to danazol for rectal irritation relief, but there is uncertainty around the estimate. No clinically significant differences in lumbosacral pain relief or in tenderness of vaginal nodules in the posterior fornix were identified when CHM (Nei Yi pills plus Nei Yi enema) and danazol were compared.

Reduction in the size and extent of endometriotic cysts

Low quality evidence from 1 RCT (n=58) found clinically significant benefit in disappearance or shrinkage of adnexal masses at the end of 3 months treatment with CHM (Nei Yi pills plus Nei Yi enema) compared to danazol.

11.2.6.10 Comparison 12: Chinese herbal medicine (Nei Yi pills plus Nei Yi enema) vs. Chinese herbal medicine (Nei Yi pills)

Pain

Very low and Low quality evidence from 1 RCT (n=58) found that there may be a clinically significant improvement in dysmenorrhoea at the end of 3 months treatment when CHM administered orally and rectally (Nei Yi pills plus Nei Yi enema) compared to oral administration of CHM alone (Nei Yi pills), but there is uncertainty around the estimate. No clinically significant differences in symptomatic relief, lumbosacral pain relief, rectal irritation

ISBN 978-1-4731-2661-9

relief or tenderness of vaginal nodules in posterior fornix were found when the 2 interventions were compared.

Reduction in the size and extent of endometriotic cysts

Low quality evidence from 1 RCT (n=58) found no clinically significant difference in disappearance or shrinkage of adnexal masses at the end of 3 months treatment when CHM administered orally and rectally (Nei Yi pills plus Nei Yi enema) and oral administration of CHM alone (Nei Yi pills) were compared.

11.2.6.11 Comparison 13: Chinese herbal medicine (Gui-Zhi-Fu-Ling-Wan) and acupuncture vs. danazol

Pain

Very low quality evidence from 1 RCT (n=78) found no clinically significant differences in dysmenorrhoea, lumbosacral pain or dyspareunia at the end of 3 months treatment when use of CHM (Gui-Zhi-Fu-Ling-Wan) and acupuncture in combination was compared to danazol.

11.2.6.12 Comparison 14: Acupuncture-like transcutaneous electrical nerve stimulation vs. self-applied transcutaneous electrical nerve stimulation

Quality of life

Very low quality evidence from 1 RCT (n=22) found no clinically significant difference in quality of life (EHP-30 total score) when use of acupuncture-like TENS was compared to self-applied TENS.

11.2.7 Evidence to recommendations

11.2.7.1 Relative value placed on the outcomes considered

The principal aim of this review is to determine the clinical and cost effectiveness of non-pharmacological therapies in reducing pain in women with endometriosis or suspected endometriosis. The Committee prioritised relief of endometriosis-related pain, health-related quality of life and adherence to the treatment programme as critical outcomes when considering recommendations. The remaining outcomes of improvement in fertility rates (live birth), reduction in the size and extent of endometriotic cysts, improvement of endometriosis-related symptoms apart from pain (e.g. fatigue), adverse effects resulting from the intervention, rates of reoccurrence and activities of daily living were considered to be important.

11.2.7.2 Consideration of clinical benefits and harms

The Committee agreed that the evidence on non-pharmacological treatments for endometriosis-related pain management was very uncertain and of limited value.

The Committee noted that some of the non-pharmacological medicines, particularly CHM, are not available within the NHS or are not applicable in the UK setting. The Committee discussed and agreed that there is some evidence that CHM may be effective but expressed their concern regarding standardisation, regulation, efficacy and safety of these medicines.

The Committee's opinion regarding recommending non-pharmacological treatments was divided: some of the Committee members would not discourage women who would like to try alternative treatment options but would warn them to be cautious, for example, regarding

ISBN 978-1-4731-2661-9

CHM or a particular diet; other Committee members felt that they would not encourage women to try alternative treatments and noted their potentially negative impact on health and interactions with standard treatment.

Some of the Committee members expressed concern that, for example, physiotherapy pain-management interventions are not necessarily disease-specific (for example, a population of women with chronic pelvic pain may also include some with endometriosis-related pain), therefore search criteria applied in this guideline may have led to an impression that there is no evidence regarding physio-related pain management interventions. Further, pain management (such as psychological and behaviour interventions) would be more broadly applicable to people with other reasons for chronic pain. However, when finalising the protocol, the Committee specified a threshold of 66% of women have a diagnosis of endometriosis that for studies with mixed populations of women with chronic pelvic pain.

Some of the Committee members, based on their experience, suggested that physiotherapy and psychological pain management approaches are definitely effective. However, the Committee stressed that there is no evidence for these approaches.

11.2.7.3 Consideration of economic benefits and harms

The Committee discussed the results of the health economic model which demonstrated that Herbal Medicine is both unlikely to be cost-effective on average and unlikely to benefit any woman more than placebo. Acupuncture is likely to be cost-effective versus placebo on average and moderately likely to be cost-effective for an individual patient (especially at a threshold of £30,000 / QALY). However the Committee agreed that this is only true when looking at non-pharmacological interventions in isolation. The Committee noted that the economic evidence clearly indicates that a willingness to pay for acupuncture implies a willingness to pay for surgery if these methods are inappropriate and therefore agreed not to recommend acupuncture on the basis of cost implications as it would only be appropriate if a woman could not tolerate any other treatment considered in the guideline and such a woman would have so idiopathic a condition that these recommendations would probably not apply to her.

The Committee discussed how certain interventions on the protocol but for which no evidence were found had a high probability of being cost-effective. This was especially true for behavioural interventions such as a Pain Management Programme and Psychosexual Counselling. The reason for the Committee's observation is that these programmes are offered once early in the treatment of a woman with endometriosis (or sometimes shortly following diagnosis) but are expected to 'pay off' with a steady improvement in QALYs over the rest of the woman's life. In pain management in particular, there may also be a positive economic impact if women are switched away from expensive drugs or treatments with unpleasant side effects and onto alternative methods of managing their pain. Given an expected cost of £1500 for any of these programmes the QALY gain required per year for cost-effectiveness at £20,000 would only be around 0.0025, which the Committee noted was easily achievable. Because of the extremely high potential for high value-of-information in this area, the Committee decided a Research Recommendation was especially important in this instance.

The Committee also noted that the interaction profile of certain herbal medicines was not well understood and so they might have an effect on other treatments women might want to try. Other non-pharmacological treatments, including acupuncture, would be unlikely to interact with any other treatment attempted – however there was no evidence that the full benefit of the non-pharmacological treatment would be felt in this instance. Consequently the Committee decided that there was insufficient evidence to recommend non-pharmacological treatment even in combination with pharmacological therapies.

ISBN 978-1-4731-2661-9

11.2.7.4 Quality of evidence

Evidence was not available for the majority of interventions stipulated in the protocol e.g. no evidence was available for behavioural medicine. Acupuncture and TENS were the only physical interventions examined. Diet was considered in 1 study but the intervention was insufficiently described to be used in clinical practice. The majority of the evidence was regarding Chinese herbal medicines and the Committee considered that their use was not without potential harm.

The Committee noted that several of the studies were small and that although the range in quality of the evidence was from moderate to very low, the majority of evidence was of low or very low quality.

The Committee discussed the paucity of available evidence and concluded that there was a broader evidence base regarding the effectiveness of behavioural medicine and other interventions used in pain management but that this would be drawn from studies of mixed populations of women, not uniquely those with endometriosis and hence would be excluded from the review.

The Committee concluded that there is lack of evidence on physical activity, psychological pain management and particularly dietary interventions and made recommendations for research in populations of women with endometriosis. They also stressed that it is not only important to encourage research but also to improve its quality.

The Committee agreed that they should not only focus on the evidence presented but also discuss other interventions listed in the protocol for which no evidence was found.

11.2.7.5 Other considerations

The Committee considered that no additional recommendations were necessary for equality reasons.

The Committee did not believe that their recommendations would constitute a change of practice requiring additional support for implementation. They acknowledged that pain management clinics may use interventions for women with endometriosis on the basis of practice in a broader population of people experiencing pain.

It was noted that many of the interventions specified in the protocol would be accessible to women outside the NHS. A recommendation was made to ensure that healthcare professionals advise women that the Committee considered there to be insufficient evidence to recommend their use. There were specific concerns regarding herbal medicine preparations and the Committee drew upon recommendations made in the Menopause guideline to echo these concerns for women with endometriosis.

The Committee were concerned that many of the currently used non-pharmacological treatments were not supported by evidence. The Committee intended to look for evidence on a wide range of psychological, physical and lifestyle treatments (see Appendix D). However, the Committee agreed that the lack of evidence specifically addressing a population of women with endometriosis made it difficult to draft recommendations for these management strategies and particularly for dietary interventions.

The Committee agreed that this would be an important topic for future research and made recommendations for research in populations of women with endometriosis which would hopefully inform an update of this guideline. They also stressed that it is not only important to encourage research but also to improve its quality.

They decided that there would be benefit in research investigating commonly used pain management programmes specifically in populations of women with endometriosis. Moreover, the Committee noted that women with endometriosis in support groups discuss

ISBN 978-1-4731-2661-9

lifestyle interventions that they perceive as helpful. Generally these are related to nutrition (such as the Endo diet) and exercise. The Committee further noted that these would be important interventions that could be promoted to self-manage symptoms if found to be effective. They therefore decided that lifestyle interventions should also be proposed as a research recommendation to inform future guidance.

11.2.7.6 Key conclusions

The Committee concluded that there are no non-pharmacological treatments that are clinically and cost-effective and with good evidence. They therefore decided not to recommend any particular non-pharmacological intervention but agreed that future research should be prioritised in this topic, particularly relating to pain management programmes and lifestyle changes.

11.2.8 Recommendations

39. Advise women that the available evidence does not support the use of traditional Chinese medicine or other Chinese herbal medicines or supplements for treating endometriosis.

11.2.9 Research recommendations

2. Are pain management programmes a clinically and cost-effective intervention for women with endometriosis?

Why this is important

Pain is one of the most debilitating symptoms of endometriosis. Endometriosis-related pain can be acute or chronic, and can adversely affect the woman's quality of life, ability to work, and can affect partners and their families.

Pain management programmes have been found to be effective in managing chronic pelvic pain, and can improve quality of life. However, it is unclear how much of this small evidence base can be generalised to women with endometriosis for which evidence is lacking. Furthermore, pain management programmes have not been compared with other treatments available for endometriosis. Pain management programmes promote self-management and are often provided in the community.

If found to be effective for endometriosis, pain management programmes would provide an additional or alternative treatment option for women experiencing endometriosis-related pain. Groups of particular interest are women for whom hormonal and surgical options have been exhausted, women who would prefer an alternative to a pharmacological or surgical approach, and women who may be prioritising trying to conceive.

Table 106: Research recommendation rationale

Research question	Are Pain Management Programmes (PMPs) a clinically and cost-effective intervention for women with endometriosis?
Importance to 'patients' or the population	Minimising distress and disability associated with chronic pelvic pain is of prime importance for women with endometriosis to maximise their overall quality of life and emotional wellbeing. PMPs are well established as interventions for people with chronic pain conditions, to minimise the physical disability and psychological distress associated with chronic pain by developing effective self-management techniques. PMPs may also reduce longer term healthcare costs. However, few services and programmes exist to support the specific needs of women with endometriosis. While there is some published research evaluating the

ISBN 978-1-4731-2661-9

Research question	Are Pain Management Programmes (PMPs) a clinically and cost-effective intervention for women with endometriosis?
	efficacy of pelvic pain specialist PMPs (which include women with endometriosis), there is little evidence specifically for women with endometriosis. This is a developing area within the field of pain management but many women experiencing endometriosis have valued PMPs and considered these to be an effective multidisciplinary intervention to support them.
Relevance to NICE guidance	This is relevant to NHS guidance because it could help to minimise the functional, psychological and social impact of endometriosis, thereby improving quality of life, emotional wellbeing and minimising associated financial costs. Since this is currently lacking evidence future guidelines would benefit from this information which could lead to recommendations in an update of this guideline.
Relevance to the NHS	This is highly relevant because appropriate self-management of chronic pain can reduce unnecessary repeated visits to GPs, A&Es, outpatients and repeated investigations. It can also minimise the impact on emotional wellbeing and mental health. This would reduce the requirement in health, social and educational settings and therefore also reduce costs.
National priorities	Since the Chief Medical Officer's Report of 2008, which made chronic pain a focus, the management of chronic pain has been recognised as of huge national importance. Since then a national pain summit, pain audit and pain service specifications have been achieved: https://www.england.nhs.uk/wp-content/uploads/2013/06/d08-spec-serv-pain-mgt.pdf The Royal College of GPs made treatment of chronic pain a priority (http://content.digital.nhs.uk/catalogue/PUB09300/HSE2011-Ch9-Chronic-Pain.pdf) Active support for self-management is now also seen as the first priority for commissioners (Kings Fund 2015) (https://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/10PrioritiesFinal2.pdf)
Current evidence base	There is a small evidence base for women with chronic pelvic pain. No evidence was identified that addressed this topic in women with endometriosis.
Equality	Pain management programmes will have to take into consideration any protected equalities groups, such as age, sexuality, or people with learning difficulties. Possible communication difficulties or need for interpreters may also need to be taken into consideration when designing materials. Cultural differences may also impact on the way pelvic pain is described or interpreted.
Feasibility	PMPs are audited as standard clinical practice, therefore evaluating the efficacy and outcomes is feasible. It may be difficult to collect follow up data for several years after the intervention as most clinical services do not offer follow up for longer than 1 year.
Other comments	Pain management programmes are multi-faceted and can therefore be tailored to individual needs. Therefore they actively promote equalities.

Table 107: Research recommendation PICO table

Criterion	Explanation
Population	An RCT of women with endometriosis who are suitable for and complete a specialist Pelvic Pain Management Programme. Multi-centre research may be feasible if services are matched for specialism and PMP intervention quality. Cross-over study against wait list control.
Intervention	Completion of a specialist PMP specifically designed for women with endometriosis (and/or other diagnoses which result in pelvic pain). It should be gender specific i.e. a programme for women only and

ISBN 978-1-4731-2661-9

Criterion	Explanation
	delivered by a multidisciplinary team including clinical psychologists, pain management physiotherapists and consultants in pain medicine with experience managing pelvic pain. It should specifically address issues such as sexual, bowel and bladder function in the context of pain. It may also touch on important issues for individuals for example, fertility. Specialised PMPS are currently run by NHSE-recognised Specialised Pain Services.
Comparator	Outcomes for women with endometriosis related chronic pain treated by PMP could be compared with outcome data for other patient groups attending PMPs e.g. those with other pelvic pain, musculoskeletal and neuropathic pain to benchmark effectiveness of PMPS for endometriosis against other pain conditions. Women with endometriosis who have completed medical and surgical management wait list control.
Outcome	Validated pain (physical, functional and psychological outcome measures recognised in pain management specialism such as the Brief Pain Inventory. Healthcare and medicines utilisation health related quality of life and costs associated with the delivery of the PMP.
Study design	Study design: A multi-centre RCT evaluating the outcomes and long-term efficacy of specialised pelvic PMPs on women with a diagnosis of endometriosis. The study should also collect prospective community (GP) service and hospital data to evaluate medicines and healthcare utilisation economics.
Timeframe	Within 5 years

3. Are specialist lifestyle interventions (diet and exercise) effective, compared with no specialist lifestyle interventions, for women with endometriosis?

Why this is important

Endometriosis is a long-term condition that can cause acute and chronic pain, and fatigue. It has a significant and sometimes severe impact on the woman's quality of life and activities of daily living, including relationships and sexuality, ability to work, fertility, fitness and mental health.

Supporting self-management is critical to improving quality of life for women living with endometriosis. In order to successfully self-manage the condition, women need evidence-based, easily accessible information about the condition and ways of managing it that support surgical and medical treatment. However, no high-quality research was identified on the effectiveness of lifestyle interventions such as diet or exercise and other non-medical treatments in reducing pain, fatigue and other symptoms.

Studies should aim to provide evidence-based options to support self-management of endometriosis. This would improve the quality of life of women with endometriosis, enabling them to manage pain and fatigue, and reducing the negative impact on their career, relationships, sex lives, fertility, and physical and emotional wellbeing.

Table 108: Research recommendation rationale

Research question	Are specialist lifestyle interventions (diet and exercise) effective, compared with no specialist lifestyle interventions, for women with endometriosis?
Importance to 'patients' or the population	Effective self-management is critical for the wellbeing of women with endometriosis. They receive care in a range of settings where information and support on self-management of endometriosis varies widely. They consistently report uncertainty on lifestyle interventions relating to self-

ISBN 978-1-4731-2661-9

Research question	Are specialist lifestyle interventions (diet and exercise) effective, compared with no specialist lifestyle interventions, for women with endometriosis?
	management of endometriosis. Many patients seek self-management lifestyle intervention information online. However, this information is not based on high quality research and therefore is not evidence-based guidance. High quality, evidence-based research would enable clinical staff to provide accurate, safe and consistent guidance to complement surgical and medical treatment and advice, enabling endometriosis patients to plan effective self-management options.
Relevance to NICE guidance	This is highly relevant to NHS guidance as it could help to establish the safety and effectiveness of specialist lifestyle interventions in the management of endometriosis. Since there is lack of high quality evidence, future NICE guidance would greatly benefit from the identification of appropriate strategies to self-manage the condition through specialist lifestyle interventions.
Relevance to the NHS	This is highly relevant to the NHS as effective self-management of endometriosis can reduce unnecessary repeated visits to GPs, A&Es, outpatients, repeated investigations and other interventions. This could also minimise associated financial costs if more women are empowered to self-manage their condition outside of the NHS.
National priorities	Since the Chief Medical Officer's Report of 2008, which made chronic pain a focus, the management of chronic pain has been recognised as of huge national importance. Since then a national pain summit, pain audit and pain service specifications have been achieved: https://www.england.nhs.uk/wp-content/uploads/2013/06/d08-spec-serv-pain-mgt.pdf The Royal College of GPs made treatment of chronic pain a priority (http://content.digital.nhs.uk/catalogue/PUB09300/HSE2011-Ch9-Chronic-Pain.pdf) Active support for self-management is now also seen as the first priority for commissioners (Kings Fund 2015) (https://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/10PrioritiesFinal2.pdf)
Current evidence base	There is currently no high quality research on the effectiveness of specialist lifestyle interventions.
Equality	Women have the right to accessible, safe and effective information and guidance on how to manage this long-term, life altering condition that can have a negative impact on many aspects of a woman's life.
Feasibility	There are always ethical issues in conducting studies in vulnerable populations. These would require careful consideration, but could be overcome.
Other comments	Not applicable

Table 109: Research recommendation PICO table

Criterion	Explanation
Population	An RCT of women and/or girls with diagnosed or suspected endometriosis, who are suitable for a specialist lifestyle intervention.
Intervention	Completion of a specialist diet and/or exercise intervention designed for women and/or girls with diagnosed or suspected endometriosis.
Comparator	Outcomes for women and/or girls participating in a specialist diet and/or exercise intervention could be compared with outcomes for those participating in a non-specialist diet and/or exercise intervention.
Outcome	Validated outcome measures or questionnaires recognised in lifestyle intervention specialism/field should be used to assess, for example, reduction in pain improvement in energy levels and fitness

ISBN 978-1-4731-2661-9

Criterion	Explanation
	improvement of menorrhagia and dysmenorrhoea improvement in emotional wellbeing improvement in autonomy and ability to manage activities of daily living increased ability to self-manage the condition fewer medical appointments
Study design	A multicentre RCT evaluating the outcomes and the long-term effectiveness of specialised diet and/or exercise interventions.
Timeframe	Within 5 years

11.3 Surgical management and combinations of treatment

11.3.1 Surgery, including ablation and excision (and the surgical network meta-analysis)

11.3.1.1 Introduction

Surgical treatment is an important part of the management of endometriosis, aiming to remove or destroy endometriotic deposits and divide adhesions with restoration of normal anatomy. Surgical treatments can be performed by laparoscopy (traditional or robotic) or as an open procedure (laparotomy). Current practice is to use a laparoscopic approach, as it offers several advantages when compared to open procedures, including improved visualisation, microsurgical techniques, shorter hospital stay, quicker return to normal function and cost.

Endometriotic deposits can be treated by excision (cutting them out) or ablation (destruction or evaporation using a variety of energy modalities). These techniques are used to treat endometriosis of all degrees of severity. Surgical techniques such as the choice of energy modality may be influenced by the surgeons' training and preferences. Severe endometriosis involving the bowel, bladder and ureter may require additional surgical expertise, including colorectal surgeons and urologists. Surgery has a role in the management of recurrent disease, although it is recognised that outcomes may reduce with increasing numbers of operations.

Even if all endometriosis tissue is removed by excision or ablation, the risk of recurrence is high. Relapse of symptoms occurs in 40–45% of women and up to 30% of women are readmitted for surgery within 5 years. Half of all women diagnosed with endometriosis require a second operation and just over a quarter will undergo 3 or more procedures.

Reduction of pain due to presumed recurrence currently involves the use of hormonal treatments pre- or post-surgery. The rationale for this is that hormonal treatments reduce circulating levels of oestrogen leading to lighter or no periods, theoretically causing shrinkage of existing endometriosis lesions and preventing new lesions developing. The Committee were interested in assessing the clinical and cost effectiveness of surgery as well as the effectiveness of pre- and post-surgical hormonal treatment.

The aim of the review question was to assess the evidence for excisional and ablative surgical techniques and combinations of hormonal treatments with surgery, and to compare their clinical and cost effectiveness in the management of endometriosis, including the management of ovarian endometriomas. A review on laparoscopic uterine nerve ablation (LUNA) for chronic pelvic pain was not prioritised, because there is a NICE interventional procedure guideline on this topic.

For full details, see the review protocol in Appendix D, the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I and full GRADE profiles in Appendix J.

ISBN 978-1-4731-2661-9

11.3.1.2 Network Meta-analysis

What is the effectiveness of the following treatments for endometriosis, including recurrent and asymptomatic endometriosis:

- surgery
- combined surgery and hormonal treatment?

11.3.1.3 Methods

Study selection and data collection

For full details see review and analysis protocols in Appendix D.

Outcome measures for NMA

For assessing the effectiveness of different surgical or combined surgery plus hormonal treatments, the Committee identified pain relief and health-related Quality of Life (QoL) as critical outcomes for which NMA could be used to aid decision-making. NMAs were performed on these outcomes where evidence was available. Pain relief

For pain relief, the visual analogue scale (VAS) was considered to be the most widely used useful pain scale for which data would be available.

Health-related QoL

For health-related QoL, the SF-36 was determined to the most useful scale that was widely used in the literature. However, there were not a sufficient number of studies available from the systematic review to allow for NMA. Therefore these studies were analysed using pairwise meta-analysis where appropriate.

Statistical methodology

Data were available for a number of treatments and routes of administration. Due to the sparseness of the networks, it was necessary to group treatments within different classes and assume a common class effect (Table 110). All non-surgical treatments in the table were only included in the NMA if they were administered in combination with surgery.

The common class effects were assessed to identify if it was reasonable to assume similarity of treatment effects within classes. Multi-level NMA models with treatments nested within classes were also examined, though this added complexity did not improve model fit for any of the analyses.

Table 110: Dose ranges of treatments in different classes of interventions, with abbreviations used in tables and figures within this chapter

Class	Treatment	Abbreviation
Diagnostic laparoscopy / No treatment	Diagnostic laparoscopy No treatment/Waiting list	Diagnostic/no treat
Danazol/gestrinone	Danazol (100-800 mg/d) Gestrinone	Dan/gest
Oestrogens (oral)	Oestradiol (1-2 mg/d) Conjugated equine oestrogens (0.3-1.25 mg/d)	Oest(o)
Progestogens (oral)	Norethisterone (2.5 mg/d) Medroxyprogesterone (15-30 mg/d)	Prog(o)

ISBN 978-1-4731-2661-9

Class	Treatment	Abbreviation
	Levonorgestrel (30 micrograms/d) Desogestrel (75 micrograms/d) Dienogest (2 mg/d)	
Progestogens (depot)	Medroxyprogesterone (150 mg/3m) Gestodene (5-10 mg)	Prog(i.m.)
Progestogens (subcutaneous)	Medroxyprogesterone (104 mg/3m) Promegestone	Prog(s.c.)
Progestogens (intrauterine)	Levonorgestrel (20 micrograms/d)	Prog(i.u.)
GnRH agonists (depot)	Leuprorelide (3.75 mg/m) Triptorelin (3 mg/m)	GnRHa(i.m.)
GnRH agonists (subcutaneous)	Goserelin (3.6 mg/m)	GnRHa(s.c.)
GnRH agonists (nasal spray)	Nafarelin (200 micrograms b.d.) Buserelin (300 micrograms t.d.)	GnRHa(i.n.)
GnRH antagonists	Elagolix	GnRHant
Aromatase inhibitors	Anastrozole (1 mg/d) Letrozole (2.5 mg/d)	AromaInhib
Anti-androgens	Cyproterone acetate (only in combination as combined oral contraceptive 2 mg)	Anti-And
Selective oestrogen receptor modulators	Raloxifene (60 mg/d)	SERM
Tibolone	Tibolone (2.5 mg/d)	Laparoscopy
Laparoscopic surgery	Ablation (laser, diathermy, etc.) Excision (laser, diathermy, etc.)	Laparoscopy
Nutritional supplements	Calcium Vitamin D	Supp
Chinese herbal medicine	Nei yi pills Dan'e mixture	CHM
Dietary interventions	Dietary intervention	Diet

(c) Table only includes treatments in full-text studies assessed for inclusion/exclusion. Treatments only in studies that were not included in the NMA could not be included in the network.

11.3.1.4 Summary of included studies

Studies included in the NMA

All studies included women with laparoscopic confirmation of endometriosis.

Table 111: Characteristics of included studies

First Author	Pub Date	rAFS	Surgery type	Endometriomas included	Risk of Bias
Granese	2015	III-IV	Excision/ablation	Some	High
Razzi	2007	NR	Excision	All	Mod
Sutton	1994	I-II	Ablation	None	High
Zhu	2014	I-II	Excision/ablation	None	Mod

Pub Date: Date of publication; rAFS: revised American Fertility Scale; Mod: Moderate; NR: Not reported

ISBN 978-1-4731-2661-9

11.3.1.4.1 Clinical evidence profile

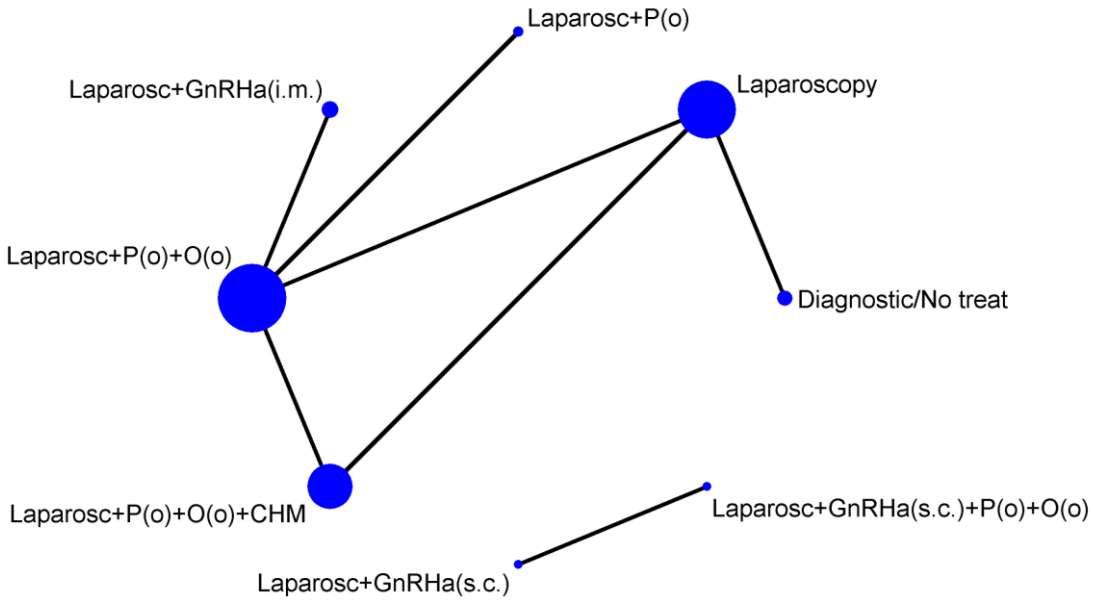
Pain Relief – VAS

Due to difficulty in achieving convergence during estimation, NMAs were conducted separately for hormonal therapies and for surgery and surgery plus hormonal treatment. The Committee felt that this was likely to be because the populations may not have been sufficiently homogeneous, as patients receiving surgical treatment were likely to have failed on hormonal treatments, thus violating the assumption of transitivity.

Surgery and combined surgery plus hormonal therapy

Four trials of 6 surgery or combined surgery plus hormonal treatment classes were included in the network for the outcome of pain relief on the VAS, with a total sample size of 267 women (Figure 20). All studies of combined surgery and hormonal treatment administered hormonal treatment within 4 weeks post-surgery. One study was at high risk of bias, 7 were at moderate risk of bias and 8 at low risk of bias. Three of the 4 trials included women with endometrioma.

Figure 20: Network for surgery and combined surgery plus hormonal therapy for pain relief (VAS)



The size of nodes is proportional to the number of women in the network who were given a particular treatment class. The thickness of connecting lines is proportional to the number of studies directly comparing 2 treatment classes. 2 treatment classes were not connected and could not be compared in the NMA (laparoscopic surgery + GnRHa (s.c.) and laparoscopic surgery + GnRHa (s.c.) + progestogen + oestrogen (oral)). For treatment name abbreviations, see Table 110.

Table 112 presents the results of the pairwise meta-analyses (direct comparisons; upper right section of table) together with the results from the NMA for every possible class comparison (lower left section of table), presented as mean differences. The VAS is a 0-100 patient-reported scale, on which a difference of 10 points has been shown to be clinically significant to patients (Gerlinger 2012). NMA results were derived from a fixed effects model. As no closed treatment loops existed that were not from the same study, incoherence could not be assessed.

ISBN 978-1-4731-2661-9

All treatments led to a clinically significant improvement when compared to diagnostic laparoscopy/no treatment. Use of a hormonal treatment after laparoscopy surgery led to a clinically significant improvement when compared to laparoscopic surgery alone, though evidence for this came exclusively from studies including a majority of women with endometrioma. There were no clinically significant differences between any of the hormonal treatments combined with laparoscopic surgery. Figure 21 graphically presents the results computed by the NMA for each treatment versus placebo.

The combined oral contraceptive pill (P(o) + O (o)) after laparoscopic surgery had the highest probability of being among the best 3 treatments (95.87%), followed by progestogen (oral) after laparoscopic surgery (85.10%) and GnRHa (i.m.) after laparoscopic surgery (84.49%) (Table 113).

Sufficient data to calculate SEs was only available in 2 of the 4 trials. However, sensitivity analyses using the upper 95% credible interval of the posterior for the imputed SEs showed that the probability of being the best treatment results were not sensitive to the imputed SEs (Appendix L). Two results compared to laparoscopic surgery (surgery plus progestogens (oral) and surgery plus the combined oral contraceptive pill plus Chinese herbal medicine) had 95% CrI that included 0, though the numerical was small and the point estimates still suggested strong clinical benefit.

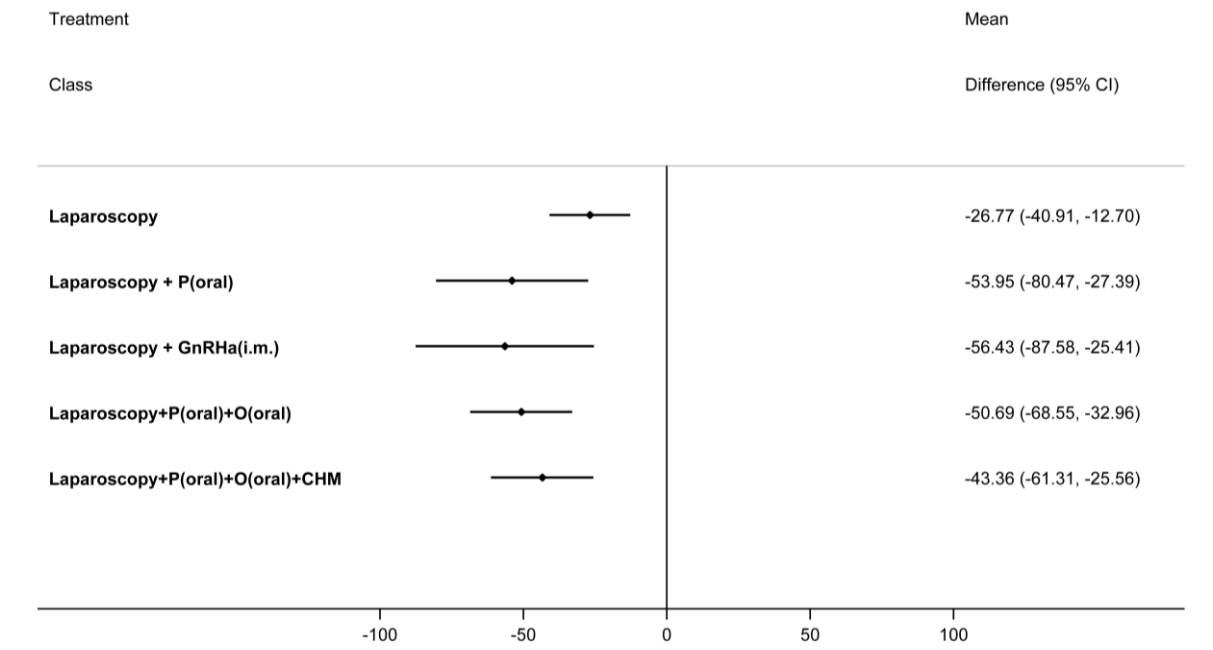
Table 112: Matrix of results for the NMA of surgery and combined surgery plus hormonal therapy for pain relief on the VAS

Diagnostic / no treatment		**-26.8 (-40.9 to -12.7)**			
-26.8 (-40.9 to -12.7)	Laparoscopic surgery			-23.9 (-35.0 to -12.9)	-16.6 (-27.7 to -5.53)
-54.0 (-80.5 to -27.4)	-27.2 (-49.8 to -4.44)	Laparosc and Prog (o)		3.25 (-16.7 to 23.1)	
-56.4 (-87.6 to -25.4)	-29.7 (-57.6 to -1.83)	-2.54 (-35.0 to 30.0)	Laparosc and GnRH (i.m.)	5.75 (-19.9 to 31.4)	
-50.7 (-68.6 to -33.0)	-23.9 (-35.0 to -12.9)	3.25 (-16.7 to 23.1)	5.75 (-19.9 to 31.4)	Laparosc and Prog (o) and Oest (o)	7.32 (-3.79 to 18.4)
-43.4 (-61.3 to -25.6)	-16.6 (-27.7 to -5.53)	10.6 (-12.1 to 33.2)	13.09 (-14.9 to 41.0)	7.32 (-3.79 to 18.4)	Laparosc and P (o) and O (o) and CMH

(a) Mean differences and 95% credible intervals from the NMA (bottom left diagonal) and conventional meta-
(b) analyses (top right diagonal) treatment effects between the column-defined and row-defined treatments. Mean
(c) differences less than 0 favour the row-defined treatment. Numbers in bold, grey-shaded cells denote results
(d) where the 95% CrI credible intervals do not include 0. For treatment name abbreviations see Table 110

ISBN 978-1-4731-2661-9

Figure 21: Forest plot showing mean differences (95% CrI) of NMA estimates for each treatment versus diagnostic laparoscopy/no treatment for pain relief on the VAS



For treatment name abbreviations see Table 110

Table 113: Probabilities of being among the best 3 treatments and the worst 3 treatments, and the rank and 95% CrI for each treatment

Treatment Class	Probability of being within the best 3 (%)	Probability of being within the worst 3 (%)	Rank (95% CrI)
Diagnostic/no treatment	0.00%	100.00%	6 (6, 6)
Laparoscopic surgery	0.08%	99.92%	5 (4, 5)
Laparosc + P (o)	85.10%	14.90%	2 (1, 4)
Laparosc + GnRH (i.m.)	84.49%	15.51%	1 (1, 4)
Laparosc + P (o) + O (o)	95.87%	4.13%	2 (1, 4)
Laparosc + P (o) + O (o) + CHM	34.46%	65.54%	(2, 4)

(a) For treatment name abbreviations see Table 110

11.3.1.5 Pairwise comparison of surgical ablation and excision

Review question: What is the effectiveness of surgery (ablation or excision) for the treatment of endometriosis, including recurrent and asymptomatic endometriosis?

11.3.1.5.1 Description of clinical evidence

The objective of this review is to determine the clinical and cost-effectiveness of surgery in improving health related quality of life and reducing adverse events.

Eight studies were included that evaluated the clinical and cost-effectiveness of ablation or excision for the management of endometriosis; 3 systematic reviews (Hart 2008; Dan 2013; Duffy 2014), of which 2 were Cochrane systematic reviews (Hart 2008 and Duffy 2014) and 5 RCTs (Abbott 2004, Carmona 2011; Wright 2005; Healey 2010; Healey 2014).

Two trials were carried out in the United Kingdom (Abbott 2004; Wright 2005), 2 in Australia (Healey, 2010; Healey, 2014) and 1 in Spain (Carmona, 2011). The included studies in the 3

ISBN 978-1-4731-2661-9

systematic reviews (Duffy 2014; Dan 2013; Hart 2008) were carried out in various countries including Australia, Canada, Egypt, Iran and the United Kingdom.

Of the 3 included systematic reviews, 1 consisted of 3 trials (Hart 2008), 1 included 10 trials (Duffy 2014) and the third included 7 trials (Dan 2013).

Two systematic reviews (Hart 2008; Dan 2013) and 1 trial (Carmona 2011) were carried out to determine whether laparoscopic surgical excision or ablation is the optimum surgical management of ovarian endometrioma with respect to pain and fertility outcomes and recurrence rate.

The effectiveness and safety of laparoscopic surgery in the treatment of painful symptoms and subfertility associated with endometriosis was assessed by 1 systematic review (Duffy 2014).

Reduction of pain following laparoscopy after ablation or excision of endometriosis was examined by 1 trial (Healey 2010). A follow-up study was performed 5 years after the operation to assess reduction in the pain score (Healey 2014).

One trial (Wright 2005) compared excisional and ablative treatment modalities for mild endometriosis in the management of chronic pelvic pain.

One trial (Abbott 2004) reported on quality of life in women with endometriosis who received excisional laparoscopy at different time periods. In this evidence report, data on quality of life outcomes were reported from the Aboott 2004 study and pain-related outcomes were reported from the Duffy 2014 systematic review.

The following comparisons were examined using the available evidence:

1. Laparoscopic treatment (excision or ablation) versus diagnostic laparoscopy
2. Excision versus diagnostic laparoscopy
3. Ablation versus diagnostic laparoscopy
4. Excisional surgery versus ablative surgery

No evidence was identified for the following outcomes:

- Effect on daily activities
- Participant satisfaction with treatment

Evidence from these studies is summarised in the clinical GRADE evidence profiles below (Table 115 to Table 117). Descriptive data from the Healey 2004 and Abbott 2004 trials are presented in Table 118 and Table 119, respectively. See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix Summary of included studies

A summary of the studies that were included in this review are presented in Table 114.

Table 114: Summary of included studies

Study	Intervention/Comparison	Population	Outcomes
Abbott 2004 UK	Excision versus diagnostic laparoscopy in a crossover trial	Women with clinically proven endometriosis N=39 Median rAFS scores (range): • AT surgery 1: • Diagnostic laparoscopy group: 27 (6 – 142)	• health related quality of life (EQ-5D, SF-12)

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes
		• Excision group: 16 (3 – 142) • At surgery 2: • Diagnostic laparoscopy group: 46 (3 –142) • Excision group: 0 (0 – 142)	
Carmona 2011 Spain	Laparoscopic cystectomy versus laser vaporisation	Women undergoing laparoscopy for adnexal mass with the diagnosis of endometrioma N=90 Median rAFS scores (range): • Endometrioma cystectomy group: 27 (19 – 96) • Drainage and laser coagulation of the inner lining group: 28 (20 – 94)	• recurrence at 12 months per woman • recurrence at 12 months per endometrioma • recurrence at 60 months per woman • recurrence at 60 months per endometrioma • pregnancy rate after surgical treatment up to 60 months • reoperation after surgical treatment up to 60 months
Dan 2013 (Systematic review)	Laparoscopic ovarian cystectomy versus fenestration/coagulation Laparoscopic ovarian cystectomy versus laser ablation	Women with endometrioma N=7 RCTs included Median rAFS scores (range) across studies ranged from 27 (16 - 136) or 27 (19 – 96) to 32 (16–133); mean (±SD) ranged from 38 ± 3.8 to 81.22 ± 11.88	• recurrence of sign/symptoms • risk of recurrence • pregnancy rate
Duffy 2014 (Systematic review)	Laparoscopic surgery compared with diagnostic laparoscopy Laparoscopic ablation versus laparoscopic excision	Women with clinical symptoms and signs suggestive of endometriosis N=973 rAFS scores one to 4	• pain • live birth or pregnancy rate • adverse events
Hart 2008 (CSR)	Planned surgical excision (stripping) of endometrioma Planned ablation of the endometrioma capsule	Women with ovarian endometrioma N=304 rAFS score not reported	• recurrence of dysmenorrhoea • recurrence of non-menstrual pelvic pain • recurrence of endometrioma • requirements for further surgery • pregnancy rate after controlled ovarian super stimulation • response to stimulation with gonadotrophins
Healey 2010 (as reported in	Ablation versus excision	Women with endometriosis N=103	• overall pain • pelvic pain • period pain

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes
Duffy 2013 CSR) (*outcomes reported only in Healey 2010)		Median rAFS scores (95% CI): • Excision group: 10 (2 – 53) • Ablation group: 7 (1 – 33)	• back pain • rectal pain • thigh pain • abdominal pain • defecation pain • voiding pain • nausea* • abdominal bloating* • vomiting* • dyspareunia*
Healey 2014 Australia	Ablation versus excision	Women of reproductive age with pelvic pain and visually proved endometriosis N=82 Median rAFS scores (range): • Excision group: 9 (2 - 45) • Ablation group: 8 (1 - 26)	• reduction in VAS score at 5 years • overall pain • pelvic pain • period pain • back pain • rectal pain • thigh pain • abdominal pain • defecation pain • voiding pain • nausea • abdominal bloating • vomiting • dyspareunia
Wright 2005 UK	Ablation versus excision	Women with mild endometriosis N=24 rAFS scores one to 2	• dysmenorrhoea • pelvic pain • dyspareunia • dyschezia • constipation • diarrhoea • back pain • fatigue • uterine mobility • tenderness • adnexal pain • symptoms • signs

rAFS: Revised American Fertility Society; EQ-5D: EuroQol 5 dimensions; SF-12: 12-Item Short Form Survey; CI: confidence Intervals; VAS: Visual Analogue Scale; CSR: Cochrane systematic review

11.3.1.5.2 Clinical evidence profile

The clinical evidence profiles for this review question are presented in Table 115 to Table 119.

ISBN 978-1-4731-2661-9

Table 115: Summary clinical evidence profile: Laparoscopic treatment (excision or ablation) versus diagnostic laparoscopy for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Diagnostic laparoscopy	Excision/ablation			
Overall pain better or improved - At 6 months	429 per 1,000	399 more per 1,000 (from 99 more-870 more)	RR 1.93 (1.23 to 3.03)	69 (1 study)	⊕⊖⊖⊖ Very low1,2,3
Overall pain better or improved - At 12 months	214 per 1,000	516 more per 1,000 (from 137 more-1,000 more)	RR 3.41 (1.64 to 7.11)	69 (1 study)	⊕⊕⊖⊖ Low1,2
Live birth or ongoing pregnancy	205 per 1,000	135 more per 1,000 (from 29 more-291 more)	RR 1.66 (1.14 to 2.42)	382 (2 studies)	⊕⊖⊖⊖ Very low2,3
Miscarriage per pregnancy	108 per 1,000	5 fewer per 1,000 (from 60 fewer-118 more)	RR 0.95 (0.44 to2.09)	112 (2 studies)	⊕⊖⊖⊖ Very low5,6

CI: confidence interval; RR: risk ratio

Table 116: Summary clinical evidence profile: Excision versus diagnostic laparoscopy for endometriosis

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Diagnostic laparoscopy	Excision/ablation			
Overall pain better or improved – At 6 months	316 per 1,000	483 more per 1,000 (from 82 more to 1,000 more)	RR 2.53 (1.26 to 5.09)	39 (1 study)	⊕⊕⊕⊕ High1
Overall pain score - At 6 months	-	The mean overall pain score - at 6 months in the intervention groups was 0.9 higher (0.31 to 1.49 higher)	MD 0.90 (0.31 to 1.49)	16 (1 study)	⊕⊖⊖⊖ Very low2,3
Overall pain score - At 12 months	-	The mean overall pain score - at 12 months in the intervention groups was 1.65 higher	MD 1.65 (1.11 to 2.19)	16 (1 study)	⊕⊕⊕⊖ Moderate3

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
		(1.11 to 2.19 higher)			
Pelvic pain scores - At 6 months	-	The mean pelvic pain scores - at 6 months in the intervention groups was 5.1 lower (16.64 lower to 6.44 higher)	MD -5.10 (-16.64 to 6.44)	39 (1 study)	⊕⊕⊕⊖ Moderate1, 3
Dysmenorrhoea pain score - At 6 months	-	The mean dysmenorrhoea pain score - at 6 months in the intervention groups was 2.4 higher (6.18 lower to 10.98 higher)	MD 2.40 (-6.18 to 10.98)	39 (1 study)	⊕⊕⊕⊖ Moderate1, 3
Dyspareunia pain score - At 6 months	-	The mean dyspareunia pain score - at 6 months in the intervention groups was 6.3 higher (8.18 lower to 20.78 higher)	MD 6.30 (-8.18 to 20.78)	39 (1 study)	⊕⊕⊕⊖ Moderate1, 3
EQ-5D index summary score - At 6 months	-	The mean EQ-5D index summary at 6 months in the intervention groups was 0.03 higher (0.12 lower to 0.18 higher)	MD 0.03 (-0.12 to 0.18)	39 (1 study)	⊕⊕⊖⊖ Low1,5
EQ-5D VAS – At 6 months	-	The mean EQ-5D VAS at 6 months in the intervention groups was 17.7 higher (7.02 to 28.38 higher)	MD 17.7 (7.02 to 28.38)	39 (1 study)	⊕⊕⊕⊖ Moderate1, 3
SF-12 Physical component score - At 6 months	-	The mean SF-12 physical component score at 6 months in the intervention groups was 2.7 higher	MD 2.7 (-2.9 to 8.3)	39 (1 study)	⊕⊕⊕⊖ Moderate1, 3

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Partici-pants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
		(2.9 lower to 8.3 higher)			
SF-12 Mental component score – At 6 months	-	The mean FS-12 mental component score at 6 months in the intervention groups was 2.3 higher (4.5 lower to 9.1 higher)	MD 2.3 (-4.5 to 9.1)	39 (1 study)	⊕⊕⊕⊖ Moderate1, 3

CI: confidence interval; RR: Risk ratio; MD: mean difference; EQ-5D: EuroQol five dimensions, SF-12: 12-Item Short Form Survey
1 Unclear if selective reporting
2 Evidence was downgraded by two due to performance bias (blinding of participants and personnel and attrition bias (incomplete outcome data)
3 Evidence was downgraded by one due to serious imprecision as 95%CI crossed one default MID
4 No blinding of participants and personnel and incomplete outcome data
5 Evidence was downgraded by two due to very serious imprecision as 95%CI crossed two default MIDs

Table 117: Summary clinical evidence profile: Excisional surgery versus ablative surgery for endometriosis and endometrioma

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Diagnostic laparoscopy	Excisional/ablation			
Endometriosis					
Pain score (reduction in VAS at 12 months) - Overall	-	The mean pain score (reduction in VAS at 12 months) - overall in the intervention groups was 0 higher (1.22 lower to 1.22 higher)	MD 0 (-1.22 to 1.22)	103 (1 study)	⊕⊕⊖⊖ Low1
Pain score (reduction in VAS at 12 months) - Pelvic	-	The mean pain score (reduction in VAS at 12 months) - pelvic in the intervention groups was 0.1 lower (1.3 lower to 1.1 higher)	MD -0.1 (-1.3 to 1.1)	103 (1 study)	⊕⊕⊖⊖ Low1
Pain score (reduction in VAS at 12 months) - Dyspareunia	-	The mean pain score (reduction in VAS at 12 months) - dyspareunia in the intervention groups was	MD 1.3 (-0.29 to 2.89)	103 (1 study)	⊕⊖⊖⊖ Very low1,2

ISBN 978-1-4731-2661-9

		1.3 higher (0.29 lower to 2.89 higher)			
Unintended effects (reduction from VAS score by 12 months after operation (nausea, vomiting) - Nausea	-	The mean unintended effects (reduction from VAS score by 12 months after operation (nausea, vomiting) - nausea in the intervention groups was 1.1 higher (0.14 lower to 2.34 higher)	MD 1.1 (-0.14 to 2.34)	103 (1 study)	⨁⨁⨀⨀ Low2,3
Unintended effects (reduction from VAS score by 12 months after operation (nausea, vomiting) - Vomiting	-	The mean unintended effects (reduction from VAS score by 12 months after operation (nausea, vomiting) - vomiting in the intervention groups was 0.2 higher (0.71 lower to 1.11 higher)	MD 0.2 (-0.71 to 1.11)	103 (1 study)	⨁⨁⨁⨀ Moderate3
Unintended effects (reduction from VAS score by 12 months after operation (nausea, vomiting) - Bloating	-	The mean unintended effects (reduction from VAS score by 12 months after operation (nausea, vomiting) - bloating in the intervention groups was 0.9 higher (0.3 lower to 2.1 higher)	MD 0.9 (-0.3 to 2.1)	103 (1 study)	⨁⨁⨀⨀ Low2,3
Endometrioma					
Recurrence of pelvic pain - Dysmenorrhoea	548 per 1,000	389 fewer per 1,000 (from 247 fewer to 466 fewer)	RR 0.29 (0.15 to 0.55)	104 (2 studies)	⨁⨁⨁⨀ Moderate3
Recurrence of pelvic pain - Non-menstrual pelvic pain	529 per 1,000	429 fewer per 1,000 (from 127 fewer to 503 fewer)	RR 0.19 (0.05 to 0.76)	37 (1 study)	⨁⨁⨀⨀ Low2,3
Pregnancy rate after surgical treatment	233 per 1,000	242 more per 1,000 (from 56 more to 552 more)	RR 2.04 (1.24 to 3.37)	138 (3 studies)	⨁⨁⨁⨀ Moderate2,4
Recurrence of endometrioma - At 12 months	256 per 1,000	146 fewer per 1,000 (from 69 fewer to 192 fewer)	RR 0.43 (0.25 to 0.73)	258 (4 studies)	⨁⨁⨁⨁ High

ISBN 978-1-4731-2661-9

Recurrence of endometrioma - At 60 months	368 per 1,000	147 fewer per 1,000 (from 261 fewer to 96 more)	RR 0.6 (0.29 to 1.26)	74 (1 study)	⊕⊕⊖⊖ Low5
Reoperation after surgical treatment up to 60 months follow-up	94 per 1,000	59 fewer per 1,000 (from 85 fewer to 33 more)	RR 0.37 (0.1 to 1.35)	174 (2 studies)	⊕⊖⊖⊖ Very low4,5

CI: confidence interval; RR: Risk ratio; MD: mean difference
1 Evidence was downgraded by two due to performance bias (lack of blinding) and attrition bias.
2 Evidence was downgraded by one due to serious imprecision as 95%CI crossed one default MID
3 Evidence was downgraded by one due to lack of blinding.
4 Taking into account weighting in a meta-analysis and the likely contribution from each component, evidence was downgraded by one due to lack of blinding.
5 Evidence was downgraded by two due to very serious imprecision as 95%CI crossed two default MIDs.

The data provided by Wright 2005 comparing ablation with excision for pelvic pain associated with mild endometriosis demonstrated good symptom relief at 6 months for the majority of participants irrespective of the treatment modality. However, their data could not be included in the meta-analysis because the data were obtained using a ranked ordinal scale.

Table 118: Reduction in VAS score by 5 years after surgery (Healey 2014)

Outcome	Excision group, median (range)	Ablation group, median (range)	P-Value (Mann-Whitney U test)
Overall pain	5.8 (-3.4 to 10.0)	5.5 (-0.2 to 10.0)	0.46
Pelvic pain	6.2 (-2.6 to 9.3)	5.9 (-3.9 to 10.0)	0.81
Dyspareunia	6.0 (0 to 10.0)	3.2 (-4.3 to 10.0)	0.03

VAS: Visual Analogue Scale

Table 119: Change in quality of life: excision versus diagnostic laparoscopy at 6-month follow-up (Abbott 2004)

Outcome	DSG (mean (SD))	ISG (mean (SD))	DSG vs. ISG p-value (t-test)
EQ-5D index summary	0.74 (0.23)	0.77 (0.25)	0.07
EQ-5D VAS summary score	65.9 (21.3)	83.6 (10.8)	0.01
SF-12 physical component score	45.5 (10.0)	48.2 (7.6)	0.36
SF-12 mental component score	45.3 (11.8)	47.6 (9.7)	0.55

EQ-5D: EuroQol 5 dimensions questionnaire, DSG: Delayed Surgery Group, ISG: Immediate Surgery Group

11.3.2 Economic evidence

No health economic studies were found contrasting ablation to excisional surgery for endometriosis.

One RCT was found looking at the costs of ablation compared to hormonal treatment (Lalchandani, 2005). This found an expected saving for surgery over hormonal treatment of £595 per patient. However this trial did not consider the opportunity cost of the use of equipment or clinician time and so was not appropriate for inclusion in a NICE Guideline.

Four large database studies were found estimating the costs of laparoscopic surgery in different healthcare systems. These studies were Allaire (2014) in Canada, Chvatal (2010) in Germany and Fuldeore (2010, 2011) in the US. Together these studies incorporated 94,605 women with endometriosis. As these looked at cost rather than cost-effectiveness and were conducted in non-NHS settings, they were considered less appropriate as a source of costs

ISBN 978-1-4731-2661-9

than the NHS Reference Costs, but were included to serve as a source of variation for sensitivity analysis.

Table 120: Estimates of cost of laparoscopic surgery from different sources

Study	Population	Cost estimate in local currency	Cost estimate in 2016 GBP
Allaire (2014)	57,879 Canadian women recruited over 5 years receiving laparoscopic surgery	$1529.89 CAD	£949.09
Chvatal (2010)	20,835 German women receiving inpatient treatment of any kind for endometriosis	3056.12 EUR	£3189.23
Fuldeore (2010)	15,891 US women receiving therapeutic laparoscopy and 63,564 controls	$5886 USD	£5506.23
Fuldeore (2011)	As above – paper is re-analysis of Fuldeore (2010)	$6856 USD	£6096.35

EUR: Euro, CAD: Canadian Dollar, GBP: British Pound, USD: US Dollar

Finally 1 US study was found contrasting laparoscopic surgery to laparotomy (Luciano, 1992). This trial was excluded as laparotomy vs. laparoscopy was not a comparison of interest to the Committee and the data were very out of date, although it should be noted that the total cost for a laparoscopy was £3004, which is consistent with other estimates of the cost of the procedure.

Because of the importance of this question to the Committee, it was prioritised for de novo health economic modelling. The model found that – relative to no treatment – surgical interventions increased the average cost of treatment by £703.36 and average lifetime QALYs by 0.47. By typical cost-effectiveness standards, paying £1506.40 per QALY would be considered cost-effective, but the Committee considered evidence from the model suggesting that pairing a laparoscopic treatment with a more sensitive diagnostic test could reduce the cost/QALY relative to the same test with no treatment quite substantially. Table 121 demonstrates that in particular MRI and laparoscopic diagnosis reduce costs greatly compared to no treatment. The table also shows that – in general – laparoscopy and adjunct hormonal treatment costs slightly more and adds more QALYs than laparoscopy alone. In fact, laparoscopic diagnosis & laparoscopy with adjunct hormonal treatment adds the most possible lifetimes QALYs, since it pairs the most effective treatment with the most sensitive diagnostic strategy (empirical diagnosis and laparoscopic treatment adds nearly as many, because it has identical sensitivity but patients do not benefit from the therapeutic effects of a diagnostic laparoscopy described elsewhere). This indicates that there is always some cost-effectiveness threshold at which the NHS would consider this treatment, although the model describes how the NHS would only consider this treatment at cost-effectiveness thresholds of >£160,000.

ISBN 978-1-4731-2661-9

Table 121: Estimates of cost of laparoscopic surgery from different sources

Strategy	Lifetime Cost	Lifetime QALY	ICER vs same diagnostic strategy, no treatment
Empirical Diagnosis & No Treatment	£22,752.60	18.12	N/A
Empirical Diagnosis & Laparoscopy + Hormonal	£31,626.43	18.86	£12,034.14
Empirical Diagnosis & Laparoscopic Treatment	£28,052.06	18.47	£15,156.19
Pelvic MRI & No Treatment	£24,929.53	18.12	N/A
Pelvic MRI & Laparoscopy + Hormonal	£25,772.03	18.77	£1,288.34
Pelvic MRI & Laparoscopic Treatment	£24,783.78	18.42	-£478.24
Nerve fibre & No Treatment	£25,795.25	18.12	N/A
Nerve fibre & Laparoscopy + Hormonal	£26,875.57	18.78	£1,630.21
Nerve fibre & Laparoscopic Treatment	£26,222.93	18.46	£1,274.89
Laparoscopy & No Treatment	£35,933.66	18.14	N/A
Laparoscopy & Laparoscopy + Hormonal	£33,344.74	18.87	-£3,562.28
Laparoscopy & Laparoscopic Treatment	£31,899.07	18.52	-£10,637.82
Peritoneal biopsy & No Treatment	£25,362.71	18.12	N/A
Peritoneal biopsy & Laparoscopy + Hormonal	£27,422.18	18.79	£3,069.05
Peritoneal biopsy & Laparoscopic Treatment	£25,079.29	18.46	-£829.43
Transabdominal Ultrasound & No Treatment	£24,775.14	18.12	N/A
Transabdominal Ultrasound & Laparoscopy + Hormonal	£24,562.05	18.65	-£403.77
Transabdominal Ultrasound & Laparoscopic Treatment	£23,948.36	18.37	-£3,262.92
CA-125 & No Treatment	£25,201.29	18.12	N/A
CA-125 & Laparoscopy + Hormonal	£25,381.47	18.51	£467.87
CA-125 & Laparoscopic Treatment	£24,377.37	18.33	-£3,964.13
Pelvic MRI & No Treatment	£24,929.53	18.12	N/A
Pelvic MRI & Laparoscopy + Hormonal	£25,772.03	18.77	£1,288.34
Pelvic MRI & Laparoscopic Treatment	£24,783.78	18.42	-£478.24

ISBN 978-1-4731-2661-9

11.3.3 Clinical evidence statements

11.3.3.1 Endometriosis

11.3.3.1.1 Laparoscopic treatment (excision or ablation) versus diagnostic laparoscopy for endometriosis

Overall pain at 6 months

Very low quality evidence from 1 study of 69 women with endometriosis showed a clinically significant improvement in overall pain at 6 months associated with laparoscopic treatment compared with diagnostic laparoscopy for endometriosis.

Overall pain at 12 months

Low quality evidence from 1 study of 69 women with endometriosis found a clinically significant improvement in overall pain at 12 months associated with laparoscopic treatment compared with diagnostic laparoscopy for endometriosis.

11.3.3.1.2 Live birth or ongoing pregnancy

Very low quality evidence from 2 studies of 382 women found no clinically significant difference in live birth or ongoing pregnancy between laparoscopic treatment and diagnostic laparoscopy for endometriosis.

11.3.3.1.3 Clinical pregnancy

Very low quality evidence from 3 studies including 528 women with endometriosis found no clinically significant difference between laparoscopic treatment and diagnostic laparoscopy for the outcome of clinical pregnancy.

11.3.3.1.4 Miscarriage per pregnancy

Very low quality evidence from 2 studies including 112 women with endometriosis found no clinically significant difference between laparoscopic treatment and diagnostic laparoscopy for miscarriages per pregnancy.

11.3.3.1.5 Excision versus diagnostic laparoscopy for endometriosis

Overall pain at 6 months

High quality evidence from 1 study including 39 women with endometriosis found a clinically significant improvement in overall pain at 6 months associated with excision compared with diagnostic laparoscopy.

Overall pain score at 6 months

Very low quality evidence from 1 study including 16 women with endometriosis found a clinically significant reduction in overall pain score at 6 months associated with diagnostic laparoscopy compared with excision.

Overall pelvic pain score at 12 months

Moderate quality evidence from 1 study including 16 women with endometriosis found a clinically significant reduction in overall pain score at 12 months' follow-up associated with diagnostic laparoscopy compared with excision.

ISBN 978-1-4731-2661-9

Pelvic pain score at 6 months

Moderate quality evidence from 1 study including 39 women with endometriosis found no clinically significant difference in pelvic pain scores at 6 months associated with excision compared with diagnostic laparoscopy.

Dysmenorrhoea pain score at 6 months

Moderate quality evidence from 1 study including 39 women with endometriosis found that there was no clinically significant difference in dysmenorrhoea pain score at 6 months associated with excision compared with diagnostic laparoscopy.

Dyspareunia pain score at 6 months

Moderate quality evidence from 1 study including 39 women with endometriosis found that there was no clinically significant difference in dyspareunia pain score at 6 months associated with excision compared with diagnostic laparoscopy.

Health-related quality of life

Low quality evidence from 1 study including 39 women with endometriosis reported that there was no clinically significant difference in the mean EQ-5D index summary score at 6-month follow -up in the excision groups compared with the diagnostic laparoscopy group. Moderate quality evidence from the same study reported a clinically significant increase in the mean EQ-5D VAS summary score at 6 months associated with excision compared with diagnostic laparoscopy, but no clinically significant difference in the mean SF-12 physical and mental component scores at 6-month follow-up associated with excision compared with diagnostic laparoscopy.

11.3.3.2 Excisional surgery versus ablative surgery for endometriosis

Pain scores (improvement from baseline in VAS scores at 12 months)

Low to very low quality evidence from 1 randomised controlled trial comprising 103 women with endometriosis showed similar improvement in pain score in the laparoscopic excision and laparoscopic ablation groups for global pain as well as pelvic pain and dyspareunia at 12 months follow-up. One study reported the reduction in VAS score at 5-year follow-up, however, the clinical significance of reported outcomes could not be calculated.

Unintended effects of treatment (improvement from baseline in VAS score at 12 months follow up)

Moderate to low quality evidence from 1 randomised controlled trial comprising 103 women with endometriosis showed no clinically significant differences between the 2 treatments in nausea, vomiting and bloating at 12 months follow-up.

11.3.3.3 Endometrioma

Excisional surgery versus ablative surgery for endometrioma

Recurrence of pelvic pain

Moderate to low quality evidence from 2 randomised controlled trials with a total of 104 women with endometriosis showed clinically significant lower rates of recurrence of dysmenorrhoea and non-menstrual pelvic pain associated with laparoscopic excision when compared to laparoscopic ablation of endometrioma.

ISBN 978-1-4731-2661-9

Pregnancy rate after surgical treatment

Moderate quality evidence from 3 randomised controlled trials with a total of 138 women with endometriosis showed higher rates of pregnancy associated with laparoscopic excision compared to laparoscopic ablation after surgical treatment of endometrioma, but there is some uncertainty around this finding which makes judgment of clinical benefit unclear.

Recurrence of endometrioma (at 12 months and at 60 months)

High quality evidence from 4 randomised controlled trials with a total of 258 women with endometriosis showed lower rates of recurrence of endometrioma associated with laparoscopic excision when compared to laparoscopic ablation at 12 months follow up. However, this result did not reach clinical significance. Low quality evidence from 1 randomised controlled trial comprising 74 women with endometriosis showed similar rates of recurrence of endometrioma in the laparoscopic excision and laparoscopic ablation groups at 60 months follow-up.

Reoperation after surgical treatment (up to 60 months)

Very low quality evidence from 2 randomised controlled trials comprising together of 174 women with endometriosis showed higher rates of reoperations associated with laparoscopic excision when compared to laparoscopic ablation up to 60 months follow up. However, this result did not reach clinical significance.

11.3.4 Evidence to recommendations

11.3.4.1 Relative value placed on the outcomes considered

As pain relief is the primary reason for patients seeking treatment, this was the most critical outcome for this review. Health-related quality of life was also critical as this might be considered to give a more broad reflection of patient experience than pain relief alone.

Rate of success, surgical complications, satisfaction with treatment, effect on daily activities, absence from work, number of women requiring more surgery and reduction in size and extent of endometriotic cysts were considered important outcomes as they were less clear indicators of effectiveness.

11.3.4.2 Consideration of clinical benefits and harms

Throughout the care pathway, the Committee stressed the importance of a full discussion with women of their symptoms and priorities with respect to pain and fertility. This was particularly important in gynaecology services and specialist endometriosis services (endometriosis centres) when discussing the benefits and harms of laparoscopic surgery. Such a discussion should highlight the potential negative impact of laparoscopic treatment on ovarian reserve.

The Committee recognised that a woman might be referred from a GP for a consultation with a general gynaecologist, a gynaecologist with a specialist interest or at a specialist centre and noted that women with suspected rectovaginal endometriosis would require the expertise available at a specialist centre.

The Committee discussed what a referral would provide for a woman and agreed that the gynaecologists would firstly discuss the woman's symptoms and priorities with her and what her treatment options would be. For example, her primary symptom could be pain in which case offering alternative hormonal therapy to that offered by the GP might be appropriate - the type of hormone and its duration of effect being determined on an individual basis

ISBN 978-1-4731-2661-9

considering the woman's preferences. However, this treatment might not be appropriate if the woman's primary concern was fertility.

The Committee discussed whether a diagnostic laparoscopy should be offered prior to further management and concluded that the decision for a diagnostic laparoscopy would be on individual symptoms and priorities (and may require a further referral). The Committee agreed that diagnostic laparoscopy is a valuable tool which provides the most accurate diagnosis and also provides the opportunity to treat. The Committee noted that once diagnosed (either by laparoscopy or incidental other confirmatory findings from ultrasound, MRI or biomarkers), the most suitable long-term treatment options can then be discussed with the women with the aim to tailor these to her needs and priorities.

The Committee agreed that if a diagnostic laparoscopy was performed and minor endometriosis was found, it should be treated during the laparoscopy by a suitably trained surgeon. To describe the minor type endometriosis they agree that there were 2 types that could be treated at the same time. Peritoneal endometriosis not involving the bowel, bladder or ureter and uncomplicated endometriomas. They intentionally used 'uncomplicated' to allow for clinical judgement, based on the surgeon's skill and experience, since it would be difficult to define all possible complex cases of endometriomas. Treatment of an uncomplicated endometrioma could be performed at the initial procedure by a suitably trained surgeon, but decision-making would be influenced by the findings at laparoscopy; more extensive surgery, for example, treatment of a large endometrioma or bilateral endometriomas would not be performed at the time of diagnostic laparoscopy. It was noted that for instance an endometrioma that is adherent to its surrounding structures may be a complex procedure and may require further surgery after referral to a specialist centre. That the diagnostic laparoscopy may include treatment should be agreed with women prior to the procedure. Therefore the discussion with the woman was key to guide surgical decision-making. The Committee further noted that surgical diagnosis with treatment might not be suitable for all women (e.g. young women with mild disease) and that therapeutic treatment at laparoscopy would only be performed with mild or moderate endometriosis and not if there was extensive disease.

Excisional treatment was recommended over ablative treatment as the evidence showed that there was lower risk of recurrence of endometrioma and the Committee suggested that ablative surgery had a greater negative impact on ovarian reserve.

In reviewing pre-operative pharmacological treatment, the Committee noted the limitations of the evidence and discussed whether there was a role for hormonal therapies in women with severe (deep infiltrating) endometriosis and felt that this should be considered, based on their surgical expertise and experience, and on discussion with the woman to ensure that she understands the possible benefits, risks and complications of the treatment (particularly highlighting the side effect profile of GnRH agonists). The Committee acknowledged that there was another school of thought that hypothesises that this adjunct treatment could lead to co-occuring superficial endometriosis being missed. However, consensus was reached that this prior treatment would facilitate surgery (by reducing bleeding and inflammation) and therefore reduce reoperation rate.

11.3.4.3 Consideration of economic benefits and harms

The difference between endometrial excision and ablation surgery is highly unlikely to carry a significant cost, as the most significant cost of the operations is not the technique itself but the cost of the support network required to employ the technique – the surgical time, operating theatre use and recovery time. The Committee believed that these would be similar for both techniques and that any difference between the 2 techniques would come down to individual patient / disease characteristics (such as location of the endometriosis), or possibly the familiarity of the surgeon with a particular piece of equipment. Consequently health economic evidence was not used to inform the discussion of ablation vs. excision.

ISBN 978-1-4731-2661-9

The difference between diagnostic, therapeutic and no laparoscopy was considered sufficiently important to warrant de novo economic modelling. Details of this model are available in Appendix K.

The difference in cost between diagnostic and therapeutic laparoscopy is not strictly relevant to health economic analysis as they are not competing alternatives – the NHS could offer one, both or neither. Additionally, the Committee suggested it would be very common to offer minor therapeutic surgery during a 'diagnostic' laparoscopy and a therapeutic laparoscopy would – by definition – require a diagnosis of the pathology that was the target of the surgery. Consequently in health economic terms the distinction between the 2 forms of surgery is a little artificial.

Committee members suggested that there might be some value in a diagnostic laparoscopy that went beyond the placebo effect, for example, receiving a definitive diagnosis might have positive psychological consequences. Although the economic model tries to account for this in sensitivity analysis, it is likely the economic benefit of a diagnostic laparoscopy will vary depending on the other potential diagnoses a woman might be considering and the value she places on knowing her condition for certain. The Committee took this fact into account when making recommendations, arguing that although the diagnostic health economic model typically found laparoscopy fell outside the range that NICE would typically pay for, the fact that the laparoscopy had other benefits, could be used to rule out malignancy and was anyway required for a therapeutic laparoscopy justified its inclusion.

Laparoscopic treatment (with or without subsequent hormonal treatment) is the gold-standard for treating endometriosis. Consequently there is always some willingness-to-pay threshold at which laparoscopic treatment becomes cost-effective. In general, NICE considers treatments more than £20,000 / QALY to be poor candidates for being cost-effective, and so the Committee observed that whether therapeutic laparoscopy was cost-effective or not depended on whether the woman was able to take hormonal therapy (and whether the treatment was having any positive effect). If the woman was currently on hormonal treatment and this was improving her symptoms relative to no treatment, the cost per QALY of operating was just above £20,000 / QALY and might be regarded as borderline cost-effective. But if the woman could not tolerate hormonal and neuromodulator treatment or was not receiving any benefit from the therapy the cost per QALY of operating was around £14,000 / QALY, which would normally be considered cost-effective. The Committee noted that the situation where this was most likely to occur was where a woman was trying to conceive and therefore could not take contraceptives.

Surgery for endometriosis that is not well controlled with other treatments is extremely common and consequently it is not thought that the Committee's recommendations will lead to a substantial change of resources.

11.3.4.4 Quality of the evidence

The quality of the evidence used to make recommendations on combined surgery plus hormonal treatments for pain relief was generally moderate. Although the majority of studies were appropriately blinded, they rarely reported appropriate allocation concealment or details of the randomisation procedure. Two of the 4 studies in the NMA did not report measures of variability or uncertainty in their estimates, which meant that statistical imputation of missing information was needed. However, a variety of sensitivity of analyses were performed to test assumptions made during modelling and the results seemed robust.

For comparison between different surgical techniques the quality of the evidence was very low. The Committee discussed the difficulty of conducting high quality randomised studies, particularly as randomising patients to either excisional or ablative laparoscopic treatment can be impractical especially where there is deep endometriosis affecting bowel, bladder and ureter.

ISBN 978-1-4731-2661-9

Evidence of the effectiveness of hormonal treatment combined with surgery only came from studies that followed surgery with hormonal treatment.

11.3.4.5 Other considerations

One of the key considerations throughout treatment for pain relief in endometriosis is women's fertility. Fertility may be a strongly influencing factor in many women's treatment choices and a timely discussion on how different treatments will impact this is essential. The Committee suggested that a particular point to highlight in such a discussion is that laparoscopic treatment (ablation or excision) of ovarian endometrioma may negatively affect ovarian reserve.

The different treatment options recommended here are based on RCT evidence from a number of different studies, which was in agreement with the experience of the Committee. Recommendations on information provision and the pathway of care were developed primarily from Committee experience and opinion, supported in part by the literature.

The Committee was aware of an ongoing trial investigating the effectiveness of post-surgical hormonal treatments. They agreed that the results of the NMA were consistent with their experience that hormonal treatments after surgery delay recurrence of endometriosis. However, they also noted that there was still some uncertainty around the size of the effect and that results from the ongoing trial would be important by adding to the evidence base and thus informing future guidance.

The Committee considered whether any additional recommendations were necessary for adolescent women. It was concluded that none were required, but that it was important to highlight that treatment options may be different for these women and that there was an even greater need to minimise repeat surgery in this population.

Although there were no studies that looked specifically at pre-surgical hormonal treatment, the Committee felt that this may often be the case in practice, as women may have been receiving ovulation suppression for many months/years prior to surgery. The Committee concluded that there was insufficient evidence to recommend hormonal therapy as a standard treatment prior to surgery although they acknowledged that that there may be some benefit for some women with deep endometriosis involving the bowel, bladder or ureter, as based on their clinical experience and knowledge, pre-operative GnRH agonists can reduce surgical complications such as bleeding. However, the Committee agreed that the decision to use GnRH agonists pre-operatively should be made on an individual level based on surgeon and patient preference and informed consent.

The Committee also discussed whether this topic should be prioritised for a research recommendation. They decided that there was a gap in the evidence with regards to the effectiveness of ablation or excision related to peritoneal endometriosis. The research recommendations are provided below.

The Committee discussed whether to cross-refer to recommendations in the laparoscopic uterosacral nerve ablation (LUNA) for chronic pelvic pain NICE interventional procedure guideline (IPG). They agreed that there was considerable uncertainty about the conclusions of this IPG for women with endometriosis and therefore did not feel sufficiently confident to refer to this.

11.3.4.6 Key conclusions

The Committee concluded that clinicians should discuss with women whether they would like uncomplicated endometriomas or peritoneal endometriosis to be treated if found during diagnostic laparoscopy. The discussion should highlight the potential risks and benefits of the laparoscopy and allow women to make an informed choice regarding their treatment.

ISBN 978-1-4731-2661-9

As there was evidence that post-surgical hormonal therapy gave additional benefit over surgery alone, the Committee recommended that this be offered after surgery. Although there was no evidence available regarding the use of GnRH agonists prior to surgery, the Committee agreed that a recommendation should be made to support this because based on their experience and knowledge, pre-operative GnRH agonists can reduce surgical complications such as bleeding. The decision to use GnRH agonists pre-operatively should be made on an individual patient basis and only in severe deep disease

11.3.5 Recommendations

40. Ask women with suspected or confirmed endometriosis about their symptoms, preferences and priorities with respect to pain and fertility, to guide surgical decision-making.

41. Discuss surgical management options with women with suspected or confirmed endometriosis. Discussions may include:

- what a laparoscopy involves
- that laparoscopy may include surgical treatment (with prior patient consent)
- how laparoscopic surgery could affect endometriosis symptoms
- the possible benefits and risks of laparoscopic surgery
- the possible need for further surgery (for example, for recurrent endometriosis or if complications arise)
- the possible need for further planned surgery for deep endometriosis involving the bowel, bladder or ureter.

42. Perform surgery for endometriosis laparoscopically unless there are contraindications.

43. During a laparoscopy to diagnose endometriosis, consider laparoscopic treatment of the following, if present:

- peritoneal endometriosis not involving the bowel, bladder or ureter.
- uncomplicated ovarian endometriomas.

44. As an adjunct to surgery for deep endometriosis involving the bowel, bladder or ureter, consider 3 months of gonadotrophin-releasing hormone agonists[b] before surgery.

45. Consider excision rather than ablation to treat endometriomas, taking into account the woman's desire for fertility and her ovarian reserve. Also see ovarian reserve testing in the NICE guideline on fertility problems.

11.3.6 Research recommendations

4. Is laparoscopic treatment (excision or ablation) of peritoneal disease in isolation effective for managing endometriosis-related pain?

b At the time of publication (September 2017), not all gonadotrophin-releasing hormone agonists have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Prescribing guidance: prescribing unlicensed medicines for further information.

ISBN 978-1-4731-2661-9

Why this is important

Isolated peritoneal endometriosis can be an incidental finding in women who may or may not experience pain or other symptoms.

Research is needed to determine whether laparoscopic treatment of isolated peritoneal endometriosis in women with endometriosis-related pain results in a clinical and cost-effective improvement in symptoms.

The current literature does not provide a clear answer because the stage of endometriosis is often not sufficiently clearly defined in research studies, and the treatment modalities used are multiple and varied. The resultant amalgamation of various stages of endometriosis and variable treatment modalities leads to loss of certainty of outcome in this specific group of women.

Establishing whether treating isolated peritoneal endometriosis is cost effective is important, because this forms a large part of the workload in general gynaecology, and uses considerable resources.

Table 122: Research recommendation rationale

Research question	Is laparoscopic treatment of peritoneal disease in minimal and mild endometriosis cost-effective for the management of suspected endometriosis-associated pain?
Importance to 'patients' or the population	This is important as all surgery carries with it a potential morbidity and mortality. Thus any surgical interventions needs to have a likelihood of relief of symptoms to be clinically justified.
Relevance to NICE guidance	The answer to this question has not been able to be determined from the current literature available. The impact of surgery on this subset of women with endometriosis needs to be known so that the cost-effectiveness of surgery can be determined.
Relevance to the NHS	Pain associated with endometriosis costs the NHS significant amounts of money. In addition surgical time and bed usage are limited resources so that cost-effective utilisation is essential.
National priorities	This is a large group of women who require evidence based care.
Current evidence base	Not available regards this group in relation to pain outcomes.
Equality	A study population should include the full age spectrum of women who suffer endometriosis associated pain. Adolescents as well as adults will need to be studied.
Feasibility	This has been done in relation to fertility outcomes and hence the same research model could be adopted.
Other comments	Not applicable

Table 123: Research recommendation PICO table

Criterion	Explanation
Population	Women with proven isolated peritoneal disease which is classed as minimal or mild endometriosis. This should be determined at diagnostic laparoscopy and the position and extent of the endometriosis described as accurately and fully as possible.
Intervention	Complete laparoscopic excision of all peritoneal endometriosis, with histological confirmation.
Comparator	Laparoscopy without excision of any endometriosis
Outcome	Standardised patient symptom questionnaire and Quality of Life using a validated measurement system at 6 months and annually after surgery. Secondary outcomes would include additional surgical and/or medical treatment required by the patient in the follow up interval.

ISBN 978-1-4731-2661-9

Criterion	Explanation
Study design	Randomised controlled trial, ideally with participant blinding to treatment allocation.
Timeframe	Two years of randomisation and 2 years of follow up, providing recruitment numbers are sufficient to achieve population numbers of sufficient size to answer the research question.

11.3.7 Pairwise comparison of combinations of treatments

Review question: What is the effectiveness of hormonal treatment before or after surgery for treatment of endometriosis?

11.3.7.1 Description of clinical evidence

The aim of this review is to determine the clinical and cost effectiveness of pharmacological therapy in combination with surgery in women with endometriosis. Pharmacological therapy specifically included hormonal suppression treatments available in the UK and 4 comparisons are examined:

- pharmacological therapy before surgery vs. placebo or no pharmacological therapy before surgery
- pharmacological therapy after surgery vs. placebo or no pharmacological therapy after surgery
- pharmacological therapy before surgery vs. pharmacological therapy after surgery
- pharmacological therapy before and after surgery vs. placebo or no pharmacological therapy before and after surgery

For full details see review protocol in Appendix D.

In total 12 studies were included, but the evidence of these studies only addressed 1 of the 4 possible comparisons.

No studies were identified for inclusion for the following 3 comparisons:

- pharmacological therapy before surgery vs. placebo or no pharmacological therapy before surgery
- pharmacological therapy before surgery vs. pharmacological therapy after surgery
- pharmacological therapy before and after surgery vs. placebo or no pharmacological therapy before and after surgery

For the comparison 'pharmacological therapy after surgery vs. placebo or no pharmacological therapy after surgery' 12 studies were included in total (Abou-Setta 2013, Alborzi 2011, Bianchi 1999, Busacca 2001, Furness 2011, Loverro 2008, Mettler 2014, Muzii 2000, Parazzini 1994, Serrachioli 2010, Sesti 2007, Sesti 2009).

Two were Cochrane systematic reviews (Abou-Setta 2013 and Furness 2011) and the remaining 10 studies were randomised controlled trials (Alborzi 2011, Bianchi 1999, Busacca 2001, Loverro 2008, Mettler 2014, Muzii 2000, Parazzini 1994, Serrachioli 2010, Sesti 2007, Sesti 2009). The full text of all trials included in the Cochrane reviews were considered for inclusion according to the protocol. Additional outcomes from 6 trials were also included in this review (Bianchi 1999, Busacca 2001, Loverro 2008, Muzii 2000, Parazzini 1994, Sesti 2007).

Of the remaining 4 trials, one reported outcomes relevant to this review (Sesti 2009) and 3 were published subsequently to the searches performed within the Furness 2011 Cochrane review (Alborzi 2011, Mettler 2014, Serrachioli 2010).

ISBN 978-1-4731-2661-9

All studies included women who had confirmed endometriosis and who had undergone surgery prior to being randomised to hormonal suppression treatment compared to no treatment or placebo. Available details of the surgery performed are noted in Table 124.

The post-surgical hormonal suppression treatments in the intervention arms were gonadotropin-releasing hormone agonists (GnRHa) (including leuprorelin, triptorelin, goserelin, nafarelin, and decapeptyl), letrozole, combined oral contraceptives, medroxyprogesterone acetate and danazol and the 2 trials in the Abou-Setta 2013 Cochrane review used a levonorgestrel-releasing intrauterine device (LGN-IUS).

Evidence was available for the critical outcomes of pain relief and health related quality of life. Evidence was available for the important outcome of rate of success (subsequent reoperation rate and disease recurrence) and participant satisfaction with treatment. No evidence was available for effect on daily activities or number of live births. Evidence relating to fertility is covered in the Network Meta-Analysis (NMA).

Evidence from the included studies are summarised in the clinical GRADE evidence profile below (Table 124).

Stratified analysis according to the specifications in the protocol was not possible due to the presentation of the available data.

See also the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G. Summary of included studies

A summary of the studies that were included in this review are presented in Table 124.

Table 124: Summary of included studies

Study	Intervention/Comparison	Population	Outcomes
Abou-Setta 2013 (CSR)	Postoperative insertion of the LNG-IUS versus • no postoperative treatment, • placebo (inert IUD), • or any other active systemic treatment	2 Trials comparing insertion of the LNG-IUS versus no postoperative treatment, placebo (inert IUD), or any other active systemic treatment in women undergoing surgery for endometriosis Tanmahasamut 2012 trial – using ASRM staging. • 10 women stage 1 • 7 women stage 2 • 8 women stage 3 • and 29 women stage 4 Vercellini 2003 trial – women were AFS stages 1-4	• Recurrence of painful periods • Patient satisfaction with results as described by women
Alborzi 2011 Iran	Letrozole for 2 months and triptorelin for 2 months (2 arms) versus no treatment	Women who had been infertile for at least 12 months and some of whom had symptoms (dysmenorrhoea,	• Pain recurrence at 12 months • Endometriosis at 12 months

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes
		dyspareunia and pelvic pain. • 65 women were AFS stages 1&2 • and 59 women were AFS stages 3&4	
Bianchi 1999 Italy	Surgery: Cook and Rock technique of laparoscopy (conservative surgery) was used Pharmacological comparison: Danazol, 600 mg/day versus no medical therapy for 3 months.	Inclusion criteria: < 40 yrs. • All women were AFS stage 3 (N=65) • or AFS stage 4 (N=12) Exclusion criteria: medical or surgical treatment for endometriosis, concurrent disease that might affect fertility or cause pelvic pain, women without pain symptoms, women not seeking pregnancy, liver or endocrine disease N randomised = 77 N analysed = 77	Included in Furness 2011 and additionally reported • Reoperation
Busacca 2001 Italy	Surgery: Cook and Rock technique of laparoscopy (conservative surgery) was used Pharmacological comparison: GnRHa (leuprolide) versus no medical therapy every 4 weeks for a period of 12 weeks	Inclusion criteria: < 40 years, laparoscopic diagnosis of endometriosis • ASRM stage 3 N=59 or stage 4 • N=30 Exclusion criteria: previous medical or surgical therapy for endometriosis, other diseases that might affect fertility or cause pelvic pain; liver, endocrine or neoplastic disease N randomised = 89 N analysed = 89	Included in Furness 2011 and additionally reported • Reoperation
Furness 2011 (CSR)	All systemic medical treatments for the hormonal suppression of endometriosis including GnRHas, danazol, progestogens, gestrinone or the oral contraceptive pill (or combinations of these) administered after surgery to no medical	Trials (N=12) The study population included women of reproductive age who were undergoing surgery for endometriosis	Pain recurrence (VAS) • Pelvic pain • Dysmenorrhoea • Deep Dyspareunia Pain recurrence • at 12 months • at 13-24 months • at 60 months Endometriosis • at 12 months

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes
	treatment, or placebo were studied. The use of medical therapy was considered at any dosage and for a period of at least 3 months duration before or after surgery. Only agents used with the aim of hormonal suppression were included.		• at 24 months Endometrioma • at 13-36 months • at 5 years
Loverro 2008 Italy	Surgery: Laparoscopic diathermy, laser vaporisation or surgical excision of endometriomas Pharmacological comparison: Triptorelin depot versus placebo over a 3 month period	Inclusion criteria: women of reproductive age with stage III - IV endometriosis, associated with chronic pelvic pain, adnexial mass or infertility, who had undergone complete laparoscopic excision, had rAFS score > 15 and no previous hormonal treatment • AFS stage 3 N=33 • AFS stage 4 N=21 N randomised = 60 N analysed = 54	Included in Furness 2011 and additionally reported • Endometrioma recurrence at 5 years
Mettler 2014 Germany	Surgery: laparoscopic excision of endometrial foci, removal of adhesions and restoration of normal reproductive anatomy. Ureter and superficial bowel lesions were removed Pharmacological comparison: Leuprorelin depot subcutaneously injected monthly or no treatment over a 3 month period	Inclusion criteria: Women with symptomatic endometriosis (18-44 years old) in whom 2 consecutive laparoscopic interventions were to be assessed. • EEC stage 0, N=0 • EEC stage I, N=185 • EEC stage II, N=127 • EEC stage III, N=85	• Pain recurrence (questionnaire based) at 12 months post treatment completion: - abdominal pain - dysmenorrhoea - dyspareunia • Disease recurrence at 5-6 months
Muzii 2000 Italy	Surgery: Laparoscopic excision of ovarian endometriomas with drainage, adhesionolysis or bipolar coagulation if necessary Pharmacological comparison: Cyclic monophasic combined oral contraceptives versus	Inclusion criteria: 20-35 years, moderate to severe dysmenorrhoea and/or chronic pelvic pain, not desiring fertility. Mean AFS scores 43.4 SD 22.3 in treatment group and 46.1 SD 23.9 in control group. Exclusion criteria: treatment for	Included in Furness 2011 and additionally reported • Endometrioma recurrence at 13-36 months

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes
	no medical therapy for 6 months	endometriosis in previous 6 months. N randomised = 70 N analysed = 68	
Parazzini 1994 Italy	Surgery: Laparotomy as first surgical treatment for debulking or radical surgery of endometriotic lesions Pharmacological comparison: Intranasal nafarelin (400 µG/day) versus placebo over a period of 3 months	Inclusion criteria: age < 38 yrs, normal medical examination, unexplained infertility for at least 1 year, with/without chronic pelvic pain, endometriosis AFS stage III-IV, partners with normal sperm analysis and post-coital tests. • AFS stage 3, N=37 • AFS stage 4, N=28 Exclusion criteria: previous laparoscopic/clinical diagnosis of endometriosis, other diseases that might cause infertility or pelvic pain, previous treatment for endometriosis or infertility N randomised =75 N analysed (pain scores) =68	Included in Furness 2011 and additionally reported • Pelvic pain recurrence (Andersch and Milsom*)
Seracchioli 2010 Italy	Surgery: Laparoscopic excision of ovarian endometriomas using the classic stripping technique Pharmacological comparison: 2 groups using continuous low dose monophasic oral contraceptives and cyclic therapy (combined in this analysis) vs. no treatment for 24 months	Inclusion criteria: Nulliparous women (20-40 years old) not attempting to conceive at study entre of for at least 2 years post-surgery. No previous surgical or medical treatment of endometriosis and no receipt of oral contraceptives for at least 6 months prior to surgery. • AFS stage 3, N=99 • AFS stage 4, N=118	• Endometrioma recurrence at 12 months post treatment completion (24 months)
Sesti 2007 Italy	Surgery: Conservative pelvic surgery Pharmacological comparison: GnRHa (either triptorelin or leuprorelin) or continuous oestroprogestin	Inclusion criteria: women of reproductive age <40, with endometriosis related symptoms (dysmenorrhoea, pelvic pain, deep dyspareunia),	Included in Furness 2011 and additionally reported • Quality of life using SF-36 (Results presented in graph - narrative

ISBN 978-1-4731-2661-9

Study	Intervention/Comparison	Population	Outcomes
	(cOCP) versus placebo for 6 months	laparoscopic diagnosis of St III -IV endometriosis, desiring pregnancy, nulliparous. • AFS 3, N=100 • AFS stage 4, N=87 Exclusion criteria: concurrent disease, such as cancer or pelvic inflammatory disease, previous surgery for endometriosis, contraindications to estrogens/progestins N randomised = 234 N analysed = 222	interpretation given in this review)
Sesti 2009 Italy	Surgery: Laparoscopic removal of endometriomas with enucleation of the entire cyst and stripping from the normal ovarian tissue and with drainage, adhesionolysis and bipolar coagulation if necessary Pharmacological comparison: Tryptorelin or leuprorelin and continuous low dose monophasic oral contraceptives (2 arms) vs. placebo for 6 months	Women of reproductive age, up to 40 years at time of surgery, US evidence of endometrioma, moderate to severe endometriosis, laparoscopic diagnosis of endometrioma first laparoscopic surgery for endometriosis, conservative treatment, complete excision of all evident ovarian and peritoneal disease, UC and clinical follow up after surgery. • AFS stage I, N=26 • stage 2, N=71 • stage 3, N=53 • stage 4, N=28	• Endometrioma: • at 13-36 months • Reoperation

CSR: Cochrane systematic review; N: number of participants in study
** Pelvic pain was assessed using Andersch and Milsom's multidimensional verbal rating scale, which defines pain according to limitation of ability to work (unaffected, 0; rarely affected, 1; moderately affected, 2; clearly inhibited, 3), presence of systemic symptoms (absent, 0; present, 1), and need for analgesics (no, 0; rarely, 1; regularly, 2). The score of each dimension is added to provide an overall summary score; AFS: American Fertility Society Score*

11.3.7.2 Clinical evidence profile

The clinical evidence profile for this review question (Post-surgical pharmacological therapy versus placebo or no treatment) is presented in Table 125.

ISBN 978-1-4731-2661-9

Table 125: Summary clinical evidence profile for Comparison: Pharmacological therapy after surgery vs. placebo or no pharmacological therapy after surgery

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Control	Post-surgical pharmacological therapy			
Pain recurrence (VAS)cm - Pelvic pain Follow-up: 12 months	Control group mean 6.2 (SD 0.9)	The mean pain recurrence (VAS) - pelvic pain in the intervention groups was 1.2 lower (1.47 to 0.93 lower)	MD -1.2 (-1.47 to -0.93)	187 (1 study)	⊕⊕⊕⊖ Moderate1
Pain recurrence (VAS) cm-Dysmenorrhoea Follow-up: 12 months	Control group mean 6.4 (SD 1.3)	The mean pain recurrence (VAS) - dysmenorrhoea in the intervention groups was 0.7 lower (1.04 to 0.36 lower)	MD -0.7 (-1.04 to -0.36)	187 (1 study)	⊕⊕⊖⊖ Low1,2
Pain recurrence (VAS) cm - Deep dyspareunia Follow-up: 12 months	Control group mean 4.8 (SD 1.2)	The mean pain recurrence (VAS) - deep dyspareunia in the intervention groups was 0.4 lower (0.76 to 0.04 lower)	MD -0.4 (-0.76 to -0.04)	187 (1 study)	⊕⊖⊖⊖ Very low1,3
Pain recurrence (questionnaire based) - Abdominal pain at 12 months post treatment completion	569 per 1,000	404 per 1,000 (279 to 586)	RR 0.71 (0.49 to 1.03)	120 (1 study)	⊕⊕⊖⊖ Low2,4
Pain recurrence (questionnaire based) - Dysmenorrhoea at 12 months post treatment completion	346 per 1,000	301 per 1,000 (190 to 471)	RR 0.87 (0.55 to 1.36)	158 (1 study)	⊕⊖⊖⊖ Very low3,4
Pain recurrence (questionnaire based) - Dyspareunia at 12 months post treatment completion	304 per 1,000	161 per 1,000 (85 to 301)	RR 0.53 (0.28 to 0.99)	144 (1 study)	⊕⊕⊖⊖ Low2,4
Pain recurrence (Andersch and Milsom) - Pelvic pain Follow-up: 12 months	Control group mean 4 (SD 3.6)	The mean pain recurrence (Andersch and Milsom) - pelvic pain in the intervention groups was 0.4 lower (2.15 lower to 1.35 higher)	MD -0.4 (-2.15 to 1.35)	53 (1 study)	⊕⊕⊖⊖ Low3

ISBN 978-1-4731-2661-9

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
Pain recurrence (dichotomous) Follow-up: 12 months	216 per 1,000	168 per 1,000 (119 to 241)	RR 0.78 (0.55 to 1.12)	476 (4 studies)	⊕⊖⊖⊖ Very low2,5
Pain recurrence (dichotomous) Follow-up: 13-24 months	286 per 1,000	200 per 1,000 (134 to 294)	RR 0.7 (0.47 to 1.03)	312 (3 studies)	⊕⊖⊖⊖ Very low2,6
Pain recurrence (dichotomous) Follow-up: 60 months	480 per 1,000	446 per 1,000 (254 to 797)	RR 0.93 (0.53 to 1.66)	54 (1 study)	⊕⊖⊖⊖ Very low3,7
Dysmenorrhoea Follow-up: 12 months	383 per 1,000	84 per 1,000 (31 to 230)	RR 0.22 (0.08 to 0.6)	95 (2 studies)	⊕⊕⊕⊖ Moderate8
Reoperation (women with endometriosis)	30 per 1,000	35 per 1,000 (12 to 101)	RR 1.17 (0.4 to 3.4)	327 (3 studies)	⊕⊖⊖⊖ Very low3,9
Endometriosis recurrence (dichotomous) - Disease recurrence at 5-6 months Follow-up: 5-6 months	401 per 1,000	397 per 1,000 (301 to 530)	RR 0.99 (0.75 to 1.32)	285 (1 study)	⊕⊖⊖⊖ Very low3,4
Endometriosis recurrence (dichotomous) Follow-up: 12 months	70 per 1,000	101 per 1,000 (20 to 515)	RR 1.44 (0.28 to 7.36)	310 (3 studies)	⊕⊖⊖⊖ Very low3,10,11
Endometriosis recurrence (dichotomous) Follow-up: 24 months	133 per 1,000	29 per 1,000 (1 to 500)	RR 0.22 (0.01 to 3.75)	45 (1 study)	⊕⊖⊖⊖ Very low3,12
Endometrioma recurrence (dichotomous) - Recurrence at 13-36 months	189 per 1,000	104 per 1,000 (68 to 163)	RR 0.55 (0.36 to 0.86)	463 (3 studies)	⊕⊕⊖⊖ Low2,13,14
Endometrioma recurrence (dichotomous) Follow-up: 60 months	125 per 1,000	210 per 1,000 (44 to 1,000)	RR 1.68 (0.35 to 8.03)	35 (1 study)	⊕⊕⊖⊖ Low3,7
Patient Satisfaction	Not estimable	Not estimable	RR 1.21 (0.80 to 1.82)	95 (2 studies)	See comment

1 Blinding: unclear risk. Placebo is not described and seems unlikely that blinding could be maintained when the interventions are depot and oral hormonal treatments
2 95% Confidence Interval crosses 1imprecision threshold
3 95% Confidence Interval crosses 2 imprecision thresholds
4 Randomisation, Allocation concealment: unclear risk. No information provided. Blinding: High risk. No placebo used

ISBN 978-1-4731-2661-9

5 Allocation concealment: unclear risk. Not mentioned in Alborzi 2011, Loverro 2001 or Bianchi 1999. Blinding: high risk. No placebo used in Alborzi 2011, Bianchi 1999 or Vercellini 1999. Incomplete data reporting: unclear risk. 22% withdrawal overall in Vercellini 1999 due to reasons other than symptom recurrence or major protocol violations (similar in each group). 18% withdrawal overall in Alborzi 2011 after randomisation due to "poor patients follow up" with reasons not reported and unequal loss across groups(11/58 letrozole group, 18/58 dipherelin group and 1/59 no treatment group)
6 Allocation concealment: unclear risk. Not mentioned in Busacca 2001 or Muzii 2000. Blinding: high risk. No placebo use in Busacca 2001, Muzii 2000 or Vercellini1999. Incomplete data reporting: unclear risk. 22% withdrawal overall in Vercellini 1999 due to reasons other than symptom recurrence or major protocol violations (similar in each group). Other bias: unclear risk. No baseline characteristics reported in Muzii 2000
7 Allocation concealment: unclear risk. Not mentioned.
8 Blinding: unclear/high risk of performance bias. Unclear how patients were blinded to IUD presence in Tanmahasamut 2012 and Vercellini 2003 reported as an open label study with outcome assessors not blinded to treatment group (high risk of detection bias).
9 Allocation concealment: unclear risk. Not mentioned in Bianchi 1999, Busacca 2001 or Sesti 2009. Blinding: high risk. No placebo use in Bianchi 1999 or Busacca 2001.
10 Allocation concealment: unclear risk. Not mentioned in Alborzi 2011, Bianchi 1999 or Busacca 2001. Blinding: high risk. No placebo used in Alborzi 2011, Bianchi 1999 or Busacca 2001. Incomplete data reporting: unclear risk. 18% withdrawal overall in Alborzi 2011 after randomisation due to "poor patients follow up" with reasons not reported and unequal loss across groups (11/58 letrozole group, 18/58 dipherelin group and 1/59 no treatment group)
11 Using random effects model. Heterogeneity: $Chi^2 = 5.72$, $df = 2$ ($P = 0.06$); $I^2 = 65\%$. Removal of Alborzi 2011 (RR = 16.48 95%CI 0.99 - 272.92) from the pooled analysis removes inconsistency (Heterogeneity: $Chi^2 = 0.38$, $df = 1$ ($P = 0.54$); $I^2 = 0\%$) and the pooled fixed effects result for Bianchi 1999 and Busacca 2001 becomes RR = 0.76 (95%CI 0.30 - 1.90)
12 Blinding: high risk. No placebo used. Incomplete data reporting: high risk. 4/15 (27%) loss to follow up in treatment group in Tsai 2004.
13 Allocation concealment: unclear risk. Not mentioned in Muzii 2000 or Sesti 2009. Blinding: unclear risk - no placebo use in Muzii 2000 or in Seracchioli 2010 (although outcome assessors were blinded to treatment group. Incomplete data reporting: unclear risk. 8% withdrawal overall in relevant treatment arms in Sesti 2009. Other bias: unclear risk. No baseline characteristics reported in Muzii 2000
14 Using fixed effects model Heterogeneity: $Chi^2 = 3.25$, $df = 2$ ($P = 0.20$); $I^2 = 39\%$

Table 126 narratively reported results

Study ID	Hormone treatment groups (GnRH agonist or estroprogestin n=77)	Control Group (Placebo n=110)
Sesti 2007 Italy	Short form 36 general health survey: Improvement of scores in all domains at 12 months	Short form 36 general health survey: Improvement of scores in all domains at 12 months

Even though this outcome could not be assessed using GRADE, this would be rated as very low quality evidence because of outcome reporting bias and lack of detail provided.

11.3.8 Economic evidence

One paper was found contrasting the use of medical treatments following surgery. Additionally, 1 paper was found looking at the costs of medical treatments before surgery.

Sanghera (2016)

This paper refers to a de novo economic model intended to assess the use of levonorgestrel-releasing intrauterine system, depot-medroxyprogesterone acetate, combined oral contraceptive pill (cOCP) and 'no treatment' after conservative surgery to prevent recurrence of endometriosis.

The model was a Markov Chain design with a time horizon of 36 months and a cycle length of 1 month. A discount rate of 3.5% is used and no half-cycle correction is applied as the model is based on discrete transitions. Cost estimates were taken from a 'recent primary parallel study' and appear to be taken from standard sources such as the NHS Reference Costs and NICE evaluations, uprated to 2016 values. Costs for increase indirect health resource use were not estimated. Utility estimates were taken from clinical consensus using a modified visual analogue scale.

ISBN 978-1-4731-2661-9

LNG-IUS cost £650.94 and generated 1.88 quality adjusted life years (QALYs), Depot medroxyprogesterone acetate (DMPA) cost £622.56 and generated 1.92 QALYs, cOCP cost £599.93 and also generated 1.92 QALYs and no treatment cost £371.34 and generated 2.27 QALYs. This indicates that no treatment dominates, as it is both the most effective and cheapest option following surgery. The paper is significantly limited by having to rely on estimates for the utility values of treatment states, as the results are heavily influenced by estimates of these values. As the result is highly counter-intuitive and contradicts estimates made by members of the Committee, less weight is put on this finding in the economic model and subsequent Committee discussion.

The paper also conducts a literature review and finds no other papers conducting an economic evaluation of hormonal treatment for endometriosis following conservative surgery, consistent with our findings.

Ferracini & Nakada (2013)

This paper is a cost minimisation study contrasting 3 possible pre-surgical hormonal treatment strategies. These strategies were: dienogest then surgery, leuprorelin acetate then surgery and finally one drug, then the other drug if no effect, then surgery in any case.

The source of cost data was 'Brazilian official data' and the time horizon was 6 months. The unit of cost measure was the Brazilian Real, BRL. As the time horizon was below 1 year, no discount rate has been applied. The Brazilian system is partially privately funded, so the authors disaggregated these costs.

For the comparison of dienogest vs. leuprorelin, the private cost of dienogest was 1020.42 Brazilian Real (BRL) (~£250) and the public cost was 1461.22 BRL (~£350) while the private cost of leuprorelin was 2328.94 (~£580) BRL and the public cost 2377.52 BRL (~£585)

For the comparison of both drugs together, the private cost of dienogest first was 882.74 BRL (~£200) and public cost 942.18 BRL (~£220), whereas the private cost of leuprorelin first was 768.13 BRL (~£170) and the public cost 856.77 BRL (~£210).

This does not provide good evidence on whether it is cost-effective to offer hormonal treatment before surgery, but does indicate there is a cost saving to providing robust hormonal treatment if hormonal treatment is offered before surgery.

Summary of findings from economic model

The cost of providing hormonal treatment after surgery is assumed to be simply the cost of the surgery itself plus the cost of a course of hormonal treatment to follow. The literature is inconsistent around which drug should be provided and for how long – a range of 3 months to 24 months has been identified in the clinical review. An estimate of 12 months of additional treatment with danazol is used for the purpose of economic modelling, making the total cost £1,546.42 for the initial surgery and a maximum of £597.56 for the subsequent hormonal treatment (this could be less if the woman relapses before the end of the full course of treatment) – the maximum cost of this technique is therefore £2,143.98.

An important health economic issue is whether the addition of hormonal treatment delays the recurrence of endometriosis. For example, doubling average recurrence time would halve the number of operations required to treat a woman over the course of her lifetime, with clear cost implications. The Committee thought it biologically plausible that such an effect might occur, but conceded that the evidence was not strong enough to recommend one way or the other.

Table 127 demonstrates that considering surgical treatments alone, combination treatment with an MRI or laparoscopic diagnosis extendedly dominates surgery alone. This is

ISBN 978-1-4731-2661-9

unsurprising as the NMA finds that combination treatment is extremely effective at controlling symptoms of pain.

Table 127: Cost and effectiveness of all non-dominated treatment strategies containing a combination treatment for pain

Treatment	Cost	QALY	ICER	Pr. cost-effective vs. no treatment (£20k / QALY)	Pr. cost-effective vs. no treatment (£30k / QALY)
Empirical Diagnosis and No Treatment	£22,899.35	18.739	Base Case	N/A	N/A
CA-125 and Laparoscopic Treatment	£25,368.67	19.016	Extendedly Dominated	71.43%	71.43%
Pelvic MRI and Laparoscopic Treatment	£26,686.53	19.085	Extendedly Dominated	85.71%	85.71%
Transabdominal Ultrasound and Laparoscopy + Hormonal	£27,908.07	19.251	Extendedly Dominated	76.19%	80.95%
Pelvic MRI and Laparoscopy + Hormonal	£28,125.90	19.404	£7,864.31	80.95%	85.71%
Laparoscopy and Laparoscopy + Hormonal	£34,123.57	19.493	£67,337.99	90.48%	95.24%

Table 128 that the opposite effect is true if the primary concern of the woman is to preserve fertility. The NMA showed that hormonal treatment suppressed fertility, so the most effective method of accruing quality adjusted life year (QALYs) for a woman (which were highly conditional on a live birth) was to offer surgery alone, without subsequent hormones.

The final column (probability of live birth) demonstrates that the effect of surgery appears greater than the effect of subsequent hormonal treatment; by the end of their lives more women on a combination treatment plan had had a live birth than women on no treatment at all. However since treatment with no subsequent hormonal treatment is cheaper than hormonal treatment and hormonal treatment suppress fertility the overall effect is for laparoscopy + hormonal treatment to be dominated by laparoscopy alone.

Table 128: Cost and effectiveness of all treatment strategies containing a combination treatment for fertility

Treatment	Cost	QALY	ICER	Pr. cost-effective vs. no treatment (£20k / QALY)	Pr. cost-effective vs. no treatment (£30k / QALY)	Pr. Live Birth
Empirical Diagnosis and No Treatment	£16,028.47	19.184	Base Case	N/A	N/A	11.90%
CA-125 and Laparoscopy + Hormonal	£16,564.63	19.202	Dominated	61.90%	61.90%	15.48%

ISBN 978-1-4731-2661-9

Treatment	Cost	QALY	ICER	Pr. cost-effective vs. no treatment (£20k / QALY)	Pr. cost-effective vs. no treatment (£30k / QALY)	Pr. Live Birth
Transabdominal Ultrasound and Laparoscopy + Hormonal	£18,216.92	19.208	Dominated	63.69%	63.69%	15.48%
CA-125 and Laparoscopic Treatment	£14,605.81	19.227	-£33,216.05	64.29%	64.29%	15.48%
Empirical Diagnosis and Laparoscopy + Hormonal	£34,692.83	19.294	Dominated	66.07%	66.67%	15.48%
Laparoscopy and Laparoscopy + Hormonal	£29,321.88	19.320	Dominated	59.52%	59.52%	20.24%
Pelvic MRI and Laparoscopy + Hormonal	£22,248.26	19.350	Dominated	63.10%	63.10%	18.45%
Empirical Diagnosis and Laparoscopic Treatment	£27,712.68	19.355	Dominated	62.50%	62.50%	19.64%
Transabdominal Ultrasound and Laparoscopic Treatment	£17,058.07	19.407	£13,607.13	66.67%	67.26%	22.02%
Laparoscopy and Laparoscopic Treatment	£27,444.47	19.409	Dominated	63.69%	63.69%	22.02%
Pelvic MRI and Laparoscopic Treatment	£19,424.28	19.415	£300,633.64	58.33%	58.33%	23.21%

11.3.9 Clinical evidence statements

11.3.9.1 Pain

Pain recurrence

Low quality evidence from 1 trial (n= 53) reported that there is no clinically significant difference between intranasal nafarelin and placebo after surgery for pain recurrence (measured using Andersch and Milsom scale).

Very low quality evidence from 4 trials (n= 476) found that there is no clinically significant difference between hormonal treatment (triptorelin, goserelin, decapeptyl, letrozole and danazol) and no treatment after surgery for pain recurrence at 12 months.

ISBN 978-1-4731-2661-9

Very low quality evidence from 3 trials (n= 312) reported that there is no clinically significant difference between hormonal treatment (leuprolide, goserelin and cyclic combined oral contraceptives) and no treatment after surgery for pain recurrence at 13 to 24 months.

Very low quality evidence from 1 trial (n=54) reported that there is no clinically significant difference between triptorelin treatment and no treatment after surgery for pain recurrence at 5 years.

Pelvic pain

Moderate evidence from 1 trial (n=187) found a clinically significant beneficial effect of hormonal treatments (triptorelin, leuprorelin and oestroprogestin) compared with placebo for pelvic pain (measured using VAS) after surgery although there was low and very low quality evidence of no clinically significant difference between the 2 interventions for dysmenorrhoea and deep dyspareunia.

Dyspareunia

Low quality evidence from 1 trial (n=120) found a clinically significant beneficial effect of leuprorelin treatment compared with no treatment for dyspareunia (measured using a questionnaire) after surgery at 12 months although there was low and very quality evidence of no clinically significant difference between the 2 interventions for abdominal pain or dysmenorrhoea.

Dysmenorrhoea

Moderate quality evidence from 2 trials (n= 95) found a clinically significant beneficial effect of LGN-IUS treatment compared with no treatment after surgery for dysmenorrhoea at 12 months.

Recurrence of endometriosis

Very low quality evidence from 1 trial (n=285) reported that there is no clinically significant difference between leuprolide treatment and no treatment after surgery for recurrence of endometriosis at 5-6 months after starting treatment.

Very low quality evidence from 3 trials (n=310) reported that there is no clinically significant difference between hormonal treatment (triptorelin, letrozole, leuprolide and danazol) and no treatment after surgery for recurrence of endometriosis at 12 months

Very low quality evidence from 1 trial (n=45) reported that there is no clinically significant difference between hormonal treatment (danazol or an unspecified GnRH agonist) compared with no treatment after surgery for endometriosis recurrence at 24 months.

11.3.9.2 Recurrence of endometrioma

Low quality evidence from 3 trials (n= 463) reported a clinically significant beneficial effect of between hormonal treatment (triptorelin, leuprolide and combined oral contraceptives) and placebo or no treatment after surgery for endometrioma recurrence at 13-36 months.

Very low quality evidence from 1 trial (n=35) reported that there is no clinically significant difference between triptorelin treatment and no treatment after surgery for endometrioma recurrence at 5 years.

11.3.9.3 Health related quality of life

Very low quality evidence from 1 trial (n=187) reported that women receiving hormone treatment with GnRH agonist or oestroprogestin (oestradiol plus medroxyprogesterone) and

ISBN 978-1-4731-2661-9

women receiving placebo had improved quality of life (improved scores in all domains of the SF-36 general health survey) at 12 months.

Satisfaction

Low quality evidence from 2 trials (n=95) reported no clinically significant difference in patient satisfaction with treatment results when LGN-IUS treatment was compared with no treatment after surgery.

11.3.9.4 Reoperation rates

Very low quality evidence from 3 trials (n=327) reported that there is no clinically significant difference between hormonal treatment (triptorelin, leuprolide, danazol and oestroprogestin) and placebo or no treatment after surgery on reoperation rates.

11.3.10 Evidence to recommendations

11.3.10.1 Relative value placed on the outcomes considered

The Committee prioritised pain relief, health related quality of life and adverse events as critical outcomes for their decision making and number of women requiring more surgery, absence from work and other activities of daily living and fertility as important outcomes. However, when the outcomes for the NMA were discussed subsequently, it was decided that adverse events causing withdrawal from the study and fertility would be more appropriately considered as outcomes in the NMA.

11.3.10.2 Consideration of clinical benefits and harms

In view of the high rate of recurrence of endometriosis, affecting long-term quality of life for many women, improvement in long-term control of the condition was felt by the Committee to be clinically very important. The Committee were aware of the high rate of reoperation for endometriosis with associated risks of surgery and, as there was strong evidence to support this, considered that avoidance of repeat surgery by the use of long -term medical therapy would be beneficial. The Committee noted that the duration of follow-up in most studies was insufficient, but brought additional clinical experience to the discussion. Based on the evidence, the beneficial effect of all hormonal therapies was similar (probably because all work through similar mechanisms) and so the Committee considered the adverse effects of the various treatments in making their recommendation, as there are known side effects with hormonal treatments that some women may wish to avoid.

In general, the Committee considered that the combined oral contraceptive pill or long-acting reversible progestogen contraceptives were the most acceptable treatments. The Committee noted that these would not be appropriate for women who were trying to conceive, although they could be used by women who were planning pregnancy at some time in the future. They also felt it was important to note that GnRHa are only licensed for 6 months due to a loss of bone density.

11.3.10.3 Consideration of economic benefits and harms

The Committee discussed how the addition of hormonal treatments either before or after surgery was likely to carry a very low direct cost to the NHS and therefore could be recommended if the clinical evidence was thought strong enough to support such a recommendation. Many studies identified in the literature review used a more expensive hormonal treatment such as GnRHa in their pre / post-surgical dosage, and the economic evidence for this is more equivocal – although the model suggests that it would be cost-effective to offer such treatment at £20,000 / QALY, the Committee were told that cheaper

ISBN 978-1-4731-2661-9

hormonal treatments like the combined oral contraceptive pill were likely to be more cost-effective

The above holds true for fertility treatments too. If fertility outcomes are improved by adding a hormonal treatment then this could be considered as it is likely to be cost-effective, but the Committee thought in this instance that the clinical evidence did not support offering a hormonal treatment to women attempting to become pregnant.

The likely resource impact of these recommendations is somewhere between low and negative – most women who are able to tolerate hormonal treatments already receive these for endometriosis, so the Committee's recommendations were not a significant departure from current practice. Even if they were, the cost of long-acting reversible contraception is not substantial. The recommendations may cause a small cost saving, since Committee opinion is that some clinicians prescribe more expensive hormonal treatments such as GnRHas before trialling combined oral contraceptive pill or long-acting reversible progestogen contraceptives. These recommendations should prevent that unwarranted clinical variation, saving NHS resources.

11.3.10.4 Quality of evidence

Evidence was available from 12 studies in total and the quality ranged from moderate to very low. Studies that reported pain as dichotomously or results from scoring systems not included in the NMA were included in the pairwise reviews. Pain outcomes using the Biberoglu and Behrman scale (B&B) were also reported in the pairwise analysis where these were presented as separate components (dysmenorrhoea, dyspareunia and pelvic pain).

The Committee commented that the descriptions of the surgery performed were poor and that the included studies had been published over a 30 year period. Although the techniques used over this time had not changed greatly, there had been significant improvement in laparoscopic technology resulting in a surgeon's ability to remove more diseased tissue through improved visualisation. Thus it was difficult to draw overall conclusions from the included studies regarding the quality of the surgery performed. The Committee further noted that this might also affect assessment of the effectiveness of the additional hormonal suppression therapy as women might have a comparatively greater treatment effect where less diseased tissue had been removed by surgery.

The Committee noted that 3 trials had used excision techniques to remove endometrioma rather than ablative techniques and that excision had been demonstrated to be superior to ablation in a separate review.

Various hormone suppression therapies were examined in the included studies. The Committee questioned the relevance and accuracy of reporting of dysmenorrhoea and dyspareunia pain outcomes in trials using GnRH analogues. These therapies can suppress menstruation and decrease libido to such an extent that assessment of pain associated with menstruation or sexual intercourse might be irrelevant if neither were occurring and studies did not report any confirmation of questioning women as to whether either were.

Further, different types of GnRHa therapies have different routes of administration. For example, leuprorelin is administered as a depot injection which diminishes uncertainty regarding dose received and user compliance compared to intranasal administration of nafarelin where there can be variability in the dose retained and which needs to be administered every 12 hours or so,

The results of 1 trial conducted in 1994 were particularly unreliable as surgery had been performed using laparotomy combined with intranasal nafarelin.

ISBN 978-1-4731-2661-9

11.3.10.5 Other considerations

The Committee gave special consideration to young women (aged 17 and under) and discussed whether any additional recommendations were necessary but concluded that none were required.

Based on consensus the Committee agreed that hormonal treatment prior to surgery would only be suitable for women with deep endometriosis involving the bowel, bladder or ureter. The Committee noted that this would usually lead to less bleeding and would therefore aid the surgical procedure.

11.3.10.6 Key conclusions

The Committee based their recommendations on the findings of the NMA, which demonstrated that adding hormonal treatment following surgery (laparoscopic excision or ablation) reduces the risk of recurrence and symptoms, so it should be offered to women post-surgery unless they want to conceive.

11.3.11 Recommendations

46. After laparoscopic excision or ablation of endometriosis, consider hormonal treatment (with, for example, the combined oral contraceptive pill)[c], to prolong the benefits of surgery and manage symptoms.

11.4 Hysterectomy (with or without oophorectomy) in combination with surgical management

Review question: What is the effectiveness of hysterectomy with or without oophorectomy, including recurrent and asymptomatic endometriosis, in managing endometriosis?

11.4.1 Introduction

Hysterectomy combined with surgical excision/ablation of endometriosis is currently offered for the treatment of endometriosis when medical and hysterectomy sparing surgical options have been offered, failed or are inappropriate. Hysterectomy is associated with potential morbidity and a very low risk of mortality. Due to the fact that endometriosis is thought to be a predominantly oestrogen-dependent disease, women can opt to have their ovaries removed at the time of hysterectomy, often depending on the severity and location of their endometriosis. However, it is unclear whether a hysterectomy without oophorectomy may be as clinically effective as with oophorectomy and there is currently variation in clinical practice.

In either case it is critical that women are appropriately counselled about the fact that they will no longer be able to have children after a hysterectomy, the risks of early oophorectomy (e.g. osteoporosis, cardiovascular disease), the effects of a surgical menopause, the need for hormone replacement until the age of natural menopause and the potential for recurrence of the disease. There are also different routes by which this could be carried out, i.e. laparoscopic or abdominal. However, individual assessment and the experience of the clinician are very important because patient characteristics and surgical expertise are determinants of the chosen approach. Hysterectomy is not currently offered for the treatment

c At the time of publication (September 2017), not all hormonal treatments (including not all combined oral contraceptive pills) have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Prescribing guidance: prescribing unlicensed medicines for further information.

ISBN 978-1-4731-2661-9

of asymptomatic endometriosis. The effectiveness of hysterectomy with and without oophorectomy is discussed below.

11.4.2 Description of clinical evidence

The objective of this review is to determine the clinical and cost-effectiveness of hysterectomy with or without oophorectomy in reducing pain, improving health-related quality of life and reducing adverse events.

For full details see the review protocol in Appendix D.

Two observational studies were included in this review (Shakiba 2008, Namnoum 1995). No other evidence was identified.

Both of the included studies were retrospective cohort studies that were carried out in the USA.

In both studies, a retrospective review of medical records was completed. In Shakiba 2008, records were searched for women who had surgery for chronic pelvic pain with histological confirmation of endometriosis, of any stage and severity between January 1995 and December 2003. Women who had surgery for infertility or menorrhagia as the primary indication were excluded from the study. Follow-up information was obtained in 2006 from medical records (operative reports, pathology reports, outpatient charts and a telephone survey consisting of a questionnaire about reoperation, pain clinic visit, medical treatment and level of satisfaction). Surgery was only performed if other therapies failed to control symptoms.

In Namnoum 1995 the inclusion criteria were women who underwent a hysterectomy with a diagnosis of endometriosis (unclear diagnostic method) between 1979 and 1991. The study excluded women who were older than 45 years at the time of hysterectomy in order to prevent confounding the results by including data from women with menopausal changes. Follow-up data were obtained primarily from outpatient charts and telephone questionnaires. However, written questionnaires were sent if the patient could not be reached by telephone

Shakiba 2008 evaluated the need for further surgery after laparoscopic excision of endometriosis or hysterectomy. Even though the focus of the study does not match our current protocol, women who had hysterectomy (n=97) were divided into 2 subgroups depending on whether they had bilateral oophorectomy. For this review we selected data for hysterectomy subgroups (hysterectomy with or without oophorectomy; n=47 and n=50 respectively). The only outcome reported was the effect of ovarian preservation on reoperation-free survival for each surgery group. In this review, the data for the outcome for 7 years follow-up in the 2 hysterectomy subgroups (hysterectomy with or without oophorectomy) was presented.

Namnoum 1995 compared the rates of reoperation and symptom recurrence (pain) between groups with some ovarian preservation (n=29) compared with women who had all ovarian tissue removed (n=109). The mean duration of follow-up was 58 months (4 years 10 months) post hysterectomy.

We did not identify any evidence for the following outcomes:

- Quality of life
- Effect on daily activities
- Unintended effects from treatment
- Participant satisfaction with treatment

Evidence is summarised in the clinical GRADE evidence profile below (Table 130). See also the study selection flow chart in Appendix F, the study selection flow chart in Appendix F,

ISBN 978-1-4731-2661-9

study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G. Summary of included studies

A summary of the studies that were included in this review are presented in Table 129.

Table 129: Summary of included studies

Study	Intervention/Comparison	Population	Outcomes	Comments
Shakiba 2008 USA	• Hysterectomy with or • without oophorectomy	Women diagnosed with endometriosis who had undergone surgery for chronic pelvic pain with histological confirmation of endometriosis N=97	Requirement of reoperation (Effect of ovarian preservation on reoperation free survival)	Endometriosis was staged according to the revised American Fertility Society Mean follow-up of 7 years.
Namnoum 1995 USA	• Hysterectomy with or • without oophorectomy	Women undergoing hysterectomy with a diagnosis of endometriosis. Women undergoing hysterectomy aged 45 years or older were excluded. N=138	Requirement of reoperation. Recurrence of symptoms (pain)	Endometriosis was staged according to the revised American Fertility Society. Mean follow-up of 4 years 10 months.

11.4.3 Clinical evidence profile

The clinical evidence profile for this review question (hysterectomy with or without oophorectomy for the treatment of endometriosis) is presented in Table 130.

In Namnoum 1995, a Cox proportional hazards model was used to investigate the relative risk of pain recurrence and relative risk of reoperation when adjusted for age at time of hysterectomy (≤35 years vs. >35 years), stage of disease (revised AFS criteria), previous medical therapy and previous surgical therapy. The results for the risk of pain recurrence showed that the relative risk for pain 6.1 (95% confidence interval (CI) 2.5% to 14.6%) with ovarian conservation compared with bilateral oophorectomy. The results for reoperation showed that the relative risk of reoperation was 8.1 (95% CI 2.1% to 31.2%) with ovarian conservation compared with bilateral oophorectomy.

In Shakiba 2008, a Cox proportional hazards ratio investigating time to reoperation when adjusted for age and stage of disease was reported for hysterectomy plus bilateral oophorectomy compared with hysterectomy only and showed that preservation of both ovaries increased the risk of reoperation by 2.44 times compared with both ovaries removed, but there was a lot of uncertainty around this result (P=0.18) with a wide 95% CI (0.65% to 9.10%), due to the small sample size. The authors reported that confounding factors such as stage of disease did not have any effect on surgery free time in either group, but age at the time of surgery was important in determining the outcome.

A Kaplan-Meier graph showed reoperation free survival estimates at 2, 5 and 7 years in the hysterectomy subgroups. In the hysterectomy only group, the 2, 5 and 7 year percentages of women who avoided reoperation were 95.7%, 86.6% and 77.0% respectively. In the

ISBN 978-1-4731-2661-9

hysterectomy with bilateral oophorectomy group, the 2, 5 and 7 year percentages of women who avoided were 96.0%, 91.7% and 91.7% respectively. It would suggest that women who had oophorectomy at the time of hysterectomy had a lower reoperation rate compared with women who had hysterectomy alone.

Table 130: Summary clinical evidence profile

Outcomes	Hazard ratio (95% CI)	Absolute effect	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
Reoperation-free survival (effect of ovarian preservation (Hysterectomy only versus hysterectomy plus bilateral oophorectomy)	HR 2.44 (0.65 to 9.10)	An absolute effect could not be calculated3	N=97 (1 study)	⊕⊖⊖⊖ Very low1,2,3	
Relative risk for reoperation (effect of ovarian preservation)	RR 8.1 (2.1 to 31.3)	An absolute effect could not be calculated5	N=138 (1 study)	⊕⊖⊖⊖ Very low4,5,6	Mean follow up 4 years 10 months
Relative risk for symptom recurrence (pain)	RR 6.1 (95% CI 2.5 to 14.6)	An absolute effect could not be calculated5	N=138 (1 study)	⊕⊖⊖⊖ Very low4,5,6,	Mean follow up 4 years 10 months

CI: confidence interval; HR: Hazard ratio; RR: risk ratio
1 Evidence was downgraded by 1 due to outcome selection bias
2 Evidence was downgraded by 2 due to very serious imprecision as 95% confidence interval crossed 2 default minimally important differences (MIDs).
3 Adjusted for age, stage of disease, or operative time predictive for reoperation. Age and time of surgery were considered important confounding factors, stage of disease did not have any effect on surgery-free time in any group, but stratification for multiple factors reduced the statistical power and even large differences may not reach statistical significance even though the size of the difference may be clinically important. The P value for the comparison was 0.18.
4 Evidence was downgraded by 2 due to risk of bias; study design was a retrospective cohort with outpatient chart review.
5 Cox proportional hazards model adjusting for adjusting for revised AFS classification of endometriosis stage, previous medical therapy, previous surgical therapy and age at time of hysterectomy (≤35 years vs. > 35 years)
6 Evidence was downgraded by 1 due to indirectness: The hysterectomies in the study took place between 1979 to 1991, which may limit the applicability of the study with regards to current surgical techniques and outcomes. In addition, women over 45 years were excluded.

11.4.4 Clinical evidence statements

Very low quality evidence from 1 retrospective cohort study with 97 participants showed that there was no clinically significant difference between the 2 interventions for reoperation free survival up to 7 years.

Very low quality evidence from 1 retrospective cohort study with 136 participants that after a mean follow-up of 4 years 10 months, there was a lower rate of reoperation after hysterectomy with oophorectomy compared to hysterectomy with ovarian conservation.

Very low quality evidence from 1 retrospective cohort study with 136 participant that after a mean follow-up of 4 years 10 months, there was a lower rate of pain recurrence after hysterectomy with oophorectomy compared to hysterectomy with ovarian conservation.

ISBN 978-1-4731-2661-9

11.4.5 Economic evidence

No health economic evidence was found on the cost-effectiveness of hysterectomy for endometriosis.

The costs of hysterectomy to the NHS are largely driven by the cost of the operation itself. Additionally, there may be complications or long-term effects of the operation which should be taken into account. Table 131 presents various costs for the initial operation given in the NHS Reference Costs for 2013/14. There is no specific code for a hysterectomy (either with or without oophorectomy), so the table presents a variety of plausible codes.

Table 131: Summary of Hysterectomy Costs

Currency Code	Procedure Name	National Average Cost
LB71Z	Total Pelvic Exenteration	£16,361
MA02C	Very Major Open, Upper or Lower Genital Tract Procedures, with CC Score 0-1	£4,013
MA07G	Major Open Upper Genital Tract Procedures with CC Score 0-2	£3,586
MA28Z	Complex, Laparoscopic or Endoscopic, Upper Genital Tract Procedures	£3,636

(b) CC: Complications and comorbidities

Excluding pelvic exenteration (which is included as an upper bound figure only), it seems the cost of a hysterectomy to the NHS is somewhere between £3500 and £4000.

Hysterectomy is likely to be a highly cost-effective treatment for endometriosis if it is clinically effective, especially if given to young women as it requires a one-off payment. However the economic harm of such a strategy is that the woman will be infertile for the rest of her life. Another harm is that this treatment will induce a surgical menopause; if menopause has Quality of life (QoL) implications (for example, affecting a woman's mental health more than if the menopause happens naturally over a number of years) so too will this strategy.

Table 132: Cost and effectiveness of all non-dominated treatment strategies containing a hysterectomy

Treatment	Cost	QALY	ICER	Pr. cost-effective vs. no treatment (£20k / QALY)	Pr. cost-effective vs. no treatment (£30k / QALY)
Empirical Diagnosis & No Treatment	£23,150.21	18.424	Base Case	62.22%	62.22%
CA-125 & Hysterectomy	£24,318.30	19.742	Extendedly Dominated	80.00%	80.00%
Pelvic MRI & Hysterectomy	£24,407.45	20.616	£573.50	91.11%	91.11%
Laparoscopy & Hysterectomy	£28,913.62	20.702	£52,403.82	88.89%	91.11%
Empirical Diagnosis & Hysterectomy	£38,856.92	20.777	£132,467.23	93.33%	95.56%

(c) ICER: incremental cost-effectiveness ratio; MRI: Magnetic resonance imaging; Pr.: Probability; QALY: quality
(d) adjusted life years

Table 132 indicates that a hysterectomy is cost-effective at a very low threshold of £574 / quality adjusted life year (QALY) and every technique is highly likely to be cost-effective vs. no treatment for a given patient. In comparison to the most cost-effective treatment in the

ISBN 978-1-4731-2661-9

model (empirical diagnosis and oral hormonal contraceptive pill), hysterectomy is cost-effective at a threshold of £4239 / QALY. Note that this model does not estimate the harm of giving a hysterectomy to a woman who does not have endometriosis (and so therefore is not cured by giving up her fertility) but this would likely be large.

11.4.6 Evidence to recommendations

11.4.6.1 Relative value placed on the outcomes considered

The Committee considered the following outcomes to be important for their decision-making:

- Pain relief
- Quality of life
- Unintended effects from treatment

Evidence was only available for pain relief in 2 old, small retrospective cohort studies. Evidence for the other critical outcomes was not identified.

11.4.6.2 Consideration of clinical benefits and harms

As only very low quality evidence was identified, the Committee based their recommendations on their experience and expertise.

The Committee was of the opinion that endometriosis by definition is endometriotic tissue outside the uterus, which means that it is not expected to be cured by hysterectomy. However, they agreed to highlight some indications for hysterectomy (for instance in presence of adenomyosis or heavy menstrual bleeding not responding to other treatments) and that it would then be important that the endometriotic lesions would be removed at the same time. Since this excision or ablation of the endometriotic tissue would be carried out laparoscopically the Committee agreed that the hysterectomy should also be carried out laparoscopically to enable both to happen in one procedure. The Committee noted that laparoscopic surgery may be contraindicated for a few women for example, those who cannot undergo procedures under anaesthetic, where there are large fibroids or where there are severe adhesions perhaps following major bowel resection, but that generally decisions regarding surgery would be based on relative harms and benefits. Also, the Committee noted that endometriosis is a hormone-dependent condition and it is therefore plausible that oophorectomy would be more effective than hysterectomy alone.

They concluded that making an informed choice was very important to women and that therefore all necessary details should be discussed. The Committee identified that women would need to know that hysterectomy was not a treatment for endometriosis but that excision of the endometriotic lesions at the same time is the actual treatment. Having the hysterectomy could then prevent or delay recurrence of endometriosis. Healthcare professionals should inform women about the procedure how it would affect their symptoms and the implications of oophorectomy (such as possible risks related to osteoporosis and cardiovascular conditions). The Committee recognised that there can be significant social / psychological effects of hysterectomy. The Committee considered cross- referencing to the NICE guideline on heavy menstrual bleeding, but because the population within that guideline would be different (they would not necessarily be suspected of having endometriosis), elected not to do so. The Committee noted that bilateral oophorectomy induces surgical menopause and therefore cross referenced the menopause guideline because the symptoms of menopause may be severe and the benefits and risks of hormone replacement therapy should be discussed.

ISBN 978-1-4731-2661-9

11.4.6.3 Consideration of economic benefits and harms

The Committee understood that health economic costs did not indicate hysterectomy should be a first-line treatment, but that hysterectomy was a cost-effective option to consider in women who would prefer this option over and above their other options which would preserve fertility.

Hysterectomy for women with endometriosis is cost-saving for the NHS and so these recommendations are likely to have a small negative resource impact

11.4.6.4 Quality of evidence

Evidence from 2 retrospective cohort studies was identified for inclusion and was of very low quality due to risk of bias in both studies (study design, outcome selection and detection bias), imprecision of results in 1 study (width of the confidence interval) and indirectness in 1 study (age of study limiting applicability for modern surgical techniques). Therefore there is uncertainty around the evidence that these studies provides.

The data in both studies were limited due to their retrospective cohort design. In 1 study the main comparison was between laparoscopic excision and hysterectomy. The results for the current review rely on a subgroup analysis only and were therefore underpowered. The second study was old (hysterectomies were conducted between 1979 and 1991) and it is unclear whether the outcomes would have changed based on modern techniques.

It is difficult to say whether the results are generalizable to all women who would have such surgery, because women who were included in the studies were from tertiary care referral centres. In both studies more than 50% of the women had advanced disease and more than 50% had at least 1 previous surgery, with the rates being as high as 77% in 1 study.

Although the result from these studies may show clinical benefit, it should be applied with caution as there are limitations in study design and the ability to be applied to the current population. In addition, the Shakiba 2008 result is not precise and also both studies had a small sample size and no other, better quality evidence has been identified.

11.4.6.5 Other considerations

The Committee noted that that the laparoscopic approach to hysterectomy is possibly safer and is a better use of resources than laparotomy which is no longer widely used.

11.4.6.6 Key conclusions

The Committee concluded that the 2 included studies provided too little and very low quality evidence to draw clear conclusions about the comparative effects between hysterectomy only and hysterectomy plus oophorectomy. The Committee therefore based the recommendations on expertise, experience and consensus highlighting examples of some possible indications, noting that the endometriotic tissue should be removed at the same time, and stating how this procedure should be carried out. The Committee also agreed that it was important to provide information to women who consider a hysterectomy and specified what a discussion about this should include.

11.4.7 Recommendations

47. If hysterectomy is indicated (for example, if the woman has adenomyosis or heavy menstrual bleeding that has not responded to other treatments), excise all visible endometriotic lesions at the time of the hysterectomy.

ISBN 978-1-4731-2661-9

48. **Perform hysterectomy (with or without oophorectomy) laparoscopically when combined with surgical treatment of endometriosis, unless there are contraindications.**

49. **For women thinking about having a hysterectomy, discuss:**
    - what a hysterectomy involves and when it may be needed
    - the possible benefits and risks of hysterectomy
    - the possible benefits and risks of having oophorectomy at the same time
    - how a hysterectomy (with or without oophorectomy) could affect endometriosis symptoms
    - that hysterectomy should be combined with excision of all visible endometriotic lesions
    - endometriosis recurrence and the possible need for further surgery
    - the possible benefits and risks of hormone replacement therapy after hysterectomy with oophorectomy (also see the NICE guideline on menopause).

ISBN 978-1-4731-2661-9

12 Pharmacological, non-pharmacological, surgical and combination management strategies - if fertility is a priority

Review question: What is the effectiveness of the following ovulation suppression treatments or surgery (or combinations of these) or non-pharmacological treatments for improving spontaneous pregnancy rates in endometriosis, including recurrent and asymptomatic endometriosis:

- hormonal medical treatments
- surgery
- non-pharmacological therapies
- combinations of surgery plus hormonal treatment?

12.1 Introduction

Endometriosis is recognised as an important cause of infertility, with a prevalence of 25–40% in infertile women, compared with 0.5–5% in fertile women (Ozkan 2008). Management of endometriosis, as well as fertility interventions, aim to improve a woman's chances of pregnancy. Since publication of the NICE guideline on fertility (CG156), which included recommendations related to the treatment of women with endometriosis wanting to conceive, further evidence has been published on first-line treatments for subfertility. These are related to surgical treatments, including surgical ablation and excision. Due to this new evidence there is therefore uncertainty about the comparative effectiveness of these interventions. Possible options for management of endometriosis include laparoscopic surgery and may also include short-term hormonal treatment pre- or post-surgery. The updated recommendations form part of this guideline.

The scope of the current guideline excluded investigation of fertility problems related to endometriosis. It also excludes management of subfertility using assisted reproductive techniques. This chapter therefore is reviewing the impact of the management strategies covered in the scope on spontaneous pregnancy rates of women with endometriosis who are trying to conceive. Many sections in the NICE's guideline on fertility problems (CG 156) are also relevant in this context of endometriosis (such as diagnostic fertility tests, ovarian reserve testing and preoperative tests) but are outside the scope of this guideline.

12.1.1 Methods for the network meta-analysis

12.1.2 Study selection and data collection

For full details see analysis protocol in Appendix D, the study selection flow chart in Appendix F, study exclusion list in Appendix H, forest plots in Appendix I, full GRADE profiles in Appendix J and study evidence tables in Appendix G.

12.1.3 Outcome measures

Spontaneous (i.e. non-assisted) pregnancy

Although the Committee highlighted that live birth was the most important outcome for sub-fertile women with endometriosis, it was agreed that evidence for this would be limited and therefore the network meta-analysis (NMA) should be of studies reporting spontaneous pregnancy, as many more studies reported this outcome.

ISBN 978-1-4731-2661-9

Relative treatment effect estimates were not found to vary over the follow-up times considered for the review and therefore treatment effects were modelled as odds ratios (ORs) (Appendix I).

12.1.4 Statistical methodology

Data were available for a number of treatments and routes of administration. Due to the sparseness of the networks, it was necessary to group treatments within different classes and assume a common class effect (Table 133). The common class effects were assessed to identify if it was reasonable to assume similarity of treatment effects within classes. Multi-level NMA models with treatments nested within classes were also examined, though this added complexity did not improve model fit for any of the analyses.

Table 133: Dose ranges of treatments in different classes of interventions, with abbreviations used in tables and figures within this chapter

Class	Treatment	Abbreviation
Placebo/diagnostic laparoscopy	Placebo Diagnostic laparoscopy	Plac/diag
Danazol/gestrinone	Danazol (100–800 mg/d) Gestrinone	Dan/gest
Oestrogens (oral)	Oestradiol (1–2 mg/d) Conjugated equine oestrogens (0.3–1.25 mg/d)	Oest(o)
Progestogens (oral)	Norethisterone (2.5 mg/d) Medroxyprogesterone (15–30 mg/d) Levonorgestrel (30 micrograms/d) Desogestrel (75 micrograms/d) Dienogest (2 mg/d)	Prog(o)
Progestogens (depot)	Medroxyprogesterone (150 mg/3m) Gestodene (5–10 mg)	Prog(i.m.)
Progestogens (subcutaneous)	Medroxyprogesterone (104 mg/3m) Promegestone	Prog(s.c.)
Progestogens (intrauterine)	Levonorgestrel (20 micrograms/d)	Prog(i.u.)
GnRH agonists (depot)	Leuprorelide (3.75 mg/m) Triptorelin (3 mg/m)	GnRHa(i.m.)
GnRH agonists (subcutaneous)	Goserelin (3.6 mg/m)	GnRHa(s.c.)
GnRH agonists (nasal spray)	Nafarelin (200 micrograms b.d.) Buserelin (300 micrograms t.d.)	GnRHa(i.n.)
GnRH antagonists	Elagolix	GnRHant
Aromatase inhibitors	Anastrozole (1 mg/d) Letrozole (2.5 mg/d)	AromaInhib
Anti-androgens	Cyproterone acetate (only in combination as combined oral contraceptive)	Anti-And
Selective oestrogen receptor modulators	Raloxifene (60 mg/d)	SERM
Tibolone	Tibolone (2.5 mg/d)	-
Laparoscopy	Ablation (laser, diathermy, etc.) Excision (laser, diathermy, etc.)	LaparoSurg
Nutritional supplements	Calcium Vitamin D	Supp

ISBN 978-1-4731-2661-9

Class	Treatment	Abbreviation
Chinese herbal medicine	Nei yi pills Dan'e mixture	CHM
Dietary interventions	Dietary intervention	Diet

Table only includes treatments in full-text studies assessed for inclusion/exclusion. Treatments only in studies that were not included in the NMA could not be included in the network.

12.2 Summary of included studies

12.2.1 Studies included in the NMA

All studies included women with laparoscopic confirmation of endometriosis who had been trying unsuccessfully to conceive for at least 12 months. All hormonal treatments were used to suppress ovulation for at least 12 weeks. Women then attempted to conceive after hormonal treatment had ceased.

Table 134: Characteristics of included studies

First Author	Pub Date	rAFS	Surgery type	Endometriomas included	Risk of bias
Alborzi	2011	III–IV	Excision	No endometriomas	High
Bayer	1988	NR	No surgery	No endometriomas	High
Bianchi	1999	III–IV	Not reported	NR	High
Burry	1989	I–IV	No surgery	No endometriomas	Moderate
Fedele	1992	I–II	No surgery	No endometriomas	High
Fedele	1989	I–IV	No surgery	No endometriomas	Moderate
Fraser	1991	I–II	No surgery	NR	Low
Gad	2012	I–II	Excision/ablation	No endometriomas	High
Loverro	2008	III–IV	Excision/ablation	Some endometriomas	Low
Marcoux	1997	I–II	Excision/ablation	NR	High
Moini	2012	I–II	Ablation	NR	Moderate
Overton	1994	III–IV	No surgery	No endometriomas	Moderate
Seibel	1982	NR	No surgery	No endometriomas	High
Thomas	1987	I–II	No surgery	NR	Moderate
Wu	2006	NR	No surgery	No endometriomas	Moderate
Zhu	2014	I–II	Excision/ablation	No endometriomas	Moderate

(e) Pub Date: date of publication; rAFS: revised American Fertility Scale; NR: not reported in study

12.2.1.1 Studies excluded from the NMA

Table 135: Table of studies excluded from the NMA for statistical reasons

First author	Publication date	Reason for exclusion
Beretta	1998	Within-class comparison
Busacca	2001	Study adds no information to network

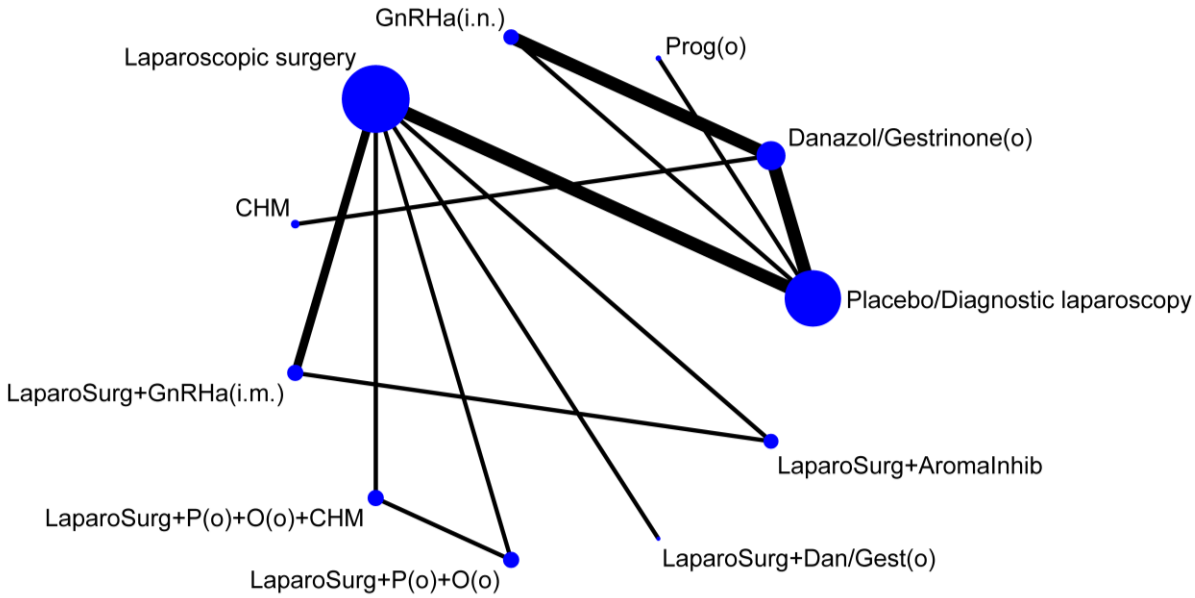
ISBN 978-1-4731-2661-9

12.3 Clinical evidence profile

12.3.1 Spontaneous pregnancy

Sixteen trials of 11 treatment classes were included in the network, with a total sample size of 1,404 women (Figure 22). Seven studies were at high risk of bias, 7 were at moderate risk of bias and 2 studies were at low risk of bias.

Figure 22: Network for spontaneous pregnancy



The size of nodes is proportional to the number of women in the network who were given a particular treatment class. The thickness of connecting lines is proportional to the number of studies directly comparing 2 treatment classes. For treatment name abbreviations, see Table 133.

Table 136 presents the results of the conventional pair-wise meta-analyses (direct comparisons; upper right section of table) together with the results from the NMA for every possible class comparison (lower left section of table), presented as odds ratios (ORs). These results were derived from a fixed effects model.

Laparoscopic surgery alone was found to lead to significantly more spontaneous pregnancies than diagnostic laparoscopy, while danazol/gestrinone led to fewer spontaneous pregnancies than placebo. For all other treatments there was considerable uncertainty regarding their effect on spontaneous pregnancy. Figure 23 graphically presents the results computed by the NMA for each treatment versus placebo/diagnostic laparoscopy.

The treatment with the highest probability of being 1 of the best 3 for improving spontaneous pregnancy was laparoscopic surgery (72.6%). Surgery plus danazol/gestrinone and surgery plus GnRHa (i.m.) also had a high probability of being among the best 3 treatments (65.3% and 47.2%, respectively), though this is likely to be due to the wide 95% credible intervals (CrI) rather than due to any evidence of a beneficial treatment effect for fertility (Table 137).

There was no evidence of incoherence in any other closed loops of treatments.

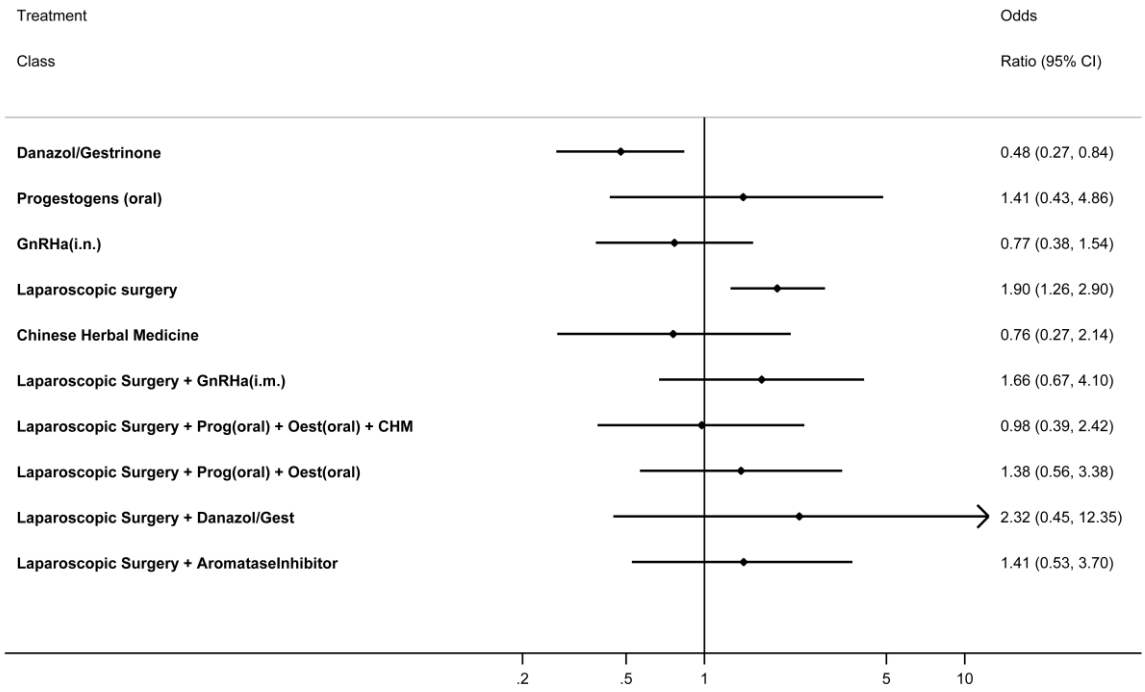
ISBN 978-1-4731-2661-9

Table 136: Matrix of results for the NMA of spontaneous pregnancy

Placebo/ diag	**0.52 (0.28 to 0.98)**	**1.41 (0.43 to 4.8)**	**0.6 (0.21 to 1.69)**	**1.91 (1.26 to 2.91)**						
0.48 (0.27 to 0.84)	Danazol/ gestrinone		1.8 (0.89 to 3.69)		1.59 (0.68 to 3.81)					
1.41 (0.43 to 4.86)	2.96 (0.8 to 11.48)	Prog (oral)								
0.77 (0.38 to 1.54)	1.61 (0.88 to 2.97)	0.54 (0.13 to 2.14)	GnRHa (i.n.)							
1.9 (1.26 to 2.9)	4 (1.99 to 8.09)	1.35 (0.37 to 4.72)	2.49 (1.11 to 5.62)	LaparoSurg		0.87 (0.39 to 1.93)	0.51 (0.22 to 1.15)	0.73 (0.33 to 1.59)	1.21 (0.25 to 6.17)	0.74 (0.3 to 1.77)
0.76 (0.27 to 2.15)	1.59 (0.68 to 3.8)	0.54 (0.11 to 2.58)	0.99 (0.35 to 2.85)	0.4 (0.13 to 1.21)	CHM					
1.66 (0.67 to 4.1)	3.49 (1.19 to 10.1)	1.17 (0.26 to 5.22)	2.16 (0.69 to 6.78)	0.87 (0.39 to 1.94)	2.19 (0.55 to 8.58)	LaparoSurg+ GnRHa(i.m.)				
0.98 (0.39 to 2.42)	2.05 (0.7 to 5.96)	0.69 (0.15 to 3.07)	1.27 (0.4 to 3.98)	0.51 (0.23 to 1.14)	1.28 (0.32 to 5.05)	0.59 (0.19 to 1.83)	LaparoSurg+ P(oral)+ O(oral)+ CHM			
1.38 (0.56 to 3.38)	2.9 (1.01 to 8.33)	0.98 (0.21 to 4.33)	1.8 (0.58 to 5.59)	0.73 (0.33 to 1.6)	1.82 (0.47 to 7.08)	0.83 (0.27 to 2.57)	1.42 (0.63 to 3.25)	LaparoSurg+ P(oral)+ O(oral)		
2.32 (0.45 to 2.35)	4.86 (0.85 to 8.28)	1.64 (0.21 to 2.74)	3.02 (0.5 to 18.47)	1.21 (0.25 to 6.11)	3.05 (0.43 to 1.68)	1.39 (0.23 to 8.54)	2.37 (0.4 to 14.46)	1.67 (0.28 to 0.16)	LaparoSurg+ Dan/gest	
1.42 (0.53 to 3.7)	2.97 (0.95 to 9.06)	1 (0.21 to 4.63)	1.84 (0.55 to 6.05)	0.74 (0.3 to 1.76)	1.86 (0.45 to 7.62)	0.85 (0.32 to 2.2)	1.45 (0.43 to 4.79)	1.02 (0.31 to 3.31)	0.61 (0.1 to 3.77)	LaparoSurg+ Aromanhib

ISBN 978-1-4731-2661-9

Figure 23: Forest plot showing odds ratios (95% CrI) of NMA estimates for each treatment versus placebo/diagnostic laparoscopy



For treatment name abbreviations, see Table 133

Table 137: Probabilities of being among the best 3 treatments and the worst 3 treatments, and the rank and 95% CrI for each treatment

Treatment class	Probability of being within the best 3 (%)	Probability of being within the worst 3 (%)	Rank (95% CrI)
Placebo/diagnostic laparoscopy	1.05%	12.12%	7 (4 to 9)
Danazol/gestrinone	0.00%	97.79%	11 (9 to 11)
Progestogens (oral)	37.59%	16.23%	5 (1 to 11)
GnRHa (i.n.)	2.10%	55.70%	9 (4 to 11)
Laparoscopic surgery	72.55%	0.00%	3 (1 to 5)
Chinese herbal medicine	6.71%	55.00%	9 (2 to 11)
Laparoscopic surgery + GnRHa(i.m.)	47.16%	4.29%	4 (1 to 9)
Laparoscopic surgery + Prog(oral) + Oest(oral) + CHM	6.95%	31.05%	7 (2 to 11)
Laparoscopic surgery + Prog(oral) + Oest(oral)	28.87%	8.14%	5 (1 to 10)

ISBN 978-1-4731-2661-9

Treatment class	Probability of being within the best 3 (%)	Probability of being within the worst 3 (%)	Rank (95% CrI)
Laparoscopic surgery + Danazol/Gestrinone	65.34%	8.97%	2 (1 to 11)
Laparoscopic surgery+ Aromatase inhibitor	31.69%	10.71%	5 (1 to 10)

(f) For treatment name abbreviations, see Table 133

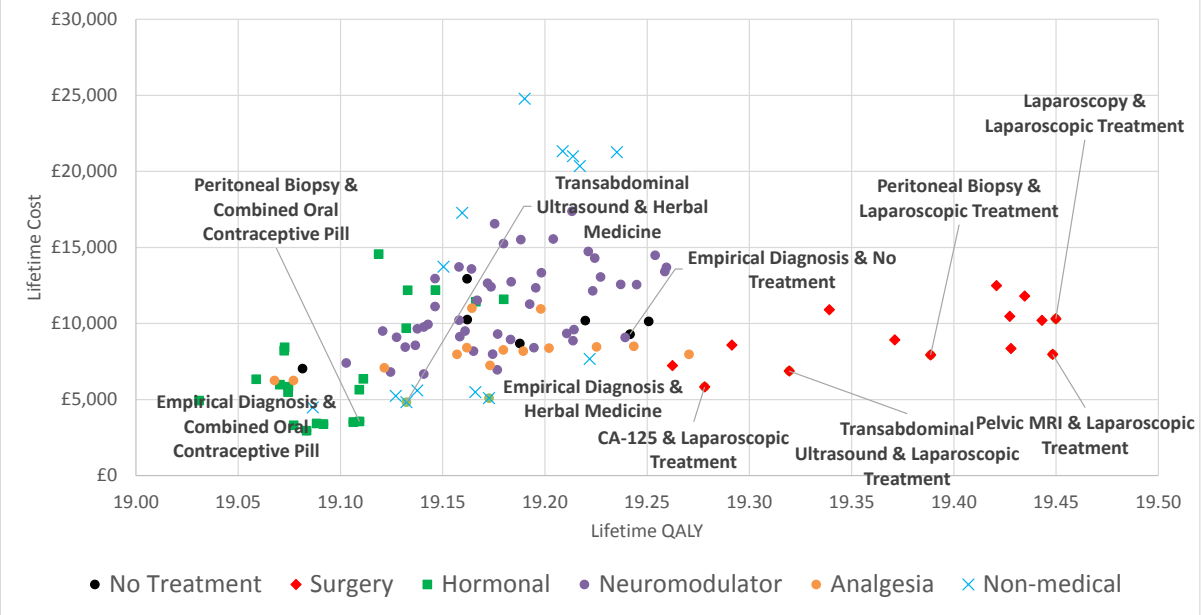
12.3.2 Economic evidence

No health economic evidence was found on the cost effectiveness of surgical or hormonal treatment to improve fertility for women with endometriosis.

Summary of relevant sections of the economic model

The results of the base case analysis are presented in Figure 24 and Figure 25. Progestogen treatments have been excluded from all analysis owing to Committee concern that the NMA shows a mean effect of progestogen treatments improving fertility when the Committee argued that this could only be an error with 1 or more of the studies as progestogen treatment is a contraceptive. For more details on the model used to generate the health economic results, please see Appendix K.

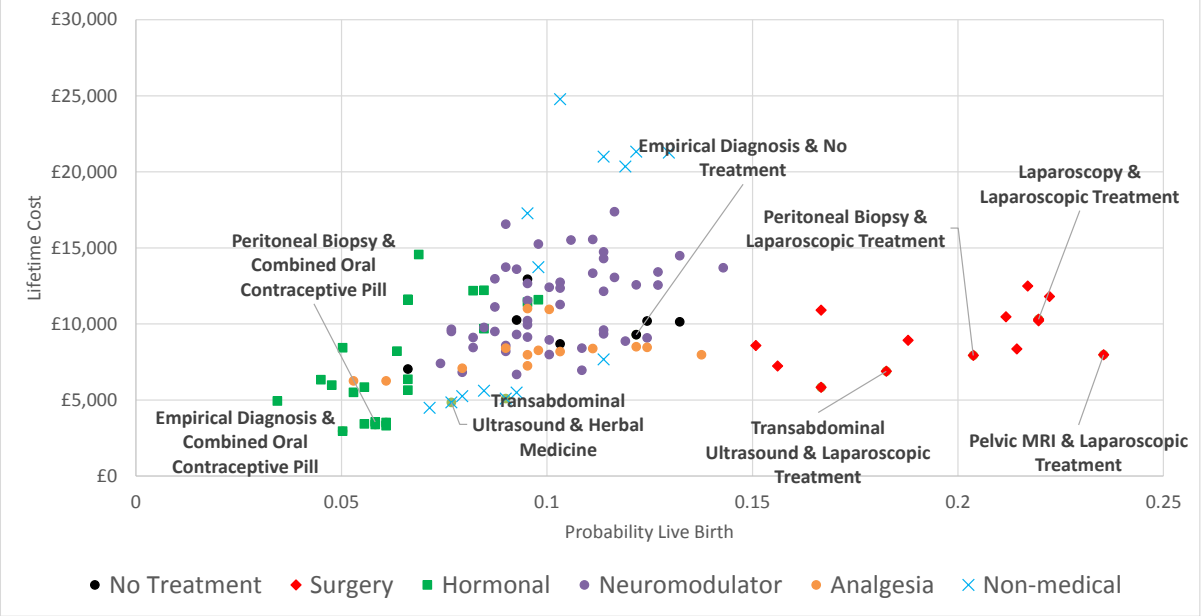
Figure 24: Base case analysis (fertility) – lifetime costs and QALYs (progestogen treatment excluded)



Source: Economic model

ISBN 978-1-4731-2661-9

Figure 25: Base case analysis (fertility) – lifetime costs and live births (progestogen treatment excluded)



Source: Economic model

The economic modelling, as demonstrated in Table 138, where every treatment more effective than the base case of doing nothing is a surgical technique – either laparoscopic excision on its own or laparoscopic excision plus hormonal therapy (although the addition of hormonal therapy harmed fertility, so there appears to be no health economic case for doing this).

Table 138: Base case analysis (fertility) – ICERs (progestogen treatment excluded)

Treatment	Cost	QALY	ICER	Probability cost-effective vs no treatment (£20,000 / QALY)	Probability cost-effective vs no treatment (£30,000 / QALY)
Empirical Diagnosis & No Treatment	£9,287.14	19.242	Base Case	100%	100%
Empirical Diagnosis & Combined Oral Contraceptive Pill	£2,951.71	19.083	Extendedly Dominated	100%	100%
Transabdominal Ultrasound & Combined Oral Contraceptive Pill	£3,382.11	19.092	Extendedly Dominated	100%	100%
Nerve fibre & Combined Oral	£3,512.64	19.106	Extendedly Dominated	100%	100%

ISBN 978-1-4731-2661-9

Treatment	Cost	QALY	ICER	Probability cost-effective vs no treatment (£20,000 / QALY)	Probability cost-effective vs no treatment (£30,000 / QALY)
Contraceptive Pill					
Peritoneal biopsy & Combined Oral Contraceptive Pill	£3,555.55	19.109	Extendedly Dominated	99%	100%
Transabdominal Ultrasound & Herbal Medicine	£4,829.00	19.132	Extendedly Dominated	100%	100%
Empirical Diagnosis & Herbal Medicine	£5,089.31	19.173	Extendedly Dominated	100%	100%
Transabdominal Ultrasound & Laparoscopic Treatment	£5,832.58	19.278	-£94,477.49	100%	100%
CA-125 & Laparoscopic Treatment	£6,876.99	19.319	Extendedly Dominated	100%	100%
Peritoneal biopsy & Laparoscopic Treatment	£7,930.55	19.389	Extendedly Dominated	100%	100%
Pelvic MRI & Laparoscopic Treatment	£7,966.94	19.448	£12,544.08	100%	100%
Laparoscopy & Laparoscopic Treatment	£10,307.01	19.450	£1,471,769.45	100%	100%

12.3.3 Evidence to recommendations

12.3.3.1 Relative value placed on the outcomes considered

The Committee considered that the most important outcomes were live births, spontaneous pregnancy (the presence of a foetal heartbeat) and miscarriage.

12.3.3.2 Consideration of clinical benefits and harms

The Committee agreed that the recommendations need to be interpreted in the context of the NICE's guideline on fertility problems (CG 156). The Committee also highlighted that the focus on spontaneous pregnancy as an outcome has limitations in that it excludes any assistive reproductive management. They therefore agreed to highlight that the management of endometriosis-related subfertility should have multidisciplinary team involvement with input

ISBN 978-1-4731-2661-9

from a fertility specialist. This means that women with endometriosis would receive the same assessment and management options (such as diagnostic tests including ovarian reserve testing, preoperative tests and assistive reproductive treatments) that other women would receive as specified in guideline CG156.

The Committee used the term surgical treatment in accordance with the evidence that was reviewed in the NMA, which included both surgical ablation and excision. The NMA showed an increase in the number of women with spontaneous pregnancy after surgery compared with women having diagnostic laparoscopy or on a waiting list (almost doubling the chances of pregnancy, RR 1.9 with a CrI from 1.3% to2.9%). There was evidence from other studies of an improvement in the number of women with live births after surgery. There was no evidence for any difference in miscarriage rate.

The studies in the NMA tended to include women with either minimal or mild endometriosis (AFS stage 1–2) or moderate or severe endometriosis (AFS stage 3–4), but there were insufficient data available to investigate fertility outcomes by severity of endometriosis. Therefore, using their knowledge and expertise, the Committee concluded that there was evidence to support the use of surgery in women with milder endometriosis to improve fertility. However, as the evidence was less clear regarding fertility outcomes for women with moderate to severe endometriosis, a comparison of surgery for assisted conception techniques had not included (as this was outside the scope of the current guideline) and as there were reports of peritonitis following egg collection and endometrioma, the Committee stipulated that surgery should only be considered (rather than offered) in conjunction with a fertility expert who would then be able to assess the ovarian reserve prior to surgery.

The Committee agreed that only those women with ovarian endometrioma who were undergoing laparoscopy should be offered cystectomy with excision of the cyst wall because this improves the chance of pregnancy. There was evidence to support ovarian cystectomy in the NMA but an amendment was made because the Committee believed that women who were selected for laparoscopy would be classed as having at least AFS stage 3 disease. Although the evidence for management of more severe endometriosis for fertility is less clear, the Committee agreed that where there were large ovarian endometriomas (>3 cm) these should be excised but that there were other considerations such as the effect of surgery on reducing ovarian reserve. Stimulation with fertility treatment where there were small endometriomas (under 3 cm) would still result in a reasonable egg yield.

The Committee agreed with the evidence pertaining to lower spontaneous pregnancy rates (not rates following assisted conception) in all women with endometriosis on hormonal treatments regardless of the severity of their condition and therefore recommended that hormonal treatment should not be offered postoperatively if fertility was the priority.

12.3.3.3 Consideration of economic benefits and harms

The Committee agreed that surgery offers the best chance of conception for a woman with endometriosis-related subfertility. In particular for those women who have endometriosis that does not involve the bowel, bladder or ureter. Surgery was also shown to be the most cost-effective management option for women trying to conceive.

The Committee discussed the use of assisted conception techniques as either an adjunct to or replacement for surgery. In particular, the Committee queried whether assisted reproductive treatment was cost-effective compared to surgery. However, this comparison had not been considered in drafting these recommendations, as this was outside the scope of the guideline. Analysis of the clinical and cost effectiveness of, and recommendations about, assisted conception techniques for all women with fertility problems are included in the NICE guideline on fertility (CG156).

ISBN 978-1-4731-2661-9

The Committee recognised that women with more severe endometriosis should be managed in collaboration with a fertility specialist in order that treatment options, including assisted conception techniques, should be considered and that there are costs attached to this. The Committee agreed that this would be a cost-effective option because it ensures that those preoperative assessments that are recommended in CG156 are carried out (for example, assessment of the ovarian reserve).

The resource impact of fertility recommendations is hard to estimate as there are several costs associated with pregnancy that only occur a long time after the initial treatment. However the Committee described how surgery was seen as the standard treatment in women who could not get pregnant due to endometriosis so the resource impact relative to current practice is likely to be low.

12.3.3.4 Quality of evidence

The NMA examined evidence on rates of spontaneous pregnancy and contained 16 studies. Of these, the risk of bias was high in 7, low in 2 and moderate in 7. GRADE criteria are currently not applied to NMA evidence, but – based on study quality – the body of the evidence would be no better than moderate quality. However, based on the effect size and the consensus and expertise of the Committee it was decided that an 'offer' recommendation should be made. Evidence about rates of live births and miscarriage was of very low quality according to GRADE criteria.

The Committee discussed the similarity of the protocols and evidence underlying the recommendations in CG156 and the evidence included in the NMA. Women included in the NMA are a subset of the guideline population because they presented with endometriosis as well as subfertility (women who had tried to conceive for 6 months) but would represent also a subset of the overall fertility guideline population because of their associated endometriosis morbidity. However, the section that is being updated in the fertility guideline is directly related to subfertility in endometriosis. Given the issue of subsets in each study and that direct diagnosis and full management of subfertility is not part of the scope, the evidence could be considered as somewhat indirect.

12.3.3.5 Other considerations

The Committee noted that a woman's symptoms would be an important factor in determining the treatment the woman would be offered (irrespective of severity of endometriosis) and the order in which assisted conception or surgery would be offered. If a woman was asymptomatic then she would be unlikely to be offered surgical laparoscopy to improve fertility because of the surgical risks of reducing ovarian reserve. The Committee considered that, dependent on other tests (for example, chlamydia antibodies), an asymptomatic woman would be more likely to be offered an ultrasound scan, tubal patency testing and expectant management before assisted conception techniques were offered. Women who had symptomatic endometriosis would be more likely to be offered laparoscopy. The Committee noted that in most of the studies included in the NMA, women did not have endometrioma but that the identification of endometrioma would also affect treatment decisions as removal of endometrioma may reduce ovarian reserve. They further noted that ovulation suppression is an attempt to delay recurrence of endometriosis, which, in the short term, could mean that conception can occur after the hormone treatment is discontinued.

The Committee discussed a comparison of surgery versus expectant management in a Cochrane Review that demonstrated no evidence of a benefit for pregnancy with either technique. Aspiration was associated with a greater number of mature oocytes retrieved and increased ovarian response compared to expectant management. Cystectomy was associated with a decreased ovarian response to controlled ovarian hyperstimulation with no evidence of an effect on the number of mature oocytes.

ISBN 978-1-4731-2661-9

The Committee highlighted that, in practice, there is multidisciplinary team involvement and usually an ultrasound scan and pre-operative checks of ovarian function should be assessed. There should be discussion with the woman about ovarian reserve and other factors before a decision about first-line treatment with assisted conception or surgery is taken.

Equality considerations and social value judgements regarding fertility treatments are considered in CG 156. Although discussed, the Committee did not consider that any amendments to recommendations were necessary to take account of such issues and agreed that the recommendations are intended to improve equality of access to the treatments covered here.

The recommendations are broadly in keeping with current practice and would not represent a significant change for typical practitioners or services in women with endometriosis. They acknowledged that not making a recommendation may limit the woman's options even if the full pathway of these women was not considered.

12.3.3.6 Key conclusions

The Committee agreed that there are limitations to the approach taken in that it addresses only a limited aspect of fertility management that is relevant to women with endometriosis. The Committee addressed this limitation by highlighting the context of NICE's guideline on fertility problems (CG 156) at the beginning of the recommendations which would safeguard that women with endometriosis would receive the same assessment and management options (such as diagnostic tests including ovarian reserve testing, preoperative tests, surgery and assistive reproductive treatments) that other women would receive.

The Committee agreed that there is strong RCT evidence to support offering surgery (surgical ablation or resection of endometriosis plus laparoscopic adhesiolysis) to women with minimal or mild (AFS stage 1 and 2) endometriosis, because this improves the chances of spontaneous pregnancy (i.e. pregnancy that is not related to assisted reproductive treatments) and only amended the previous recommendation to clarify this and to more fully describe mild or moderate endometriosis.

For women with ovarian endometriomas, the Committee, based on consensus, agreed that women who are having laparoscopy (that is, who are likely to have at least AFS stage 3 [moderate] endometriosis) should be offered laparoscopic ovarian cystectomy (excision of the endometrioma capsule) to improve the chances of pregnancy. The Committee noted that large ovarian endometriomas (>3 cm) should be excised, but acknowledged that there are risks associated with this type of surgery, such as reducing ovarian reserve, so it is not suitable for all women with endometriomas. It was further noted that for women with endometriomas the preoperative assessment of ovarian reserve in line with CG156 would be important because excision of endometriomas could impact on reserve and therefore decrease future fertility.

There is less convincing evidence to indicate that surgical treatment for women with moderate to severe (stage 3 or 4) endometriosis improves the chances of spontaneous pregnancy. The Committee also noted that there are adverse effects, such as endometrioma and peritonitis after egg collection, in this group. They therefore stipulated that surgery should be considered (rather than offered) and that a fertility expert should be involved.

The Committee agreed that evidence indicates that post-operative medical treatment does not improve spontaneous pregnancy rates in women with endometriosis, regardless of severity.

ISBN 978-1-4731-2661-9

12.3.4 Recommendations

The recommendations in this section should be interpreted within the context of NICE's guideline on fertility problems. The management of endometriosis-related subfertility should have multidisciplinary team involvement with input from a fertility specialist. This should include the recommended diagnostic fertility tests or preoperative tests, as well as other recommended fertility treatments such as assisted reproduction that are included in the NICE guideline on fertility problems.

50. **Offer excision or ablation of endometriosis plus adhesiolysis for endometriosis not involving the bowel, bladder or ureter, because this improves the chance of spontaneous pregnancy.**

51. **Offer laparoscopic ovarian cystectomy with excision of the cyst wall to women with endometriomas, because this improves the chance of spontaneous pregnancy and reduces recurrence. Take into account the woman's ovarian reserve. (Also see ovarian reserve testing in the NICE guideline on fertility problems.)**

52. **Discuss the benefits and risks of laparoscopic surgery as a treatment option for women who have deep endometriosis involving the bowel, bladder or ureter and who are trying to conceive (working with a fertility specialist). Topics to discuss may include:**
    - whether laparoscopic surgery may alter the chance of future pregnancy
    - the possible impact on ovarian reserve (also see ovarian reserve testing in the NICE guideline on fertility problems)
    - the possible impact on fertility if complications arise
    - alternatives to surgery
    - other fertility factors.

53. **Do not offer hormonal treatment to women with endometriosis who are trying to conceive, because it does not improve spontaneous pregnancy rates.**

ISBN 978-1-4731-2661-9

13 References

Abbott et al., 2004

Abbott, J., Hawe, J., Hunter, D., Holmes, M., Finn, P., Garry, R., Laparoscopic excision of endometriosis: a randomized, placebo-controlled trial, Fertility & Sterility, 82, 878-84, 2004

Abou-Setta et al., 2013

Abou-Setta, A. M., Houston, B., Al-Inany, H. G., Farquhar, C., Levonorgestrel-releasing intrauterine device (LNG-IUD) for symptomatic endometriosis following surgery, Cochrane Database of Systematic Reviews, 1, CD005072, 2013

Achana et al., 2014

Achana, F. A., Cooper, N. J., Bujkiewicz, S., Hubbard, S. J., Kendrick, D., Jones, D. R., Sutton, A. J., Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes, BMC medical research methodology, 14, 92, 2014

Acs et al., 2015

Acs, N., O'Brien, C., Jiang, P., Burke, J., Jimenez, R., Garner, E., Chwalisz, K., Treatment of endometriosis-associated pain with elagolix, an oral gnrh antagonist: Results from a phase 2, randomized controlled study, Journal of Endometriosis and Pelvic Pain Disorders, 7, 56-62, 2015

Agarwal et al., 1997

Agarwal, S. K., Hamrang, C., Henzl, M. R., Judd, H. L., Nafarelin vs. leuprolide acetate depot for endometriosis. Changes in bone mineral density and vasomotor symptoms. Nafarelin Study Group, Journal of Reproductive Medicine, 42, 413-23, 1997

Alborzi et al., 2011

Alborzi, S., Hamedi, B., Omidvar, A., Dehbashi, S., Alborzi, S., Alborzi, M., A comparison of the effect of short-term aromatase inhibitor (letrozole) and GnRH agonist (triptorelin) versus case control on pregnancy rate and symptom and sign recurrence after laparoscopic treatment of endometriosis, Archives of Gynecology and Obstetrics, 284, 105-110, 2011

Allaire et al., 2014

Allaire, C., MacRae, G. S., Nishi, C., Chen, I., Hospital-related costs for endometriosis in Canada, Journal of Minimally Invasive Gynecology, 1), S36-S37, 2014

American Society for Reproductive Medicine, 1997

American Society for Reproductive Medicine, Revised American Society for Reproductive Medicine classification of endometriosis: 1996. Fertil Steril, 67, 817–21, 1997

Anonymous, 1992

Anonymous, Nafarelin for endometriosis: a large-scale, danazol-controlled trial of efficacy and safety, with 1-year follow-up. The Nafarelin European Endometriosis Trial Group (NEET), Fertility & Sterility, 57, 514-22, 1992

ISBN 978-1-4731-2661-9

Anonymous, 1996

Anonymous, Goserelin depot versus danazol in the treatment of endometriosis the Australian/New Zealand experience, Australian & New Zealand Journal of Obstetrics & Gynaecology, 36, 55-60, 1996

Araujo et al., 2009

Araujo, D., Passos, R. B. F., Souza, C. P. R., Silva, A. P., Marques, M., Cost effectiveness analysis of goserelin empiric therapy for deep endometriosis treatment, Value in health, 12 (3), A164, 2009

Arguedas et al., 2002

Arguedas, M. R., Heudebert, G. H., Stinnett, A. A., Wilcox, C. M., Biliary stents in malignant obstructive jaundice due to pancreatic carcinoma: a cost-effectiveness analysis, American Journal of Gastroenterology, 97, 898-904, 2002

Aris, 2010

Aris, A., Endometriosis-associated ovarian cancer: A ten-year cohort study of women living in the Estrie Region of Quebec, Canada, Journal of ovarian research, 3 (1) (no pagination), 2010

Arrive et al., 1989

Arrive, L., Hricak, H., Martin, M. C., Pelvic endometriosis: MR imaging, Radiology, 171, 687-92, 1989

Attilia et al., 2013

Attilia, B. K., Griffiths, A. N., Penketh, R. J., Regression analysis of predicted and actual operating duration for advanced laparoscopic excisional endometriosis surgery, Journal of Minimally Invasive Gynecology, 1), S7-S8, 2013

Avxentyeva, 2013

Avxentyeva, M., Cost-minimization analysis of dienogest versus gonadotrophinreleasing hormone analogues or dydrogesterone for endometriosis treatment in Russia, Value in health, 16 (7), A334, 2013

Bahr et al., 2006

Bahr, A., de Parades, V., Gadonneix, P., Etienney, I., Salet-Lizee, D., Villet, R., Atienza, P., Endorectal ultrasonography in predicting rectal wall infiltration in patients with deep pelvic endometriosis: a modern tool for an ancient disease, Diseases of the Colon & Rectum, 49, 869-75, 2006

Balasch et al., 1996

Balasch, J., Creus, M., Fabregues, F., Carmona, F., Ordi, J., Martinez-Roman, S., Vanrell, J. A., Visible and non-visible endometriosis at laparoscopy in fertile and infertile women and in patients with chronic pelvic pain: a prospective study, Human Reproduction, 11, 387-91, 1996

Ballard et al., 2006

Ballard, K., Lowton, K., Wright, J., What's the delay? A qualitative study of women's experiences of reaching a diagnosis of endometriosis, Fertility & Sterility, 86, 1296-301, 2006

Bayer et al., 1988

ISBN 978-1-4731-2661-9

Bayer, S. R., Seibel, M. M., Saffan, D. S., Berger, M. J., Taymor, M. L., Efficacy of danazol treatment for minimal endometriosis in infertile women. A prospective, randomized study, Journal of Reproductive Medicine, 33, 179-83, 1988

Becker et al., 2010

Becker, I., Parlayan, S., Baumgarten, C., Dakkak, R., Rosenow, G., Von Kleinsorgen, C., Halis, G., Fuhr, N., Ebert, A. D., Inpatient costs for women with deep infiltrating endometriosis: A comparison of data from the years 2006 and 2007 of the Deutsche Endometriosezentrum Berlin (DEZB) Stufe III, Archives of Gynecology and Obstetrics, 282, S175, 2010

Bergqvist et al., 1997

Bergqvist, A., Jacobson, J., Harris, S., A double-blind randomized study of the treatment of endometriosis with nafarelin or nafarelin plus norethisterone, Gynecological Endocrinology, 11, 187-194, 1997

Bergqvist et al., 1998

Bergqvist, A., Bergh, T., Hogstrom, L., Mattsson, S., Nordenskjold, F., Rasmussen, C., Effects of triptorelin versus placebo on the symptoms of endometriosis, Fertility & Sterility, 69, 702-8, 1998

Bergqvist et al., 2000

Bergqvist, A., Scandet Group, A comparative study of the acceptability and effect of goserelin and nafarelin on endometriosis, Gynecological Endocrinology, 14, 425-32, 2000

Bianchi et al., 1999

Bianchi, S., Busacca, M., Agnoli, B., Candiani, M., Calia, C., Vignali, M., Effects of 3 month therapy with danazol after laparoscopic surgery for stage III/IV endometriosis: a randomized study, Human Reproduction, 14, 1335-7, 1999

Biberoglu & Behrman, 1981

Biberoglu, K. O., Behrman, S. J., Dosage aspects of danazol therapy in endometriosis: short-term and long-term effectiveness, American Journal of Obstetrics and Gynecology, 139, 645-654, 1981

Bodner et al., 1996

Bodner, C., Vale, L., Ratcliffe, J., Farrar, S., Using economics alongside medical audit. A case study of the management of endometriosis, Health Bulletin, 54, 204-11, 1996

Bostrom et al., 2012

Bostrom, P., Lovkvist, L., Gustafsson, M., Alexandersson, O., Bruse, C., Liffner, C., Holmberg, J., Edlund, M., Olovsson, M., Cost of illness in women with endometriosis, Value in health, 15 (7), A538, 2012

Bradburn et al., 2007

Bradburn, M. J., Deeks, J. J., Berlin, J. A., Russell Localio, A., Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events, Statistics in medicine, 26, 53-77, 2007

Brandes et al., 2010

Brandes, I., Kleine-Budde, K., Halis, G., Quality of life results from the EndoCost study, Archives of Gynecology and Obstetrics, 282, S10, 2010

ISBN 978-1-4731-2661-9

Brinton et al., 1997

Brinton, L. A., Gridley, G., Persson, I., Baron, J., Bergqvist, A., Cancer risk after a hospital discharge diagnosis of endometriosis, American Journal of Obstetrics and Gynecology, 176, 572-579, 1997

Brinton et al., 2004

Brinton, L. A., Lamb, E. J., Moghissi, K. S., Scoccia, B., Althuis, M. D., Mabie, J. E., Westhoff, C. L., Ovarian cancer risk associated with varying causes of infertility, Fertility and Sterility, 82, 405-414, 2004

Brinton et al., 2005

Brinton, L. A., Sakoda, L. C., Sherman, M. E., Frederiksen, K., Kjaer, S. K., Graubard, B. I., Olsen, J. H., Mellemkjaer, L., Relationship of benign gynecologic diseases to subsequent risk of ovarian and uterine tumors, Cancer Epidemiology Biomarkers and Prevention, 14, 2929-2935, 2005

Brinton et al., 2005a

Brinton, L. A., Westhoff, C. L., Scoccia, B., Lamb, E. J., Althuis, M. D., Mabie, J. E., Moghissi, K. S., Causes of infertility as predictors of subsequent cancer risk, Epidemiology, 16, 500-7, 2005

British Society for Gynaecological Endoscopy et al., 2016

British Society for Gynaecological Endoscopy, Accredited endometriosis centres, 2016 http://bsge.org.uk/centre/

Brown et al., 2010

Brown, J., Pan, A., Hart, R. J., Gonadotrophin-releasing hormone analogues for pain associated with endometriosis, Cochrane Database of Systematic Reviews, 12, CD008475-, 2010

Brown et al., 2012

Brown, J., Kives, S., Akhtar, M., Progestagens and anti-progestagens for pain associated with endometriosis, Cochrane Database of Systematic Reviews, 3, CD002122, 2012

Buchweitz et al., 2003

Buchweitz, O., Poel, T., Diedrich, K., Malik, E., The diagnostic dilemma of minimal and mild endometriosis under routine conditions, Journal of the American Association of Gynecologic Laparoscopists, 10, 85-9, 2003

Buis et al., 2013

Buis, C. C., Van Leeuwen, F. E., Mooij, T. M., Burger, C. W., Omega Project Group, Increased risk for ovarian cancer and borderline ovarian tumours in subfertile women with endometriosis, Human Reproduction, 28, 3358-69, 2013

Burry et al., 1989

Burry, K. A., Patton, P. E., Illingworth, D. R., Metabolic changes during medical treatment of endometriosis: nafarelin acetate versus danazol.[Erratum appears in Am J Obstet Gynecol 1989 Dec;161(6 Pt 1):1755], American Journal of Obstetrics and Gynecology, 160, 1454-1459, 1989

Burry, 1992

ISBN 978-1-4731-2661-9

Burry, K. A., Nafarelin in the management of endometriosis: quality of life assessment, American Journal of Obstetrics & Gynecology, 166, 735-9, 1992

Busacca et al., 2001

Busacca, M., Somigliana, E., Bianchi, S., De Marinis, S., Calia, C., Candiani, M., Vignali, M., Post-operative GnRH analogue treatment after conservative surgery for symptomatic endometriosis stage III-IV: a randomized controlled trial, Human Reproduction, 16, 2399-402, 2001

Calhaz-Jorge et al., 2004

Calhaz-Jorge, C., Mol, B. W., Nunes, J., Costa, A. P., Clinical predictive factors for endometriosis in a Portuguese infertile population, Human Reproduction, 19, 2126-31, 2004

Carmona et al., 2011

Carmona, F., Martinez-Zamora, M. A., Rabanal, A., Martinez-Roman, S., Balasch, J., Ovarian cystectomy versus laser vaporization in the treatment of ovarian endometriomas: a randomized clinical trial with a five-year follow-up, Fertility & Sterility, 96, 251-4, 2011

Carr et al., 2014

Carr, B., Dmowski, W. P., O'Brien, C., Jiang, P., Burke, J., Jimenez, R., Garner, E., Chwalisz, K., Elagolix, an oral GnRH antagonist, versus subcutaneous depot medroxyprogesterone acetate for the treatment of endometriosis: effects on bone mineral density, Reproductive Sciences, 21, 1341-51, 2014

Chang et al., 2014

Chang, W. H., Wang, K. C., Lee, W. L., Huang, N., Chou, Y. J., Feng, R. C., Yen, M. S., Huang, B. S., Guo, C. Y., Wang, P. H., Endometriosis and the subsequent risk of epithelial ovarian cancer, Taiwanese Journal of Obstetrics and Gynecology, 53, 530-535, 2014

Chapron et al., 1998

Chapron, C., Querleu, D., Bruhat, M. A., Madelenat, P., Fernandez, H., Pierre, F., Dubuisson, J. B., Surgical complications of diagnostic and operative gynaecological laparoscopy: a series of 29,966 cases, Human Reproduction, 13, 867-72, 1998

Chatman & Zbella, 1987

Chatman, D. L., Zbella, E. A., Biopsy in laparoscopically diagnosed endometriosis, Journal of Reproductive Medicine, 32, 855-7, 1987

Chen et al., 2012

Chen, L, Lin, Y, Yuan, L, Huang, H, Abdominal acupuncture in treating 70 cases of endometriosis dysmenorrhoea, International Journal of Clinical Acupuncture, 21, 100-2., 2012

Chvatal et al., 2010

Chvatal, R., Reichert, B., Renner, S., Oppelt, P., Cost of inpatient treatment for endometriosis in Germany: An analysis based on the G-DRG-coding, Archives of Gynecology and Obstetrics, 282, S192, 2010

Cornillie et al., 1990

Cornillie, F. J., Oosterlynck, D., Lauweryns, J. M., Koninckx, P. R., Deeply infiltrating pelvic endometriosis: histology and clinical significance, Fertility & Sterility, 53, 978-83, 1990

ISBN 978-1-4731-2661-9

Cosson et al., 2002

Cosson, M., Querleu, D., Donnez, J., Madelenat, P., Koninckx, P., Audebert, A., Manhes, H., Dienogest is as effective as triptorelin in the treatment of endometriosis after laparoscopic surgery: results of a prospective, multicenter, randomized study, Fertility and Sterility, 77, 684-692, 2002

Cox et al., 2003

Cox, H., Henderson, L., Andersen, N., Cagliarini, G., Ski, C., Focus group study of endometriosis: struggle, loss and the medical merry-go-round, International Journal of Nursing Practice, 9, 2-9, 2003

Cox et al., 2003a

Cox, H., Henderson, L., Wood, R., Cagliarini, G., Learning to take charge: women's experiences of living with endometriosis, Complementary Therapies in Nursing & Midwifery, 9, 62-8, 2003

Crosignani et al., 2006

Crosignani, P. G., Luciano, A., Ray, A., Bergqvist, A., Subcutaneous depot medroxyprogesterone acetate versus leuprolide acetate in the treatment of endometriosis-associated pain, Human Reproduction, 21, 248-56, 2006

Culley et al., 2013

Culley, L.; Hudson, N.; Mitchell, H.; Law, C.; Denny, E.; Raine-Fenning, N., Funded by the UK Economic and Social Research Council, Endometriosis: improving the wellbeing of couples. Summary report and recommendations, 2013

Dan & Limin, 2013

Dan, H., Limin, F., Laparoscopic ovarian cystectomy versus fenestration/coagulation or laser vaporization for the treatment of endometriomas: a meta-analysis of randomized controlled trials, Gynecologic & Obstetric Investigation, 76, 75-82, 2013

Davis et al., 2007

Davis, L., Kennedy, S. S., Moore, J., Prentice, A., Modern combined oral contraceptives for pain associated with endometriosis, Cochrane Database of Systematic Reviews, CD001019, 2007

De Almeida Filho et al., 2008

De Almeida Filho, D. P., de Oliveira, L. J., Amaral, V. F., Accuracy of laparoscopy for assessing patients with endometriosis, Sao Paulo Medical Journal, 126, 305-308, 2008

De Graaff et al., 2013

De Graaff, A. A., D'Hooghe, T. M., Dunselman, G. A., Dirksen, C. D., Hummelshoj, L., Werf EndoCost Consortium, Simoens, S., The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey, Human Reproduction, 28, 2677-85, 2013

De Graaff et al., 2015

De Graaff, A. A., Dirksen, C. D., Simoens, S., De Bie, B., Hummelshoj, L., D'Hooghe, T. M., Dunselman, G. A., Quality of life outcomes in women with endometriosis are highly influenced by recruitment strategies, Human Reproduction, 30, 1331-41, 2015

Denny & Mann, 2007

ISBN 978-1-4731-2661-9

Denny, E., Mann, C. H., Endometriosis-associated dyspareunia: the impact on women's lives, Journal of Family Planning & Reproductive Health Care, 33, 189-93, 2007

Denny, 2004

Denny, E., Women's experience of endometriosis, Journal of Advanced Nursing, 46, 641-8, 2004

Denny, 2009

Denny, E., I never know from one day to another how I will feel: pain and uncertainty in women with endometriosis, Qualitative Health Research, 19, 985-95, 2009

DerSimonian, 1986

DerSimonian R., Laird N. M., Meta-analysis in clinical trials, Controlled Clinical Trials, 7, 177–188, 1986

De Sousa et al., 2016

De Sousa, Tatiane Regina, De Souza, Bruna Cruz, Zomkowisk, Kamilla, da Rosa, Priscila Cibils, Sperandio, Fabiana Flores, The effect of acupuncture on pain, dyspareunia, and quality of life in Brazilian women with endometriosis: A randomized clinical trial, Complementary Therapies in Clinical Practice, 25, 114-121, 2016

Devlin & Parkin, 2003

Devlin, N., Parkin, D., Funding fertility: issues in the allocation and distribution of resources to assisted reproduction technologies, Human Fertility (Camb), 6 Suppl 1, S2-6, 2003

D'Hooghe et al., 2011

D'Hooghe, T., Hummelshoj, L., Dunselman, G. A. J., Dirksen, C. D., EndoCost Consortium, W. E. R. F., Simoens, S., The societal burden of endometriosis equals that of other chronic diseases: A prospective international study of costs, Human Reproduction, 26, i206-i207, 2011

D'Hooghe et al., 2012

D'Hooghe, T., Dirksen, C. D., Dunselman, G. A. J., De Graaff, A., EndoCost Consortium, W. E. R. F., Simoens, S., The costs of endometriosis: It's the economy, stupid, Fertility and sterility, 1), S218-S219, 2012

Diamond et al., 2014

Diamond, M. P., Carr, B., Dmowski, W. P., Koltun, W., O'Brien, C., Jiang, P., Burke, J., Jimenez, R., Garner, E., Chwalisz, K., Elagolix treatment for endometriosis-associated pain: results from a phase 2, randomized, double-blind, placebo-controlled study, Reproductive Sciences, 21, 363-71, 2014

Dias et al., 2010

Dias, S., Welton, N. J., Caldwell, D. M., Ades, A. E., Checking consistency in mixed treatment comparison meta-analysis, Statistics in medicine, 29, 932-44, 2010

Dixon-Woods et al., 2005

Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., Sutton, A., Synthesising qualitative and quantitative evidence: a review of possible methods, Journal of Health Services & Research Policy, 10, 45-53, 2005

Dlugi et al., 1990

ISBN 978-1-4731-2661-9

Dlugi, A. M., Miller, J. D., Knittle, J., Lupron depot (leuprolide acetate for depot suspension) in the treatment of endometriosis: A randomized, placebo-controlled, double-blind study, Fertility and sterility, 54, 419-427, 1990

Dmowski, 1989

Dmowski, W. P., Comparitive Study of Buserelin Versus Danazol in the Management of Endometriosis, Gynecological Endocrinology, 3, 21-31, 1989

Dousset et al., 2010

Dousset, B., Leconte, M., Borghese, B., Millischer, A. E., Roseau, G., Arkwright, S., Chapron, C., Complete surgery for low rectal endometriosis: Long-term results of a 100-case prospective study, Annals of Surgery, 251, 887-895, 2010

Duffy et al., 2014

Duffy, J. M., Arambage, K., Correa, F. J., Olive, D., Farquhar, C., Garry, R., Barlow, D. H., Jacobson, T. Z., Laparoscopic surgery for endometriosis, Cochrane Database of Systematic Reviews, 4, CD011031, 2014

Ebert et al., 2009

Ebert, A. D., Rosenow, G., Kruger, K., Liehr, R. M., Adam, U., Braunig, P., Haselmann, J., Freitag, A., Papadopolous, T., Von Kleinsorgen, C., Development of centres of excellence for endometriosis - The Berlin experiences, Molecular Human Reproduction, 24, i54-i55, 2009

Edelstam et al., 2012

Edelstam, G., Sjosten, A., Jablonowska, B., Kjellberg, S., Spira, J., Pertubation with lidocaine - a non-hormonal, long-term treatment of dysmenorrhoea due to endometriosis, Sexual and Reproductive Healthcare, 3, 93-94, 2012

El Bishry et al., 2008

El Bishry, G., Tselos, V., Pathi, A., Correlation between laparoscopic and histological diagnosis in patients with endometriosis, Journal of Obstetrics & Gynaecology, 28, 511-5, 2008

Emmert et al., 1998

Emmert, C., Romann, D., Riedel, H. H., Endometriosis diagnosed by laparoscopy in adolescent girls, Archives of Gynecology & Obstetrics, 261, 89-93, 1998

Shakiba et al. 2008

Shakiba K, Bena JF, McGill KM, Minger J, Falcone T. Surgical treatment of endometriosis: a 7-year follow-up on the requirement for further surgery. Obstetrics and Gynecology, 111, 1285-92, 2008Falcone et al., 2008

Fedele et al., 1989b

Fedele, L., Bianchi, S., Arcaini, L., Vercellini, P., Candiani, G. B., Buserelin versus danazol in the treatment of endometriosis-associated infertility, American Journal of Obstetrics and Gynecology, 161, 871-876, 1989

Fedele et al., 1992

Fedele, L., Parazzini, F., Radici, E., Bocciolone, L., Bianchi, S., Bianchi, C., Candiani, G. B., Buserelin acetate versus expectant management in the treatment of infertility associated with minimal or mild endometriosis: a randomized clinical trial, American Journal of Obstetrics and Gynecology, 166, 1345-1350, 1992

ISBN 978-1-4731-2661-9

Fernandez et al., 2004

Fernandez, H., Lucas, C., Hedon, B., Meyer, J. L., Mayenga, J. M., Roux, C., One year comparison between two add-back therapies in patients treated with a GnRH agonist for symptomatic endometriosis: a randomized double-blind trial, Human Reproduction, 19, 1465-71, 2004

Fernandez et al., 2006

Fernandez, I., Reid, C., Dziurawiec, S., Living with endometriosis: the perspective of male partners, Journal of Psychosomatic Research, 61, 433-8, 2006

Fernando et al., 2013

Fernando, S., Soh, P. Q., Cooper, M., Evans, S., Reid, G., Tsaltas, J., Rombauts, L., Reliability of visual diagnosis of endometriosis, Journal of Minimally Invasive Gynecology, 20, 783-9, 2013

Ferracini & Nakada, 2013

Ferracini, M., Nakada, C. P., Cost minimization analysis of the dienogest use in patients with endometris under Brazilian public and private perspective, Value in health, 16 (7), A334-A335, 2013

Ferreira et al., 2010

Ferreira, R. A., Vieira, C. S., Rosa, E. Silva J. C., Rosa-e-Silva, A. C., Nogueira, A. A., Ferriani, R. A., Effects of the levonorgestrel-releasing intrauterine system on cardiovascular risk markers in patients with endometriosis: a comparative study with the GnRH analogue, Contraception, 81, 117-22, 2010

Flower et al., 2011

Flower, A., Lewith, G. T., Little, P., A feasibility study exploring the role of Chinese herbal medicine in the treatment of endometriosis, Journal of Alternative & Complementary Medicine, 17, 691-9, 2011

Flower et al., 2012

Flower, A., Liu, J. P., Lewith, G., Little, P., Li, Q., Chinese herbal medicine for endometriosis, Cochrane Database of Systematic Reviews, 5, CD006568, 2012

Fourquet et al., 2010

Fourquet, J., Gao, X., Zavala, D., Orengo, J. C., Abac, S., Ruiz, A., Laboy, J., Flores, I., Patients' report on how endometriosis affects health, work, and daily life, Fertility & Sterility, 93, 2424-8, 2010

Franke et al., 2000

Franke, H. R., Van De Weijer, P. H., Pennings, T. M., Van Der Mooren, M. J., Gonadotropin-releasing hormone agonist plus "add-back" hormone replacement therapy for treatment of endometriosis: a prospective, randomized, placebo-controlled, double-blind trial, Fertility and Sterility, 74, 534-539, 2000

Fraser et al., 1991

Fraser, I. S., Shearman, R. P., Jansen, R. P., Sutherland, P. D., A comparative treatment trial of endometriosis using the gonadotrophin-releasing hormone agonist, nafarelin, and the synthetic steroid, danazol, Australian and New Zealand Journal of Obstetrics and Gynaecology, 31, 158-163, 1991

ISBN 978-1-4731-2661-9

Fuldeore et al., 2010

Fuldeore, M., Wu, N., Boulanger, L., Chwalisz, K., Marx, S., Prevalence rate and direct cost of surgical procedures among women with newly diagnosed endometriosis, Human Reproduction, 26th Annual Meeting of the European Society of Human Reproduction and Embryology, ESHRE Rome Italy. Conference Start, i158-i159, 2010

Fuldeore et al., 2011

Fuldeore, M., Chwalisz, K., Marx, S., Wu, N., Boulanger, L., Ma, L., Lamothe, K., Surgical procedures and their cost estimates among women with newly diagnosed endometriosis: a US database study, Journal of Medical Economics, 14, 115-123, 2011

Fuldeore et al., 2015

Fuldeore, M., Yang, H., Du, E. X., Soliman, A. M., Wu, E. Q., Winkel, C., Healthcare utilization and costs in women diagnosed with endometriosis before and after diagnosis: a longitudinal analysis of claims databases, Fertility & Sterility, 103, 163-71, 2015

Furness et al., 2011

Furness, S., Yap, C., Farquhar, C., Cheong, Y. C., Pre and post-operative medical therapy for endometriosis surgery, Cochrane Database of Systematic Reviews, -, 2011

Gad & Badroui, 2012

Gad, M. S., Badroui, M. H. H., Evidence-based therapy for infertility associated with early stage endometriosis, International journal of gynaecology and obstetrics, 119, 2012

Gerlinger et al, 2012

Gerlinger, Christoph, Schumacher, Ulrike, Wentzeck, René, Uhl-Hochgräber, Kerstin, Solomayer, Erich F., Schmitz, Heinz, Faustmann, Thomas, Seitz, Christian, How can we measure endometriosis-associated pelvic pain?, 4, 109, 2012

Gilmour et al., 2008

Gilmour, J. A., Huntington, A., Wilson, H. V., The impact of endometriosis on work and social participation, International Journal of Nursing Practice, 14, 443-8, 2008

Gomes et al., 2007

Gomes, M. K., Ferriani, R. A., Rosa e Silva, J. C., Japur de Sa Rosa e Silva, A. C., Vieira, C. S., Candido dos Reis, F. J., The levonorgestrel-releasing intrauterine system and endometriosis staging, Fertility & Sterility, 87, 1231-4, 2007

Granese et al., 2015

Granese, R., Perino, A., Calagna, G., Saitta, S., De Franciscis, P., Colacurci, N., Triolo, O., Cucinella, G., Gonadotrophin-releasing hormone analogue or dienogest plus estradiol valerate to prevent pain recurrence after laparoscopic surgery for endometriosis: a multi-center randomized trial, Acta Obstetricia et Gynecologica Scandinavica, 94, 637-45, 2015

Gupta et al., 2016

Gupta, Devashana, Hull, Louise M., Fraser, Ian, Miller, Laura, Bossuyt, M. M. Patrick, Johnson, Neil, Nisenblat, Vicki, Endometrial biomarkers for the non-invasive diagnosis of endometriosis, Cochrane Database of Systematic Reviews, 2016

Guzick et al., 2011

ISBN 978-1-4731-2661-9

Guzick, D. S., Huang, L. S., Broadman, B. A., Nealon, M., Hornstein, M. D., Randomized trial of leuprolide versus continuous oral contraceptives in the treatment of endometriosis-associated pelvic pain, Fertility & Sterility, 95, 1568-73, 2011

Hamedi et al., 2011

Hamedi, B., Omidvar, A., Dehbashi, S., Alborzi, S., Alborzi, M., A comparison of the effect of short-term aromatase inhibitor (letrozole) and GnRH agonist (triptorelin) versus case control on pregnancy rate and symptom and sign recurrence after laparoscopic treatment of endometriosis, Archives of Gynecology and Obstetrics, 284, 105-110, 2011

Harada et al., 2008

Harada, T., Momoeda, M., Taketani, Y., Hoshiai, H., Terakawa, N., Low-dose oral contraceptive pill for dysmenorrhoea associated with endometriosis: a placebo-controlled, double-blind, randomized trial, Fertility & Sterility, 90, 1583-8, 2008

Harada et al., 2009

Harada, T., Momoeda, M., Taketani, Y., Aso, T., Fukunaga, M., Hagino, H., Terakawa, N., Dienogest is as effective as intranasal buserelin acetate for the relief of pain symptoms associated with endometriosis--a randomized, double-blind, multicenter, controlled trial, Fertility & Sterility, 91, 675-81, 2009

Healey et al., 2010

Healey, M., Ang, W. C., Cheng, C., Surgical treatment of endometriosis: a prospective randomized double-blinded trial comparing excision and ablation, Fertility & Sterility, 94, 2536-40, 2010

Healey et al., 2014

Healey, M., Cheng, C., Kaur, H., To excise or ablate endometriosis? A prospective randomized double-blinded trial after 5-year follow-up, Journal of Minimally Invasive Gynecology, 21, 999-1004, 2014

Heinrichs & Henzl, 1998

Heinrichs, W. L., Henzl, M. R., Human issues and medical economics of endometriosis. Three- vs. six-month GnRH-agonist therapy, Journal of Reproductive Medicine, 43, 299-308, 1998

Henzl, 1989

Henzl, M. R., Role of nafarelin in the management of endometriosis, Journal of Reproductive Medicine, 34, 1021-4, 1989

Hornstein et al., 1998

Hornstein, M. D., Surrey, E. S., Weisberg, G. W., Casino, L. A., Leuprolide acetate depot and hormonal add-back in endometriosis: a 12-month study. Lupron Add-Back Study Group, Obstetrics & Gynecology, 91, 16-24, 1998

Hughes et al., 2007

Hughes, E., Brown, J., Collins, J. J., Farquhar, C., Fedorkow, D. M., Vandekerckhove, P., Ovulation suppression for endometriosis, Cochrane Database of Systematic Reviews, #2007. Article Number, -, 2007

Hunault et al., 2005

ISBN 978-1-4731-2661-9

Hunault, C. C., Laven, J. S., Van Rooij, I. A., Eijkemans, M. J., te Velde, E. R., Habbema, J. D., Prospective validation of two models predicting pregnancy leading to live birth among untreated subfertile couples, Human Reproduction, 20, 1636-41, 2005

Irvine et al., 1996

Irvine, S., Cawood, E., Richardson, D., MacDonald, E., Aitken, J., Evidence of deteriorating semen quality in the United Kingdom: birth cohort study in 577 men in Scotland over 11 years, BMJ, 312, 467-71, 1996

Jansen & Russell, 1986

Jansen, R. P., Russell, P., Nonpigmented endometriosis: clinical, laparoscopic, and pathologic definition, American Journal of Obstetrics & Gynecology, 155, 1154-9, 1986

Jelley, 1988

Jelley, R. Y., Multicentre Open Comparative Study of Buserelin and Danazol in the Treatment of Endometriosis, British journal of clinical practice. 41, Suppl, 48, 64-8, 1988

Jones et al., 2004

Jones, G., Jenkinson, C., Kennedy, S., The impact of endometriosis upon quality of life: a qualitative analysis, Journal of Psychosomatic Obstetrics & Gynecology, 25, 123-33, 2004

Kauppila & Ronnberg, 1985

Kauppila, A., Ronnberg, L., Naproxen sodium in dysmenorrhoea secondary to endometriosis, Obstetrics & Gynecology, 65, 379-83, 1985

Keltz et al., 1995

Keltz, M. D., Kliman, H. J., Arici, A. M., Olive, D. L., Endosalpingiosis found at laparoscopy for chronic pelvic pain, Fertility & Sterility, 64, 482-5, 1995

Kennedy et al., 1990

Kennedy, S. H., Williams, I. A., Brodribb, J., Barlow, D. H., Shaw, R. W., A comparison of nafarelin acetate and danazol in the treatment of endometriosis, Fertility and Sterility, 53, 998-1003, 1990

Kent et al., 2016

Kent, A., Shakir, F., Rockall, T., Haines, P., Pearson, C., Rae-Mitchell, W., Jan, H., Laparoscopic surgery for severe rectovaginal endometriosis compromising the bowel: a prospective cohort study, Journal of Minimally Invasive Gynecology, 23, 526-34, 2016

Kiesel et al., 1996

Kiesel, L., Schweppe, K. W., Sillem, M., Siebzehnrubl, E., Should add-back therapy for endometriosis be deferred for optimal results?, British Journal of Obstetrics and Gynaecology, 103 Suppl 14, 15-17, 1996

Kiilholma et al., 1995

Kiilholma, P., Tuimala, R., Kivinen, S., Korhonen, M., Hagman, E., Comparison of the gonadotropin-releasing hormone agonist goserelin acetate alone versus goserelin combined

ISBN 978-1-4731-2661-9

with estrogen-progestogen add-back therapy in the treatment of endometriosis, Fertility and Sterility, 64, 903-908, 1995

Kind et al., 1999

Kind, P., Hardman, G., Macran, S., University of York. Centre for Health Economics, UK Population Norms for EQ-5D, 1999

Klein et al., 2014

Klein, S., D'Hooghe, T., Meuleman, C., Dirksen, C., Dunselman, G., Simoens, S., What is the societal burden of endometriosis-associated symptoms? a prospective Belgian study, Reproductive Biomedicine Online, 28, 116-24, 2014

Kobayashi et al., 2007

Kobayashi, H., Sumimoto, K., Moniwa, N., Imai, M., Takakura, K., Kuromaki, T., Morioka, E., Arisawa, K., Terao, T., Risk of developing ovarian cancer among women with ovarian endometrioma: a cohort study in Shizuoka, Japan, International Journal of Gynecological Cancer, 17, 37-43, 2007

Kobayashi et al., 2008

Kobayashi, H., Sumimoto, K., Kitanaka, T., Yamada, Y., Sado, T., Sakata, M., Yoshida, S., Kawaguchi, R., Kanayama, S., Shigetomi, H., Haruta, S., Tsuji, Y., Ueda, S., Terao, T., Ovarian endometrioma--risks factors of ovarian cancer development, European Journal of Obstetrics, Gynecology, & Reproductive Biology, 138, 187-93, 2008

Kok et al., 2015

Kok, V. C., Tsai, H. J., Su, C. F., Lee, C. K., The risks for ovarian, endometrial, breast, colorectal, and other cancers in women with newly diagnosed endometriosis or adenomyosis: a population-based study, International Journal of Gynecological Cancer, 25, 968-76, 2015

Koltermann et al., 2013

Koltermann, K. C., Mpinou, L. E., Niedobitek-Kreuter, G., Kruger, K., Thiel-Moder, U., Mechsner, S., Ebert, A. D., Symptomatic and asymptomatic bowel endometriosis in health economic focus in Germany, Journal of Endometriosis, 5, S62-S63, 2013

Koninckx et al., 1996a

Koninckx, P. R., Timmermans, B., Meuleman, C., Penninckx, F., Complications of CO2-laser endoscopic excision of deep endometriosis, Human Reproduction, 11, 2263-8, 1996

Lee et al., 2015

Lee, W. L., Chang, W. H., Wang, K. C., Guo, C. Y., Chou, Y. J., Huang, N., Huang, H. Y., Yen, M. S., Wang, P. H., The risk of epithelial ovarian cancer of women with endometriosis may be varied greatly if diagnostic criteria are different: A nationwide population-based cohort study, Medicine (United States), 94, e1633, 2015

Lei, 2010

Lei, L., Economic burden of surgery on endometriosis, Journal of Minimally Invasive Gynecology, 1), S101, 2010

ISBN 978-1-4731-2661-9

Levy et al., 2011

Levy, A. R., Osenenko, K. M., Lozano-Ortega, G., Sambrook, R., Jeddi, M., Belisle, S., Reid, R. L., Economic burden of surgically confirmed endometriosis in Canada, Journal of Obstetrics & Gynaecology Canada: JOGC, 33, 830-7, 2011

Lewin et al., 2015

Lewin, S., Glenton, C., Munthe-Kaas, H., Carlsen, B., Colvin, C. J., Gulmezoglu, M., Noyes, J., Booth, A., Garside, R., Rashidian, A., Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual), PLoS Medicine / Public Library of Science, 12, e1001895, 2015

Ling, 1999

Ling, F. W., Randomized controlled trial of depot leuprolide in patients with chronic pelvic pain and clinically suspected endometriosis. Pelvic Pain Study Group, Obstetrics & Gynecology, 93, 51-8, 1999

Loverro et al., 2001

Loverro, G., Santillo, V., Pansini, M., Lorusso, F., Depalo, R., Selvaggi, L., Are GnRH agonists helpful in the therapy of endometriosis after surgical treatment?, Human Reproduction (Oxford, England), 16, 96, 2001

Loverro et al., 2008

Loverro,G., Carriero,C., Rossi,A.C., Putignano,G., Nicolardi,V., Selvaggi,L., A randomized study comparing triptorelin or expectant management following conservative laparoscopic surgery for symptomatic stage III-IV endometriosis, European Journal of Obstetrics, Gynecology, and Reproductive Biology, 136, 194-198, 2008

Luciano et al., 1992

Luciano, A. A., Lowney, J., Jacobs, S. L., Endoscopic treatment of endometriosis-associated infertility. Therapeutic, economic and social benefits, Journal of Reproductive Medicine, 37, 573-6, 1992

Lukac et al., 2011

Lukac, M., Knight, C., Bielik, J., Tomek, D., Foltan, V., Kovac, A., et al., Cost-utility analysis of dienogest versus GNRH analogue in the treatment of endometriosis-associated pelvic pain in Slovakia, Value in health, 14, A109, 2011

Lukac et al., 2011a

Lukac, M., Knight, C., Bielik, J., Tomek, D., Petri, J. C., Bojnicky, M., Kovac, A., Endometriosis-associated pelvic pain treated with dienogest or GnRH analogues: Cost-utility comparison with 5 years time horizon, Value in health, 14, A403, 2011

Mantel et al., 1959

Mantel, N., Haenszel, W., Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 22, 719–748, 1959

ISBN 978-1-4731-2661-9

Marcoux et al., 1997

Marcoux, S., Maheux, R., Berube, S., Laparoscopic surgery in infertile women with minimal or mild endometriosis, New England Journal of Medicine, 337, 217-222, 1997

Markovic et al., 2008

Markovic, M., Manderson, L., Warren, N., Endurance and contest: women's narratives of endometriosis, Health: an Interdisciplinary Journal for the Social Study of Health, Illness & Medicine, 12, 349-67, 2008

Melin et al., 2006

Melin, A., Sparen, P., Persson, I., Bergqvist, A., Endometriosis and the risk of cancer with special emphasis on ovarian cancer, Human Reproduction, 21, 1237-1242, 2006

Melin et al., 2007

Melin, A., Sparen, P., Bergqvist, A., The risk of cancer and the role of parity among women with endometriosis, Human Reproduction, 22, 3021-6, 2007

Mettler et al., 2003

Mettler, L., Schollmeyer, T., Lehmann-Willenbrock, E., Schuppler, U., Schmutzler, A., Shukla, D., Zavala, A., Lewin, A., Accuracy of laparoscopic diagnosis of endometriosis, Journal of the Society of Laparoendoscopic Surgeons, 7, 15-8, 2003

Mettler et al., 2014

Mettler, L., Ruprai, R., Alkatout, I., Impact of medical and surgical treatment of endometriosis on the cure of endometriosis and pain, BioMed Research International, 2014, 264653, 2014

Mira et al., 2015

Mira, T. A., Giraldo, P. C., Yela, D. A., Benetti-Pinto, C. L., Effectiveness of complementary pain treatment for women with deep endometriosis through Transcutaneous Electrical Nerve Stimulation (TENS): randomized controlled trial, European Journal of Obstetrics, Gynecology, & Reproductive Biology, 194, 1-6, 2015

Mirkin et al., 2007

Mirkin, D., Murphy-Barron, C., Iwasaki, K., Actuarial analysis of private payer administrative claims data for women with endometriosis, Journal of Managed Care Pharmacy, 13, 262-72, 2007

Mogensen et al., 2016

Mogensen, J. B., Kjaer, S. K., Mellemkjaer, L., Jensen, A., Endometriosis and risks for ovarian, endometrial and breast cancers: A nationwide cohort study, Gynecologic Oncology, 143, 87-92, 2016

Moini et al., 2012

Moini, A., Bahar, L., Ashrafinia, M., Eslami, B., Hosseini, R., Ashrafinia, N., Fertility outcome after operative laparoscopy versus no treatment in infertile women with minimal or mild endometriosis, International Journal of Fertility & Sterility, 5, 235-40, 2012

Muzii et al., 2000

ISBN 978-1-4731-2661-9

Muzii, L., Marana, R., Caruana, P., Catalano, G. F., Margutti, F., Panici, P. B., Postoperative administration of monophasic combined oral contraceptives after laparoscopic treatment of ovarian endometriomas: a prospective, randomized trial, American Journal of Obstetrics & Gynecology, 183, 588-92, 2000

Namnoum et al., 1995

Namnoum, A. B., Hickman, T. N., Goodman, S. B., Gehlbach, D. L., Rock, J. A., Incidence of symptom recurrence after hysterectomy for endometriosis, Fertility and Sterility, 64, 898-902, 1995

Neal & McKenzie, 2011

Neal, D. M., McKenzie, P. J., Putting the pieces together: endometriosis blogs, cognitive authority, and collaborative information behavior, Journal of the Medical Library Association, 99, 127-34, 2011

Nisenblat et al., 2016

Nisenblat, Vicki, Bossuyt, M. M. Patrick, Shaikh, Rabia, Farquhar, Cindy, Jordan, Vanessa, Scheffers, Carola S., Mol, Willem Ben, Johnson, Neil, Hull, Louise M., Blood biomarkers for the non-invasive diagnosis of endometriosis, Cochrane Database of Systematic Reviews, 2016

Nisenblat et al., 2016a

Nisenblat, Vicki, Bossuyt, M. M. Patrick, Farquhar, Cindy, Johnson, Neil, Hull, Louise M., Imaging modalities for the non-invasive diagnosis of endometriosis, Cochrane Database of Systematic Reviews, 2016

Nisolle et al., 1990

Nisolle, M., Paindaveine, B., Bourdon, A., Berliere, M., Casanas-Roux, F., Donnez, J., Histologic study of peritoneal endometriosis in infertile women, Fertility & Sterility, 53, 984-8, 1990

Nnoaham et al., 2011

Nnoaham KE, Hummelshoj L, Webster P, d'Hooghe T, de Cicco Nardone F, de Cicco Nardone C, Jenkinson C, Kennedy SH, Zondervan KT., Impact of endometriosis on quality of life and work productivity: a multicenter study across ten countries, Fertility & Sterility, 96, 366-373

Novielli et al., 2010

Novielli, N., Cooper, N. J., Abrams, K. R., Sutton, A. J., How is evidence on test performance synthesized for economic decision models of diagnostic tests? A systematic appraisal of Health Technology Assessments in the UK since 1997, Value in health, 13, 952-7, 2010

Office of National Statistics (ONS), 2016

Office of National Statistics (ONS), Office of National Statistics (ONS), Families and households in the UK: 2016, 2016
https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2016

Oppelt et al., 2012

Oppelt, P., Chavtal, R., Haas, D., Reichert, B., Wagner, S., Muller, A., Lermann, J. H., Renner, S. P., Costs of in-patient treatment for endometriosis in Germany 2006: An analysis based on the G-DRG-Coding, Gynecological Endocrinology, 28, 903-905, 2012

ISBN 978-1-4731-2661-9

Overton et al., 1994

Overton, C. E., Lindsay, P. C., Johal, B., Collins, S. A., Siddle, N. C., Shaw, R. W., Barlow, D. H., A randomized, double-blind, placebo-controlled study of luteal phase dydrogesterone (Duphaston*) in women with minimal to mild endometriosis, Fertility and Sterility, 62, 701-707, 1994

Parali et al., 2012

Parali, E., Deger, C., Marmarali, B., Memis, S. A., Sumer, F., Karakeben, K., Kayadibinli, M., Effects of different clinical practices of laparoscopic surgery for endometriosis treatment on cost of therapy, Value in health, 15 (7), A407, 2012

Parazzini et al., 1994

Parazzini, F., Fedele, L., Busacca, M., Falsetti, L., Pellegrini, S., Venturini, P. L., Stella, M., Postsurgical medical treatment of advanced endometriosis: results of a randomized clinical trial, American Journal of Obstetrics and Gynecology, 171, 1205-1207, 1994

Parazzini et al., 2000

Parazzini, F., Di Cintio, E., Chatenoud, L., Moroni, S., Ardovino, I., Struzziero, E., Falsetti, L., Bianchi, A., Bracco, G., Pellegrini, A., Bertulessi, C., Romanini, C., Zupi, E., Massobrio, M., Guidetti, D., Troiano, L., Beretta, P., Franchi, M., Estroprogestin vs. gonadotrophin agonists plus estroprogestin in the treatment of endometriosis-related pelvic pain: a randomized trial. Gruppo Italiano per lo Studio dell'Endometriosi, European Journal of Obstetrics, Gynecology, & Reproductive Biology, 88, 11-4, 2000

Peterson et al., 2013

Peterson, C. M., Johnstone, E. B., Hammoud, A. O., Stanford, J. B., Varner, M. W., Kennedy, A., Chen, Z., Sun, L., Fujimoto, V. Y., Hediger, M. L., Buck Louis, G. M., Endo Study Working Group, Risk factors associated with endometriosis: importance of study population for characterizing disease in the ENDO Study, American Journal of Obstetrics & Gynecology, 208, 451.e1-11, 2013

Petta et al., 2005

Petta, C. A., Ferriani, R. A., Abrao, M. S., Hassan, D., Rosa, E. Silva J. C., Podgaec, S., Bahamondes, L., Randomized clinical trial of a levonorgestrel-releasing intrauterine system and a depot GnRH analogue for the treatment of chronic pelvic pain in women with endometriosis, Human Reproduction, 20, 1993-8, 2005

Piniazhko et al., 2014

Piniazhko, O., Zalis'ka, O., Vernikovskyy, I., The prevalence and cost of illness in women with endometriosis in Ukraine, Value in health, 17 (7), A507, 2014

Prast et al., 2013

Prast, J., Oppelt, P., Shamiyeh, A., Shebl, O., Brandes, I., Haas, D., Costs of endometriosis in Austria: a survey of direct and indirect costs, Archives of Gynecology & Obstetrics, 288, 569-76, 2013

Pritts et al., 1999

Pritts, T. A., Nussbaum, M. S., Flesch, L. V., Fegelman, E. J., Parikh, A. A., Fischer, J. E., Implementation of a clinical pathway decreases length of stay and cost for bowel resection, Annals of surgery, 230, 728-33, 1999

Querleu et al., 1993

ISBN 978-1-4731-2661-9

Querleu, D., Chapron, C., Chevallier, L., Bruhat, M. A., Complications of gynecologic laparoscopic surgery--a French multicenter collaborative study, New England Journal of Medicine, 328, 1355, 1993

Razzi et al., 2007

Razzi, S., Luisi, S., Ferretti, C., Calonaci, F., Gabbanini, M., Mazzini, M., Petraglia, F., Use of a progestogen only preparation containing desogestrel in the treatment of recurrent pelvic pain after conservative surgery for endometriosis, European Journal of Obstetrics, Gynecology, & Reproductive Biology, 135, 188-90, 2007

Redwine et al., 1996

Redwine, D. B., Koning, M., Sharpe, D. R., Laparoscopically assisted transvaginal segmental resection of the rectosigmoid colon for endometriosis, Fertility & Sterility, 65, 193-7, 1996

Reitsma et al., 2005

Reitsma, J. B., Glas, A. S., Rutjes, A. W., Scholten, R. J., Bossuyt, P. M., Zwinderman, A. H., Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews, Journal of clinical epidemiology, 58, 982-90, 2005

Rock et al., 1993

Rock, J. A., Truglia, J. A., Caplan, R. J., Zoladex (goserelin acetate implant) in the treatment of endometriosis: A randomized comparison with danazol, Obstetrics and gynecology, 82, 198-205, 1993

Rolland & Van Der Heijden, 1990

Rolland, R., Van Der Heijden, P. F., Nafarelin versus danazol in the treatment of endometriosis, American Journal of Obstetrics & Gynecology, 162, 586-8, 1990

RomeroMartin et al., 2012

Romero Martin, R. M., Alvis Nelson, A. N., Alfonso Camilo, A. C., Karpf Elizabeth, K. E., Upegui Angie, U. A., Cost minimization analysis of the dienogest use in patients with endometris in Colombia and Argentina, Value in health, 15 (4), A195, 2012

Rotondi et al., 2002

Rotondi, M., Labriola, D., Rotondi, M., Ammaturo, F. P., Amato, G., Carella, C., Izzo, A., Panariello, S., Depot leuprorelin acetate versus danazol in the treatment of infertile women with symptomatic endometriosis, European Journal of Gynaecological Oncology, 23, 523-526, 2002

Sanghera et al., 2016

Sanghera, S., Barton, P., Bhattacharya, S., Horne, A. W., Roberts, T. E., Pharmaceutical treatments to prevent recurrence of endometriosis following surgery: A model-based economic evaluation, BMJ Open, 6 (4) (no pagination), 2016

Sayasneh et al., 2015

Sayasneh, A., Kaijser, J., Preisler, J., Smith, A. A., Raslan, F., Johnson, S., Husicka, R., Ferrara, L., Stalder, C., Ghaem-Maghami, S., Timmerman, D., Bourne, T., Accuracy of ultrasonography performed by examiners with varied training and experience in predicting specific pathology of adnexal masses, Ultrasound in Obstetrics & Gynecology, 45, 605-12, 2015

Schlaff et al., 2006

ISBN 978-1-4731-2661-9

Schlaff, W. D., Carson, S. A., Luciano, A., Ross, D., Bergqvist, A., Subcutaneous injection of depot medroxyprogesterone acetate compared with leuprolide acetate in the treatment of endometriosis-associated pain, Fertility & Sterility, 85, 314-25, 2006

Schwartz et al., 1994

Schwartz, L. B., Panageas, E., Lange, R., Rizzo, J., Comite, F., McCarthy, S., Female pelvis: Impact of MR imaging on treatment decisions and net cost analysis, Radiology, 192, 55-60, 1994

Seear, 2009

Seear, K, The third shift: Health, work and expertise among women with endometriosis, Health Sociology Review, 18, 194-206, 2009

Seibel et al., 1982

Seibel, M. M., Berger, M. J., Weinstein, F. G., Taymor, M. L., The effectiveness of danazol on subsequent fertility in minimal endometriosis, Fertility & Sterility, 38, 534-7, 1982

Seracchioli et al., 2010

Seracchioli, R., Mabrouk, M., Frasca, C., Manuzzi, L., Montanari, G., Keramyda, A., Venturoli, S., Long-term cyclic and continuous oral contraceptive therapy and endometrioma recurrence: a randomized controlled trial, Fertility & Sterility, 93, 52-6, 2010

Seracchioli et al., 2010a

Seracchioli, R., Mabrouk, M., Frasca, C., Manuzzi, L., Savelli, L., Venturoli, S., Long-term oral contraceptive pills and postoperative pain management after laparoscopic excision of ovarian endometrioma: a randomized controlled trial, Fertility & Sterility, 94, 464-71, 2010

Sesti et al., 2007

Sesti, F., Pietropolli, A., Capozzolo, T., Broccoli, P., Pierangeli, S., Bollea, M. R., Piccione, E., Hormonal suppression treatment or dietary therapy versus placebo in the control of painful symptoms after conservative surgery for endometriosis stage III-IV. A randomized comparative trial, Fertility & Sterility, 88, 1541-7, 2007

Sesti et al., 2009

Sesti, F., Capozzolo, T., Pietropolli, A., Marziali, M., Bollea, M. R., Piccione, E., Recurrence rate of endometrioma after laparoscopic cystectomy: a comparative randomized trial between post-operative hormonal suppression treatment or dietary therapy vs. placebo, European Journal of Obstetrics, Gynecology, & Reproductive Biology, 147, 72-7, 2009

Shafik et al., 2000

Shafik, A., Ratcliffe, N., Wright, J. T., The importance of histological diagnosis in patients with chronic pelvic pain and laparoscopic evidence of endometriosis, Gynaecological Endoscopy, 9, 301-304, 2000

Shah et al., 2016

Shah, A., Raval, A., Thakker, D., National trends in direct health care costs associated with endometriosis among females in the United States: 2004-2013, Value in Health, 19 (3), A30-A31, 2016

Shakeri et al., 2016

ISBN 978-1-4731-2661-9

Shakeri, B., Nadim, B., Reid, S., Condous, G., Mongelli, M., EP27.03: Cost-benefit analysis for the utilisation of detailed preoperative ultrasound in women with complex endometriosis disease, Ultrasound in Obstetrics & Gynecology 48 Suppl 1, 378, 2016

Shakeri, 2016

Shakeri, B., Cost-benefit analysis for the utilisation of detailed preoperative ultrasound in women with complex endometriosis disease, Journal of minimally invasive gynecology, 23 (7 Supplement 1), S183, 2016

Shaw, 1990

Shaw, R. W., Nafarelin in the treatment of pelvic pain caused by endometriosis, American Journal of Obstetrics and Gynecology, 162, 574-576, 1990

Shaw, 1992a

Shaw, R. W., An open randomized comparative study of the effect of goserelin depot and danazol in the treatment of endometriosis. Zoladex Endometriosis Study Team, Fertility and Sterility, 58, 265-272, 1992

Shoebotham & Coulson, 2016

Shoebotham, A., Coulson, N. S., Therapeutic affordances of online support group use in women with endometriosis, Journal of Medical Internet Research, 18, e109, 2016

Shokeir & Mousa, 2015

Shokeir, T., Mousa, S., A randomized, placebo-controlled, double-blind study of hysteroscopic-guided pertubal diluted bupivacaine infusion for endometriosis-associated chronic pelvic pain, International Journal of Gynaecology & Obstetrics, 130, 219-22, 2015

Siddle et al., 1987

Siddle, N., Sarrel, P., Whitehead, M., The effect of hysterectomy on the age at ovarian failure: identification of a subgroup of women with premature loss of ovarian function and literature review, Fertility & Sterility, 47, 94-100, 1987

Simoens et al., 2011

Simoens, S., Hummelshoj, L., Dunselman, G., Dirksen, C., Endocost Consortium, W., D'Hooghe, T., The burden of endometriosis: Costs and quality of life of women with endometriosis treated in referral centres, Value in health, 14 (7), A401, 2011

Simoens et al., 2014

Simoens, S., Dunselman, G., Dirksen, C., Hummelshoj, L., Bokor, A., Brandes, I., Brodszky, V., Canis, M., Colombo, G. L., DeLeire, T., Falcone, T., Graham, B., Halis, G., Horne, A., Kanj, O., Kjer, J. J., Kristensen, J., Lebovic, D., Mueller, M., Vigano, P., Wullschleger, M., D'Hooghe, T., Erratum: The burden of endometriosis: Costs and quality of life of women with endometriosis and treated in referral centres [Human Reproduction (2012) 27 (1292-1299)], Human Reproduction, 29, 1292, 2014

Slack et al., 2007

Slack, A., Child, T., Lindsey, I., Kennedy, S., Cunningham, C., Mortensen, N., Koninckx, P., McVeigh, E., Urological and colorectal complications following surgery for rectovaginal endometriosis, BJOG: An International Journal of Obstetrics and Gynaecology, 114, 1278-1282, 2007

Soliman et al., 2014

ISBN 978-1-4731-2661-9

Soliman, A. M., Fuldeore, M. J., Yang, H., Du, E. X., Wu, E. Q., Winkel, C. A., Significant economic burden among endometriosis patients before and after diagnosis: A retrospective claims database analysis, Endocrine Reviews. Conference: 96th Annual Meeting and Expo of the Endocrine Society, ENDO, 35, 2014

Stevens, 2011

Stevens, J. W., A note on dealing with missing standard errors in meta-analyses of continuous outcome measures in WinBUGS, Pharmaceutical statistics, 10, 374-8, 2011

Stewart et al., 2013

Stewart, L. M., Holman, C. D. J., Aboagye-Sarfo, P., Finn, J. C., Preen, D. B., Hart, R., In vitro fertilization, endometriosis, nulliparity and ovarian cancer risk, Gynecologic Oncology, 128, 260-264, 2013

Stratton et al., 2002

Stratton, P., Winkel, C. A., Sinaii, N., Merino, M. J., Zimmer, C., Nieman, L. K., Location, color, size, depth, and volume may predict endometriosis in lesions resected at surgery, Fertility & Sterility, 78, 743-9, 2002

Stripling et al., 1988

Stripling, M. C., Martin, D. C., Chatman, D. L., Zwaag, R. V., Poston, W. M., Subtle appearance of pelvic endometriosis, Fertility & Sterility, 49, 427-31, 1988

Strowitzki et al., 2010

Strowitzki, T., Marr, J., Gerlinger, C., Faustmann, T., Seitz, C., Dienogest is as effective as leuprolide acetate in treating the painful symptoms of endometriosis: a 24-week, randomized, multicentre, open-label trial, Human Reproduction, 25, 633-41, 2010

Strowitzki et al., 2010a

Strowitzki, T., Faustmann, T., Gerlinger, C., Seitz, C., Dienogest in the treatment of endometriosis-associated pelvic pain: a 12-week, randomized, double-blind, placebo-controlled study, European Journal of Obstetrics, Gynecology, & Reproductive Biology, 151, 193-8, 2010

StrzempkoButt & Chesla, 2007

Strzempko Butt, F., Chesla, C., Relational patterns of couples living with chronic pelvic pain from endometriosis, Qualitative Health Research, 17, 571-85, 2007

Sugimura et al., 1996

Sugimura, K., Imaoka, I., Okizuka, H., Pelvic endometriosis: impact of magnetic resonance imaging on treatment decisions and costs, Academic Radiology, 3 Suppl 1, S66-8, 1996

Sutton et al., 1994

Sutton, C. J., Ewen, S. P., Whitelaw, N., Haines, P., Prospective, randomized, double-blind, controlled trial of laser laparoscopy in the treatment of pelvic pain associated with minimal, mild, and moderate endometriosis, Fertility & Sterility, 62, 696-700, 1994

Taipale et al., 2009

Taipale, K., Leminen, A., Rasanen, P., Heikkila, A., Tapper, A. M., Sintonen, H., Roine, R. P., Costs and health-related quality of life effects of hysterectomy in patients with benign uterine disorders, Acta Obstetricia et Gynecologica Scandinavica, 88, 1402-1410, 2009

ISBN 978-1-4731-2661-9

Telimaa et al., 1987a

Telimaa, S., Puolakka, J., Ronnberg, L., Kauppila, A., Placebo-controlled comparison of danazol and high-dose medroxyprogesterone acetate in the treatment of endometriosis, Gynecological Endocrinology, 1, 13-23, 1987

Thomas & Cooke, 1987a

Thomas, E. J., Cooke, I. D., Successful treatment of asymptomatic endometriosis: does it benefit infertile women? British Medical Journal Clinical Research Ed., 294, 1117-9, 1987

Treloar et al., 2007

Treloar, S. A., Morley, K. I., Taylor, S. D., Hall, W. D., Why do they do it? A pilot study towards understanding participant motivation and experience in a large genetic epidemiological study of endometriosis, Community Genetics, 10, 61-71, 2007

Tuletova et al., 2014

Tuletova, A., Dochshanova, A., Skakova, R., Cost-effectiveness analysis of the therapy of endometriosis, Value in health, 17 (7), A511, 2014

Turner et al., 2012

Turner, R. M., Davey, J., Clarke, M. J., Thompson, S. G., Higgins, J. P., Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews, International journal of epidemiology, 41, 818-27, 2012

Van Der Valk et al., 2015

Van Der Valk, M. E., Mangen, M. J., Severs, M., Van Der Have, M., Dijkstra, G., Van Bodegraven, A. A., Fidder, H. H., De Jong, D. J., Pierik, M., Van Der Woude, C. J., Romberg-Camps, M. J., Clemens, C. H., Jansen, J. M., Van De Meeberg, P. C., Mahmmod, N., Van Der Meulen-de Jong, A. E., Ponsioen, C. Y., Bolwerk, C., Vermeijden, J. R., Siersema, P. D., Leenders, M., Oldenburg, B., Coin study group, Dutch Initiative on, Crohn, Colitis, Comparison of Costs and Quality of Life in Ulcerative Colitis Patients with an Ileal Pouch-Anal Anastomosis, Ileostomy and Anti-TNFalpha Therapy, Journal of Crohn's & Colitis, 9, 1016-23, 2015

Van Houwelingen et al., 1993

Van Houwelingen, H. C., Zwinderman, K. H., Stijnen, T., A bivariate approach to meta-analysis, Statistics in medicine, 12, 2273-84, 1993

Van Houwelingen et al., 2002

Van Houwelingen, H. C., Arends, L. R., Stijnen, T., Advanced methods in meta-analysis: multivariate approach and meta-regression, Statistics in medicine, 21, 589-624, 2002

Van Valkenhoef et al., 2016

Van Valkenhoef, G., Dias, S., Ades, A.E., Welton, N.J., Automated generation of node-splitting models for assessment of inconsistency in network meta-analysis, Research Synthesis Methods, 7, 80-93, 2016

Vercellini et al., 1991

Vercellini, P., Vendola, N., Bocciolone, L., Rognoni, M. T., Carinelli, S. G., Candiani, G. B., Reliability of the visual diagnosis of ovarian endometriosis, Fertility & Sterility, 56, 1198-200, 1991

Vercellini et al., 1996

ISBN 978-1-4731-2661-9

Vercellini, P., Soma, M., Moro, G. L., Gestrinone versus a gonadotropin-releasing hormone agonist for the treatment of pelvic pain associated with endometriosis: A multicenter, randomized, double-blind study, Fertility and sterility, 66, 911-919, 1996

Vercellini et al., 1996a

Vercellini, P., De Giorgi, O., Oldani, S., Cortesi, I., Panazza, S., Crosignani, P. G., Depot medroxyprogesterone acetate versus an oral contraceptive combined with very-low-dose danazol for long-term treatment of pelvic pain associated with endometriosis, American Journal of Obstetrics & Gynecology, 175, 396-401, 1996

Walch et al., 2009

Walch, K., Unfried, G., Huber, J., Kurz, C., Van Trotsenburg, M., Pernicka, E., Wenzl, R., Implanon versus medroxyprogesterone acetate: effects on pain scores in patients with symptomatic endometriosis--a pilot study, Contraception, 79, 29-34, 2009

Walter et al., 2001

Walter, A. J., Hentz, J. G., Magtibay, P. M., Cornella, J. L., Magrina, J. F., Endometriosis: correlation between histologic and visual findings at laparoscopy, American Journal of Obstetrics & Gynecology, 184, 1407-11; discussion 1411-3, 2001

Wan et al., 2014

Wan, X., Wang, W., Liu, J., Tong, T., Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range, BMC medical research methodology, 14, 135, 2014

Wang et al., 2014

Wang, K. C., Chang, W. H., Lee, W. L., Huang, N., Huang, H. Y., Yen, M. S., Guo, C. Y., Wang, P. H., An increased risk of epithelial ovarian cancer in Taiwanese women with a new surgico-pathological diagnosis of endometriosis, BMC Cancer, 14, 831, 2014

Wasiak et al., 2010

Wasiak, R., Manson, S., Ryan, J., Petri, J. C., Treatment and cost for patients with endometriosis in the United Kingdom, Value in health, 13 (7), A376, 2010

Wasiak et al., 2012

Wasiak, R., Hjalte, F., Ragnarson Tennvall, G., Olovsson, M., Cost of illness for endometriosis in Sweden, International Journal of Gynecology and Obstetrics, 119, S553, 2012

Wayne et al., 2008

Wayne, P. M., Kerr, C. E., Schnyer, R. N., Legedza, A. T. R., Savetsky-German, J., Shields, M. H., Buring, J. E., Davis, R. B., Conboy, L. A., Highfield, E., Parton, B., Thomas, P., Laufer, M. R., Japanese-style acupuncture for endometriosis-related pelvic pain in adolescents and young women: results of a randomized sham-controlled trial, Journal of Pediatric and Adolescent Gynecology, 21, 247-257, 2008

Weng et al., 2015

Weng, Q., Ding, Z. M., Lv, X. L., Yang, D. X., Song, Y. Z., Wang, F. F., Ye, Y. H., Qu, F., Chinese medicinal plants for advanced endometriosis after conservative surgery: a prospective, multi-center and controlled trial, International journal of clinical and experimental medicine, 8, 11307-11, 2015

ISBN 978-1-4731-2661-9

Wheeler et al., 1993

Wheeler, J. M., Knittle, J. D., Miller, J. D., Depot leuprolide acetate versus danazol in the treatment of women with symptomatic endometriosis: a multicenter, double-blind randomized clinical trial. II. Assessment of safety. The Lupron Endometriosis Study Group, American Journal of Obstetrics & Gynecology, 169, 26-33, 1993

Whelan, 2007

Whelan, E., 'No one agrees except for those of us who have it': endometriosis patients as an epistemological community, Sociology of Health & Illness, 29, 957-82, 2007

Whitehill et al., 2012

Whitehill, K., Yong, P. J., Williams, C., Clinical predictors of endometriosis in the infertility population: is there a better way to determine who needs a laparoscopy?, Journal of Obstetrics & Gynaecology Canada, 34, 552-7, 2012

Wickstrm et al., 2012

Wickstrm, K., Bruse, C., Sjsten, A., Spira, J., Edelstam, G., Pertubation with lignocaine as a new treatment of dysmenorrhoea due to endometriosis: A randomized controlled trial, Human Reproduction, 27, 695-701, 2012

Wickstrom & Edelstam, 2012

Wickstrom, K., Edelstam, G., Pertubation with lignocaine as a new treatment of dysmenorrhoea due to endometriosis: A randomised controlled trial, International Journal of Gynecology and Obstetrics, 119, S548-S549, 2012

Wickstrom et al., 2013

Wickstrom, K., Bruse, C., Sjosten, A., Spira, J., Edelstam, G., Quality of life in patients with endometriosis and the effect of pertubation with lidocaine - a randomized controlled trial, Acta Obstetricia et Gynecologica Scandinavica, 92, 1375-82, 2013

Wickstrom et al., 2013

Wickstrom, K., Bruse, C., Sjosten, A., Spira, J., Edelstam, G., Pertubation with lignocaine as a new treatment of dysmenorrhoea due to endometriosis: A randomized controlled trial, Obstetrical & Gynecological Survey, 68, 286-7, 2013

Wong & Tang, 2004

Wong, A. Y., Tang, L., An open and randomized study comparing the efficacy of standard danazol and modified triptorelin regimens for postoperative disease management of moderate to severe endometriosis, Fertility & Sterility, 81, 1522-7, 2004

Wong et al., 2010

Wong, A. Y., Tang, L. C., Chin, R. K., Levonorgestrel-releasing intrauterine system (Mirena) and Depot medroxyprogesterone acetate (Depoprovera) as long-term maintenance therapy for patients with moderate and severe endometriosis: a randomised controlled trial, Australian & New Zealand Journal of Obstetrics & Gynaecology, 50, 273-9, 2010

Wright et al., 2005

Wright, J., Lotfallah, H., Jones, K., Lovell, D., A randomized trial of excision versus ablation for mild endometriosis, Fertility & Sterility, 83, 1830-6, 2005

Xia, 2006

ISBN 978-1-4731-2661-9

Xia, T, Effect of acupuncture and traditional Chinese herbal medicine in treating endometriosis, International Journal of Clinical Acupuncture, 15, 145-50., 2006

Xiang et al., 2002

Xiang, D., Situ, Y., Liang, X., Cheng, L., Zhang, G., Ear acupuncture therapy for 37 cases of dysmenorrhoea due to endometriosis, Journal of Traditional Chinese Medicine, 22, 282-5, 2002

Yu et al., 2015

Yu, H. C., Lin, C. Y., Chang, W. C., Shen, B. J., Chang, W. P., Chuang, C. M., Increased association between endometriosis and endometrial cancer: A nationwide population-based retrospective cohort study, International Journal of Gynecological Cancer, 25, 447-452, 2015

Zalis'ka et al., 2014

Zalis'ka, O., Piniazhko, O., Vernikovskyy, I., Gnatyshak, L., Analysis of treatment cost for endometriosis in Ukraine, Value in health, 17 (3), A158, 2014

Zhang et al., 2014

Zhang, Y., Qiao, C., Li, L., Zhao, X., Li, Y., Serum HE4 is more suitable as a biomarker than CA125 in Chinese women with benign gynecologic disorders, African Health Sciences, 14, 913-8, 2014

Zhao et al., 1998

Zhao, S. Z., Arguelles, L. M., Wong, J. M., Davis, M. B., Gersh, G. E., Struthers, B. J., Cost comparisons between nafarelin and leuprolide in the treatment of endometriosis, Clinical Therapeutics, 20, 592-602, 1998

Zhao et al., 1998a

Zhao, S. Z., Wong, J. M., Davis, M. B., Gersh, G. E., Johnson, K. E., The cost of inpatient endometriosis treatment: An analysis based on the healthcare cost and utilization project nationwide inpatient sample, American Journal of Managed Care, 4, 1127-1134, 1998

Zhu et al., 2014

Zhu, S., Liu, D., Huang, W., Wang, Q., Wang, Q., Zhou, L., Feng, G., Post-laparoscopic oral contraceptive combined with Chinese herbal mixture in treatment of infertility and pain associated with minimal or mild endometriosis: a randomized controlled trial, BMC Complementary & Alternative Medicine, 14, 222, 2014

Zondervan et al., 2001

Zondervan, K. T., Yudkin, P. L., Vessey, M. P., Jenkinson, C. P., Dawes, M. G., Barlow, D. H., Kennedy, S. H., Chronic pelvic pain in the community--symptoms, investigations, and diagnoses, 184, 1149-55, 2001

ISBN 978-1-4731-2661-9

14 Glossary and abbreviations

Term	Definition
Abstract	Summary of a study, which may be published alone or as an introduction to a full scientific paper.
Arm (of a clinical study)	Subsection of individuals within a study who receive one particular intervention, for example, placebo arm.
Association	Statistical relationship between 2 or more events, characteristics or other variables. The relationship may or may not be causal.
Attrition bias	Systematic differences between comparison groups for withdrawal or exclusion of participants from a study.
Available case analysis (ACA)	Analysis of data that is available for participants at the end of follow-up.
Baseline	The initial set of measurements at the beginning of a study (after run-in period where applicable) with which subsequent results are compared.
Before-and-after study	A study that investigates the effects of an intervention by measuring particular characteristics of a population both before and after taking the intervention, and assessing any change that occurs.
Bias	Influences on a study that can make the results look better or worse than they really are. Bias can occur by chance, deliberately or as a result of systematic errors in the design and execution of a study. It can also occur at different stages in the research process, for example, during the collection, analysis, interpretation, publication or review of research data. For examples see Confounding factor, Performance bias, Publication bias Selection bias.
Carer (caregiver)	Someone who looks after family, partners or friends in need of help because they are ill, frail or have a disability.
Case-control study	A study to find out the cause(s) of a disease or condition. This is done by comparing a group of patients who have the disease or condition (cases) with a group of people who do not have it (controls) but who are otherwise as similar as possible (in characteristics thought to be unrelated to the causes of the disease or condition). This means the researcher can look for aspects of their lives that differ to see if they may cause the condition. Such studies are retrospective because they look back in time from the outcome to the possible causes of a disease or condition.
Case series	Report of a number of cases of a given disease, usually covering the course of the disease and the response to treatment. There is no comparison (control) group of patients.
Chronic pelvic pain	Defined as pelvic pain lasting for 6 months or longer.
Clinical audit	A systematic process for setting and monitoring standards of clinical care. Whereas ‘guidelines’ define what the best clinical practice should be, ‘audit’ investigates whether best practice is being carried out. Clinical audit can be described as a cycle or spiral. Within the cycle there are stages that follow a systematic process of establishing best practice, measuring care against specific criteria, taking action to improve care and monitoring to sustain improvement. The spiral suggests that as the process continues, each cycle aspires to a higher level of quality.
Clinical effectiveness	How well a specific test or treatment works when used in the 'real world' (for example, when used by a doctor with a patient at home), rather than in a carefully controlled clinical trial. Trials that assess clinical effectiveness are sometimes called management trials. Clinical effectiveness is not the same as efficacy.
Clinical efficacy	The extent to which an intervention is active when studied under controlled research conditions.

ISBN 978-1-4731-2661-9

Term	Definition
Clinician	A healthcare professional who provides patient care. For example, a doctor, nurse or physiotherapist.
Cochrane Review	The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of RCTs prepared by the Cochrane Collaboration).
Cohort study	A study with 2 or more groups of people – cohorts – with similar characteristics. One group receives a treatment, is exposed to a risk factor or has a particular symptom and the other group does not. The study follows their progress over time and records what happens.
Comorbidity	A disease or condition that someone has in addition to the health problem being studied or treated.
Community services	Community services include: GPs, sexual health services, practice nurses and school nurses.
Concealment of allocation	The process used to ensure that the person deciding to enter a participant into an RCT does not know the comparison group into which that individual will be allocated. This is distinct from blinding and is aimed at preventing selection bias. Some attempts at concealing allocation are more prone to manipulation than others and the method of allocation concealment is used as an assessment of the quality of a trial.
Confidence interval (CI)	There is always some uncertainty in research. This is because a small group of patients is studied to predict the effects of a treatment on the wider population. The confidence interval is a way of expressing how certain we are about the findings from a study, using statistics. It gives a range of results that is likely to include the 'true' value for the population. The CI is usually stated as '95% CI', which means that the range of values has a 95 in 100 chance of including the 'true' value. For example, a study may state that “based on our sample findings, we are 95% certain that the 'true' population blood pressure is not higher than 150 and not lower than 110”. In such a case the 95% CI would be 110 to 150. A wide confidence interval indicates a lack of certainty about the true effect of the test or treatment – often because a small group of patients has been studied. A narrow confidence interval indicates a more precise estimate (for example, if a large number of patients have been studied).
Confounding factor	Something that influences a study and can result in misleading findings if it is not understood or appropriately dealt with. For example, a study of heart disease may look at a group of people who exercise regularly and a group who do not exercise. If the ages of the people in the 2 groups are different, then any difference in heart disease rates between the 2 groups could be because of age rather than exercise. Therefore age is a confounding factor.
Continuous outcome	Data with a potentially infinite number of possible values within a given range. Height, weight and blood pressure are examples of continuous variables.
Control group	A group of people in a study who do not receive the treatment or test being studied. Instead, they may receive the standard treatment (sometimes called 'usual care') or a dummy treatment (placebo). The results for the control group are compared with those for a group receiving the treatment being tested. The aim is to check for any differences. Ideally, the people in the control group should be as similar as possible to those in the treatment group, to make it as easy as possible to detect any effects due to the treatment.
Cost–benefit analysis (CBA)	Cost-benefit analysis is one of the tools used to carry out an economic evaluation. The costs and benefits are measured using the same

ISBN 978-1-4731-2661-9

Term	Definition
	monetary units (for example, UK pounds) to see whether the benefits exceed the costs.
Cost–consequence analysis (CCA)	Cost-consequence analysis is one of the tools used to carry out an economic evaluation. This compares the costs (such as treatment and hospital care) with the consequences (such as health outcomes) of a test or treatment with a suitable alternative. Unlike cost–benefit analysis or cost-effectiveness analysis, it does not attempt to summarise outcomes in a single measure (such as the quality adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out.
Cost-effectiveness analysis (CEA)	Cost-effectiveness analysis is one of the tools used to carry out an economic evaluation. The benefits are expressed in non-monetary terms related to health, such as symptom-free days, heart attacks avoided, deaths avoided or life years gained (that is, the number of years by which life is extended as a result of the intervention).
Cost-effectiveness model	An explicit mathematical framework which is used to represent clinical decision problems and incorporate evidence from a variety of sources in order to estimate the costs and health outcomes.
Cost–utility analysis (CUA)	Cost–utility analysis is one of the tools used to carry out an economic evaluation. The benefits are assessed in terms of both quality and duration of life, and expressed as quality adjusted life years (QALYs). See also Utility.
COX proportional hazard model	In survival analysis, a statistical model that asserts that the effect of the study factors (for example, the intervention of interest) on the hazard rate (the risk of occurrence of an event) in the study population is multiplicative and does not change over time.
Credible interval (CrI)	The Bayesian equivalent of a confidence interval.
Cyst wall	The outer or capsular portion of a cyst.
Decision analysis	An explicit quantitative approach to decision-making under uncertainty, based on evidence from research. This evidence is translated into probabilities, and then into diagrams or decision trees which direct the clinician through a succession of possible scenarios, actions and outcomes.
Deep Infiltrating Endometriosis (DIE)	The nodules implant at least 5mm below the peritoneum (the lining of the pelvis). Structures penetrated can include the uterosacral ligaments (ligaments supporting the womb), bowel, bladder and ureter.
Dichotomous outcomes	Outcome that can take one of 2 possible values, such as dead/alive, smoker/non-smoker, present/not present (also called binary data).
Discounting	Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be experienced in the future rather than the present.
Dominance	A health economics term. When comparing tests or treatments, an option that is both less effective and costs more is said to be 'dominated' by the alternative.
Drop-out	A participant who withdraws from a trial before the end.
Economic evaluation	An economic evaluation is used to assess the cost effectiveness of healthcare interventions (that is, to compare the costs and benefits of a healthcare intervention to assess whether it is worth doing). The aim of an economic evaluation is to maximise the level of benefits – health effects – relative to the resources available. It should be used to inform

ISBN 978-1-4731-2661-9

Term	Definition
	and support the decision-making process; it is not supposed to replace the judgement of healthcare professionals. There are several types of economic evaluation: cost–benefit analysis, cost–consequence analysis, cost-effectiveness analysis, cost-minimisation analysis and cost–utility analysis. They use similar methods to define and evaluate costs, but differ in the way they estimate the benefits of a particular drug, programme or intervention.
Effect (as in effect measure, treatment effect, estimate of effect, effect size)	A measure that shows the magnitude of the outcome in 1 group compared with that in a control group. For example, if the absolute risk reduction is shown to be 5% and it is the outcome of interest, the effect size is 5%. The effect size is usually tested, using statistics, to find out how likely it is that the effect is a result of the treatment and has not just happened by chance.
Effectiveness	How beneficial a test or treatment is under usual or everyday conditions?
Efficacy	How beneficial a test, treatment or public health intervention is under ideal conditions (for example, in a laboratory).
Epidemiological study	The study of a disease within a population, defining its incidence and prevalence and examining the roles of external influences (for example, infection, diet) and interventions.
EQ-5D (EuroQol 5 dimensions)	A standardised instrument used to measure health-related quality of life. It provides a single index value for health status.
Equivalence study	A trial designed to determine whether the response to 2 or more treatments differs by an amount that is clinically unimportant. This is usually demonstrated by showing that the true treatment difference is likely to lie between a lower and an upper equivalence level of clinically acceptable differences.
Evidence	Information on which a decision or guidance is based. Evidence is obtained from a range of sources including RCTs, observational studies, expert opinion (of clinical professionals or patients).
Exclusion criteria (clinical study)	Criteria that define who is not eligible to participate in a clinical study.
Exclusion criteria (literature review)	Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence.
Extended dominance	If Option A is both more clinically effective than Option B and has a lower cost per unit of effect when both are compared with a do-nothing alternative, then Option A is said to have extended dominance over Option B. Option A is therefore more cost-effective and should be preferred, other things remaining equal.
Extrapolation	An assumption that the results of studies of a specific population will also hold true for another population with similar characteristics.
False negative	A diagnostic test result that incorrectly indicates that an individual does not have the disease of interest, when they do actually have it.
False positive	A diagnostic test result that incorrectly indicates that an individual has the disease of interest, when they actually do not have it.
Fixed-effect model	In meta-analysis, a model that calculates a pooled effect estimate using the assumption that all observed variation between studies is caused by random sample variability. Studies are assumed to estimating the same overall effect.
Follow-up	Observation over a period of time of an individual, group or initially defined population whose appropriate characteristics have been assessed in order to observe changes in health status or health-related variables.
Forest plot	A graphical representation of the individual results of each study included in a meta-analysis together with the combined meta-analysis

ISBN 978-1-4731-2661-9

Term	Definition
	result. The plot also allows readers to see the heterogeneity among the results of the studies. The results of individual studies are shown as squares centred on each study's point estimate. A horizontal line runs through each square to show each study's confidence interval. The overall estimate from the meta-analysis and its confidence interval are shown at the bottom, represented as a diamond. The centre of the diamond represents the pooled point estimate, and its horizontal tips represent the confidence interval.
Generalisability	The extent to which the results of a study hold true for groups that did not participate in the research.
Gold standard	A method, procedure or measurement that is widely accepted as being the best available to test for or treat a disease.
GRADE, GRADE profile	A system developed by the GRADE Working Group to address the short-comings of present grading systems in healthcare. The GRADE system uses a common, sensible and transparent approach to grading the quality of evidence. The results of applying the GRADE system to clinical trial data are displayed in a table known as a GRADE profile.
Harms	Adverse effects of an intervention.
Hazard ratio	A hazard is the rate at which events happen, so that the probability of an event happening in a short time interval is the length of time multiplied by the hazard. Although the hazard may vary with time, the assumption in proportional hazard models for survival analysis is that the hazard in one group is a constant proportion of the hazard in the other group. This proportion is the hazard ratio.
Health economics	Study or analysis of the cost of using and distributing healthcare resources.
Health-related quality of life (HRQoL)	A measure of the effects of an illness to see how it affects someone's day-to-day life.
Heterogeneity	The term is used in meta-analyses and systematic reviews to describe when the results of a test or treatment (or estimates of its effect) differ
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of effect.
Incidence	The incidence of a disease is the rate at which new cases occur in a population during a specified period.
Inclusion criteria (clinical study)	Specific criteria that define who is eligible to participate in a clinical study.
Inclusion criteria (literature review)	Explicit criteria used to decide which studies should be considered as potential sources of evidence.
Incremental cost	The extra cost linked to using one test or treatment rather than another. Or the additional cost of doing a test or providing a treatment more frequently.
Incremental cost-effectiveness ratio (ICER)	The difference in the mean costs in the population of interest divided by the differences in the mean outcomes in the population of interest for one treatment compared with another.
Incremental net benefit (INB)	The value (usually in monetary terms) of an intervention net of its cost compared with a comparator intervention. The INB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the INB is calculated as: (£20,000×QALYs gained) minus incremental cost.
Indirectness	The available evidence is different to the review question being addressed, in terms of population, intervention, comparison and outcome (PICO).
Intention-to-treat analysis (ITT)	An assessment of the people taking part in a clinical trial, based on the group they were initially (and randomly) allocated to. This is regardless

ISBN 978-1-4731-2661-9

Term	Definition
	of whether or not they dropped out, fully complied with the treatment or switched to an alternative treatment. Intention-to-treat analyses are often used to assess clinical effectiveness because they mirror actual practice: that is, not everyone complies with treatment and the treatment people receive may be changed according to how they respond to it.
Intervention	In medical terms this could be a drug treatment, surgical procedure, diagnostic or psychological therapy. Examples of public health interventions could include action to help someone to be physically active or to eat a more healthy diet.
Kappa statistic	A statistical measure of inter-rater agreement that takes into account the agreement occurring by chance
Length of stay	The total number of days a patient stays in hospital.
Licence	See Product licence.
Life years gained	Mean average years of life gained per person as a result of the intervention compared with an alternative intervention.
Likelihood ratio	The likelihood ratio combines information about the sensitivity and specificity. It tells you how much a positive or negative result changes the likelihood that a patient would have the disease. The likelihood ratio of a positive test result (LR+) is sensitivity divided by (1 minus specificity).
LNG-IUS	Levonorgestrel-releasing intrauterine system is a contraceptive device fitted in the uterus that releases a form of progestogen
Loss to follow-up	Patients who have withdrawn from the clinical trial at the point of follow-up.
Managed clinical networks	Linked groups of healthcare professionals from primary, secondary and tertiary care providing a coordinated patient pathway. Responsibility for setting up these networks will depend on existing service provision and location.
Markov model	A method for estimating long-term costs and effects for recurrent or chronic conditions, based on health states and the probability of transition between them within a given time period (cycle).
Mean	An average value, calculated by adding all the observations and dividing by the number of observations.
Mean difference	In meta-analysis, a method used to combine measures on continuous scales (such as weight), where the mean, standard deviation and sample size in each group are known. The weight given to the difference in means from each study (for example, how much influence each study has on the overall results of the meta-analysis) is determined by the precision of its estimate of effect.
Median	The value of the observation that comes half-way when the observations are ranked in order.
Meta-analysis	A method often used in systematic reviews. Results from several studies of the same test or treatment are combined to estimate the overall effect of the treatment.
Minimal important difference (MID)	Threshold for clinical importance which represents the minimal important difference for benefit or for harm; for example, the threshold at which drug A is less effective than drug B by an amount that is clinically important to patients.
Monte Carlo	A technique used to approximate the probability of certain outcomes by running multiple simulations using random variables.
Multivariate model	A statistical model for analysis of the relationship between 2 or more predictors, (independent) variables and the outcome (dependent) variable.

ISBN 978-1-4731-2661-9

Term	Definition
Net monetary benefit (NMB)	The value (usually in monetary terms) of an intervention net of its cost. The NMB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the NMB is calculated as: (£20,000×QALYs gained) minus cost.
Network meta-analysis	Meta-analysis in which multiple treatments (that is, 3 or more) are being compared using both direct comparisons of interventions within RCTs and indirect comparisons across trials based on a common comparator.
Non-inferiority trial	A trial designed to determine whether the effect of a new treatment is not worse than a standard treatment by more than a pre-specified amount. A one-sided version of an equivalence trial.
Number needed to treat (NNT)	The average number of patients who need to be treated to get a positive outcome. For example, if the NNT is 4, then 4 patients would have to be treated to ensure 1 of them gets better. The closer the NNT is to 1, the better the treatment. For example, if you give a stroke prevention drug to 20 people before 1 stroke is prevented, the number needed to treat is 20.
Observational study	Individuals or groups are observed or certain factors are measured. No attempt is made to affect the outcome. For example, an observational study of a disease or treatment would allow 'nature' or usual medical care to take its course. Changes or differences in one characteristic (for example, whether or not people received a specific treatment or intervention) are studied without intervening. There is a greater risk of selection bias than in experimental studies.
Odds ratio (OR)	Odds are a way to represent how likely it is that something will happen (the probability). An odds ratio compares the probability of something in one group with the probability of the same thing in another. An odds ratio of 1 between 2 groups would show that the probability of the event (for example, a person developing a disease, or a treatment working) is the same for both. An odds ratio greater than 1 means the event is more likely in the first group. An odds ratio less than 1 means that the event is less likely in the first group. Sometimes probability can be compared across more than 2 groups – in this case, one of the groups is chosen as the 'reference category' and the odds ratio is calculated for each group compared with the reference category. For example, to compare the risk of dying from lung cancer for non-smokers, occasional smokers and regular smokers, non-smokers could be used as the reference category. Odds ratios would be worked out for occasional smokers compared with non-smokers and for regular smokers compared with non-smokers. See also Confidence interval, Relative risk.
Opportunity cost	The loss of other healthcare programmes displaced by investment in or introduction of another intervention. This may be best measured by the health benefits that could have been achieved had the money been spent on the next best alternative healthcare intervention.
Outcome	The impact that a test, treatment, policy, programme or other intervention has on a person, group or population. Outcomes from interventions to improve the public's health could include changes in knowledge and behaviour related to health, societal changes (for example, a reduction in crime rates) and a change in people's health and wellbeing or health status. In clinical terms, outcomes could include the number of patients who fully recover from an illness or the number of hospital admissions, and an improvement or deterioration in someone's health, functional ability, symptoms or situation. Researchers should decide what outcomes to measure before a study begins.

ISBN 978-1-4731-2661-9

Term	Definition
Ovarian cystectomy	Ovarian cystectomy is a surgical excision of an ovarian endometriotic cyst. An ovarian endometrioma is a cystic mass arising from ectopic endometrial tissue within the ovary.
p value	The p value is a statistical measure that indicates whether or not an effect is statistically significant. For example, if a study comparing 2 treatments found that one seems more effective than the other, the p value is the probability of obtaining these results by chance. By convention, if the p value is below 0.05 (that is, there is less than a 5% probability that the results occurred by chance) it is considered that there probably is a real difference between treatments. If the p value is 0.001 or less (less than a 1% probability that the results occurred by chance), the result is seen as highly significant. If the p value shows that there is likely to be a difference between treatments, the confidence interval describes how big the difference in effect might be.
Paediatric and adolescent gynaecology service	Paediatric and adolescent gynaecology services are hospital-based, multidisciplinary specialist services for girls and young women (usually aged under 18).
Performance bias	Systematic differences between intervention groups in care provided apart from the intervention being evaluated. Blinding of study participants (both the recipients and providers of care) is used to protect against performance bias.
Peritoneal endometriosis	The peritoneum is the lining of the pelvis. Peritoneal endometriosis occurs when endometrial cells travel to and implant in the peritoneal wall.
Placebo	A fake (or dummy) treatment given to participants in the control group of a clinical trial. It is indistinguishable from the actual treatment (which is given to participants in the experimental group). The aim is to determine what effect the experimental treatment has had over and above any placebo effect caused because someone has received (or thinks they have received) care or attention.
Placebo effect	A beneficial (or adverse) effect produced by a placebo and not due to any property of the placebo itself.
Post-hoc analysis	Statistical analyses that are not specified in the trial protocol and are generally suggested by the data.
Power (statistical)	The ability to demonstrate an association when one exists. Power is related to sample size; the larger the sample size, the greater the power and the lower the risk that a possible association could be missed.
Prevalence	The prevalence of a disease is the proportion of a population that are cases at a point in time.
Primary care	Healthcare delivered outside hospitals. It includes a range of services provided by GPs, nurses, health visitors, midwives and other healthcare professionals and allied health professionals such as dentists, pharmacists and opticians.
Primary outcome	The outcome of greatest importance, usually the one in a study that the power calculation is based on.
Product licence	An authorisation from the Medicines and Healthcare Products Regulatory Agency (MHRA) to market a medicinal product.
Prognosis	A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is associated with a high rate of undesirable outcomes.
Prospective study	A research study in which the health or other characteristic of participants is monitored (or 'followed up') for a period of time, with events recorded as they happen. This contrasts with retrospective studies.

ISBN 978-1-4731-2661-9

Term	Definition
Protocol (review)	A document written prior to commencing a review that details exactly how evidence to answer a review question will be obtained and synthesised. It defines in detail the population of interest, the interventions, the comparators/controls and the outcomes of interest (PICO).
Publication bias	Publication bias occurs when researchers publish the results of studies showing that a treatment works well and do not publish those showing it did not have any effect. If this happens, analysis of the published results will not give an accurate idea of how well the treatment works. This type of bias can be assessed by a funnel plot.
Quality of life	See Health-related quality of life.
Quality adjusted life year (QALY)	A measure of the state of health of a person or group in which the benefits, in terms of length of life, are adjusted to reflect the quality-of-life. One QALY is equal to 1 year of life in perfect health. QALYS are calculated by estimating the years of life remaining for a patient following a particular treatment or intervention and weighting each year with a quality-of-life score (on a scale of 0 to 1). It is often measured in terms of the person's ability to perform the activities of daily life, and freedom from pain and mental disturbance.
Random effect model	In meta-analysis, a model that calculates a pooled effect estimate using the assumption that each study is estimating a different true treatment effect due to real differences between studies. Observed variation in effects are therefore caused by a combination of random sample variability (within-study variation) and heterogeneity between studies (between-study variation). The overall effects is an average of the estimated true study effects.
Randomisation	Assigning participants in a research study to different groups without taking any similarities or differences between them into account. For example, it could involve using a random numbers table or a computer-generated random sequence. It means that each individual (or each group in the case of cluster randomisation) has the same chance of receiving each intervention.
Randomised controlled trial (RCT)	A study in which a number of similar people are randomly assigned to 2 (or more) groups to test a specific drug or treatment. One group (the experimental group) receives the treatment being tested, the other (the comparison or control group) receives an alternative treatment, a dummy treatment (placebo) or no treatment at all. The groups are followed up to see how effective the experimental treatment was. Outcomes are measured at specific times and any difference in response between the groups is assessed statistically. This method is also used to reduce bias.
Reference standard	The test that is considered to be the best available method to establish the presence or absence of the outcome – this may not be the one that is routinely used in practice.
Relative risk (RR)	The ratio of the risk of disease or death among those exposed to certain conditions compared with the risk for those who are not exposed to the same conditions (for example, the risk of people who smoke getting lung cancer compared with the risk for people who do not smoke). If both groups face the same level of risk, the relative risk is 1. If the first group had a relative risk of 2, subjects in that group would be twice as likely to have the event happen. A relative risk of less than 1 means the outcome is less likely in the first group. Relative risk is sometimes referred to as risk ratio.
Reporting bias	See Publication bias.
Resource implication	The likely impact in terms of finance, workforce or other NHS resources.

ISBN 978-1-4731-2661-9

Term	Definition
Retrospective study	A research study that focuses on the past and present. The study examines past exposure to suspected risk factors for the disease or condition. Unlike prospective studies, it does not cover events that occur after the study group is selected.
Review question	The plan or set of steps to be followed in a study. A protocol for a systematic review describes the rationale for the review, the objectives and the methods that will be used to locate, select and critically appraise studies, and to collect and analyse data from the included studies.
Secondary care	Care provided in hospitals.
Secondary outcome	An outcome used to evaluate additional effects of the intervention deemed a priori as being less important than the primary outcomes.
Selection bias	Selection bias occurs if: • The characteristics of the people selected for a study differ from the wider population from which they have been drawn; or • There are differences between groups of participants in a study in terms of how likely they are to get better.
Sensitivity	How well a test detects the thing it is testing for. If a diagnostic test for a disease has high sensitivity, it is likely to pick up all cases of the disease in people who have it (that is, give a 'true positive' result). But if a test is too sensitive it will sometimes also give a positive result in people who do not have the disease (that is, give a 'false positive'). For example, if a test were developed to detect if a woman is 6 months pregnant, a very sensitive test would detect everyone who was 6 months pregnant but would probably also include those who are 5 and 7 months pregnant. If the same test were more specific (sometimes referred to as having higher specificity), it would detect only those who are 6 months pregnant and someone who was 5 months pregnant would get a negative result (a 'true negative'). But it would probably also miss some people who were 6 months pregnant (that is, give a 'false negative'). Breast screening is a 'real-life' example. The number of women who are recalled for a second breast screening test is relatively high because the test is very sensitive. If it were made more specific, people who don have the disease would be less likely to be called back for a second test but more women who have the disease would be missed.
Sensitivity analysis	A means of representing uncertainty in the results of an analysis. Uncertainty may arise from missing data, imprecise estimates or methodological controversy. Sensitivity analysis also allows for exploring the generalisability of results to other settings. The analysis is repeated using different assumptions to examine the effect on the results. • One-way simple sensitivity analysis (univariate analysis) – each parameter is varied individually in order to isolate the consequences of each parameter on the results of the study. • Multi-way simple sensitivity analysis (scenario analysis) – 2 or more parameters are varied at the same time and the overall effect on the results is evaluated. • Threshold sensitivity analysis – the critical value of parameters above or below which the conclusions of the study will change are identified. • Probabilistic sensitivity analysis – probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (for example, Monte Carlo simulation).
Significance (statistical)	A result is deemed statistically significant if the probability of the result occurring by chance is less than 1 in 20 ($p<0.05$).
Specificity	The proportion of true negatives that are correctly identified as such. For example, in diagnostic testing the specificity is the proportion of non-cases correctly diagnosed as non-cases. In terms of literature searching

ISBN 978-1-4731-2661-9

Term	Definition
	a highly specific search is generally narrow and aimed at picking up the key papers in a field and avoiding a wide range of papers. See also Sensitivity.
Spontaneous pregnancy	Pregnancy that was not assisted by reproductive treatment.
Stakeholder	An organisation with an interest in a topic on which NICE is developing a clinical guideline or piece of public health guidance. Organisations that register as stakeholders can comment on the draft scope and the draft guidance. Stakeholders may be: • manufacturers of drugs or equipment • national patient and carer organisations • NHS organisations • organisations representing healthcare professionals.
Standard deviation (SD)	A measure of the spread or dispersion of a set of observations, calculated as the average difference from the mean value in the sample.
Standardised incidence ratio	Standardised incidence ratio is the incidence rate in the endometriosis group (number of new cases of cancer in the endometriosis patients) is compared to a population incidence rate. A value >1.00 indicates a higher incidence rate in the endometriosis group, i.e. higher risk of cancer if a woman has endometriosis.
Subfertility	Any form or grade of reduced fertility with prolonged time of unwanted non-conception.
Subgroup analysis	An analysis in which the intervention effect is evaluated in a defined subset of the participants in a trial, or in complementary subsets.
Systematic review	A review in which evidence from scientific studies has been identified, appraised and synthesised in a methodical way according to predetermined criteria. It may include a meta-analysis.
Time horizon	The time span over which costs and health outcomes are considered in a decision analysis or economic evaluation.
Treatment allocation	Assigning a participant to a particular arm of a trial.
True negative	A diagnostic test result that correctly indicates that an individual does not have the disease of interest when they actually do not have it.
True positive	A diagnostic test result that correctly indicates that an individual has the disease of interest when they do actually have it.
Univariate	Analysis which separately explores each variable in a data set.
Utility	In health economics, a utility is the measure of the preference or value that an individual or society places upon a particular health state. It is generally a number between 0 (representing death) and 1 (perfect health). The most widely used measure of benefit in cost-utility analysis is the quality-adjusted life year, but other measures include disability-adjusted life years (DALYs) and healthy year equivalents (HYEs).

ISBN 978-1-4731-2661-9

ISBN 978-1-4731-2661-9

15 Appendices (see separate files)

Appendix A: Scope

Appendix B: Stakeholders

Appendix C: Declarations of interest

Appendix D: Review protocol

Appendix E: Search strategies

Appendix F: Summary of identified studies (Prisma charts)

Appendix G: Evidence tables

Appendix H: Excluded studies

Appendix I: Forest Plots

Appendix J: GRADE tables

Appendix K: Health economics

Appendix L: Network Meta-Analysis

ISBN 978-1-4731-2661-9