National Institute for Health and Care Excellence

Version 1.0 Consultation

Endometriosis: diagnosis and management

Appendix G

Appendix

Evidence Tables

19 January 2017

Draft for Consultation

Developed by the National Guidelines Alliance, hosted by the Royal College of Obsetricians 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.

ISBN:

Contents

1

Appendix G:

G.1 Review question: Specialist services

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

No clinical evidence was identified for this review.

G.2 Review question: Timing: association between duration of symptoms before laparoscopy and treatment outcomes

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

No clinical evidence was identified for this review.

G.3 Review question: Signs and symptoms of endometriosis (monitoring and referral)

- **What are the signs and symptoms of endometriosis?**
- 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?

<table>
<tr><th>Study details</th><th>Participants</th><th>Risk factor</th><th>Methods</th><th>Outcome and result</th><th>Comments</th></tr>
<tr>
<td>Full citation
Calhaz-Jorge, C., Mol, B. W., Nunes, J., Costa, A. P., Clinical predictive factors for endometriosis</td>
<td>Sample size
N=1079 (n=488 endometriosis, n=591 no endometriosis)

Characteristics</td>
<td>Risk factor
Pelvic pain (chronic pelvic pain)
Uterus: pain (dysmenorrhoea),</td>
<td>Method of measurement of risk factor
Personal interview a standard questionnaire regarding general characteristics (age at laparoscopy,</td>
<td>Outcome
<u>Results of the multivariate analysis</u>
<table>
<tr><td>Characteristic</td><td>OR endometriosis AFS any type</td><td>OR endometriosis AFS grade III/IV</td></tr>
</table>
</td>
<td>Limitations
<u>NICE prognostic study checklist</u>
Overall: Moderate quality</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Risk factor</th><th>Methods</th><th>Outcome and result</th><th>Comments</th></tr>
<tr>
<td>in a Portuguese infertile population, Human Reproduction, 19, 2126-31, 2004

Country/ies where study was carried out
Portugal

Study type
Prospective cohort

Study dates
1993-2000, Unit of Human Reproduction, Department of Obstetrics and Gynaecology, Hospital de Santa Maria in Lisbon

Aim of the study
To investigate factors that may be related to either minimal/mild or</td>
<td>
<table>
<tr><td>Characteristic</td><td>No endometriosis n=591</td><td>AFS grade I/II n=358</td><td>AFS grade III/IV n=130</td></tr>
<tr><td>Age, years (SD)</td><td>30.9 (4.2)</td><td>30.9 (3.9)</td><td>30.7 (4.0)</td></tr>
<tr><td>Dysmenorrhoea
No
Mild
Moderate
Severe</td><td>194 (64%)
219 (60%)
142 (45%)
36 (38%)</td><td>86 (28%)
116 (32%)
124 (%)
32 (34%)</td><td>23 (8%)
29 (8%)
51 (16%)
27 (28%)</td></tr>
<tr><td>Dyspareunia
No
Sometimes
Always
missing value</td><td>470 (56%)
100 (52%)
17 (49%)
4</td><td>278 (33%)
69 (36%)
11 (31%)
0</td><td>97 (11%)
24 (12%)
7 (20%)
2</td></tr>
<tr><td>Chronic pelvic pain (no/yes)</td><td>525/66</td><td>333/25</td><td>105/25</td></tr>
<tr><td>Menstrual flow
Mild
Moderate
Severe</td><td>161 (66%)
338 (51%)</td><td>70 (29%)
232 (35%)
56 (32%)</td><td>13 (5%)
91 (14%)
26 (15%)</td></tr>
</table>
</td>
<td>abnormal bleeding (prolonged and heavy)
Vaginal pain (dyspareunia)</td>
<td>weight and height, race, education), lifestyle habits (smoking), reproductive history (obstetric history, duration of subfertility and use or oral contraceptives), menstrual characteristics (age at menarche, average duration of bleeding and average cycle length), presence and intensity of pelvic symptomatology (dysmenorrhoea, dyspareunia and pelvic pain)
Dysmenorrhoea definition: mild (mild discomfort with no use of analgesic medication), moderate (significant pain with need of analgesic medication most of the time), severe (intense pain with a need for medication every menstrual flow, with or without a need for bed rest and absence from work)</td>
<td>
<table>
<tr><td>Negroid women</td><td>0.50 (0.30-0.83)</td><td></td></tr>
<tr><td>Dysmenorrhoea any type</td><td></td><td>2.5 (1.2-5.2)</td></tr>
<tr><td>Mild dysmenorrhoea</td><td>0.62 (0.46-0.83)</td><td></td></tr>
<tr><td>Moderate dysmenorrhoea</td><td></td><td>1.7 (1.1-2.7)</td></tr>
<tr><td>Severe dysmenorrhoea</td><td></td><td>2.8 (1.5-5.1)</td></tr>
<tr><td>Recently intensified dysmenorrhoea</td><td></td><td>2.4 (1.3-4.5)</td></tr>
<tr><td>Primary dysmenorrhoea</td><td>1.4 (1.0-1.9)</td><td></td></tr>
<tr><td>Dysmenorrhoea day 1-2</td><td>1.4 (1.1-1.7)</td><td></td></tr>
<tr><td>Chronic pelvic pain</td><td></td><td>2.0 (1.2-3.4)</td></tr>
<tr><td>Generally regular menstrual cycle</td><td></td><td>0.60 (0.38-0.94)</td></tr>
</table>
</td>
<td>See following row for details</td>
</tr>
</table>

Study details	Participants	Risk factor	Methods	Outcome and result	Comments
moderate/severe endometriosis. To evaluate whether data from the clinical history and symptomatology could predict the presence of endometriosis at laparoscopy. Source of funding None described.	<table><tr><td></td><td>92 (53%)</td><td></td><td></td></tr><tr><td>OAC never ever</td><td>176 (64%) 415 (51%)</td><td>76 (28%) 282 (35%)</td><td>21 (8%) 109 (14%)</td></tr><tr><td>Duration of OAC use (per year)</td><td>3.5 (3.2)</td><td>3.9 (3.2)</td><td>4.6 (3.2)</td></tr><tr><td>Duration of menstrual flow (SD)</td><td>4.5 (1.7)</td><td>4.4 (1.3)</td><td>4.5 (1.4)</td></tr></table> **Inclusion criteria** • Subfertile women who underwent either diagnostic or therapeutic laparoscopy (subfertile definition: period of at least 12 months without conception despite unprotected intercourse) • previous pelvic surgery not excluded **Exclusion criteria** • Medical treatment within 3 months prior to laparoscopy		**Outcome ascertainment measure** Laparoscopy- any day of the menstrual cycle except during menstruation Endometriosis definition: direct visualization or biopsy of lesions No blind biopsies of apparently normal peritoneum was taken Staging according to American Society for Reproductive Medicine (AFS, 1985) **Statistical method** Classed as no endometriosis, minimal to mild, moderate to severe endometriosis Logistic regression analysis. Dependent variable: endometriosis Potential predictors: data from the medical history and clinical symptoms	<table><tr><td>Irregular cycle</td><td>0.60 (0.43-0.84)</td><td>0.29 (0.15-0.54)</td></tr><tr><td>BMI <20kg/m2</td><td>1.7 (1.2-2.5)</td><td></td></tr><tr><td>BMI 25-30kg/m2</td><td>0.65 (0.47-0.91)</td><td></td></tr><tr><td>BMI >30kg/m2</td><td>0.33 (0.18-0.59)</td><td></td></tr><tr><td>Smoker 1-10 cigarettes/ day</td><td>0.57 (0.39-0.79)</td><td></td></tr><tr><td>Smoker 11-20 cigarettes/ day</td><td>0.52 (0.34-0.79)</td><td>0.47 (0.22-1.02)</td></tr><tr><td>Smoker >20 cigarettes/ day</td><td>0.56 (0.32-0.99)</td><td></td></tr><tr><td>Previous pregnancy</td><td>0.65 (0.49-0.87)</td><td>0.58 (0.37-0.92)</td></tr><tr><td>Ever use of oral contraceptives</td><td>1.6 (1.2-2.3)</td><td>2.2 (1.3-3.7)</td></tr><tr><td>AUC</td><td>0.71</td><td>0.74</td></tr></table> Calibration of the model reported as good.	

Study details	Participants	Risk factor	Methods	Outcome and result	Comments
			Univariate and multivariate analysis (performed twice; presence of any type of endometriosis, presence of moderate to severe endometriosis) MVA: stepwise logistic regression, p value of 0.5 as entry criterion, p value of 0.1 for a variable to stay in the model • AUC calculated • Calibration of the model **Confouders included in multivariate analysis model** <u>Critical confounders</u> • OAC use • Age **Length of follow-up** NA		
NICE prognostic study checklist for: 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 <u>The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results</u> Are the source population or the population of interest adequately described with respect to key characteristics? Yes Are the sampling frame and recruitment adequately described, possibly including methods to identify the sample (number and type used; for example, referral patterns in healthcare), period of recruitment and place of recruitment (setting and geographical location)? consecutive recruitment Are inclusion and exclusion criteria adequately described (for example, including explicit diagnostic criteria or a description of participants at the start of the follow-up period)? yes					

Study details	Participants	Risk factor	Methods	Outcome and result	Comments
Is participation in the study by eligible individuals adequate? yes Is the baseline study sample (that is, individuals entering the study) adequately described with respect to key characteristics? yes Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias Is the response rate (that is, proportion of study sample completing the study and providing outcome data) adequate? No women were reported not to participate/ having inadequate data. Some missing data at baseline but minimal. Are attempts to collect information on participants who dropped out of the study described? NA Are reasons for loss to follow-up provided? NA Are the key characteristics of participants lost to follow-up adequately described? NA Are there any important differences in key characteristics and outcomes between participants who completed the study and those who did not? NA The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias Is a clear definition or description of the prognostic factor(s) measured provided (including dose, level, duration of exposure, and clear specification of the method of measurement)? Only definition of dysmenorrhoea given. Are continuous variables reported, or appropriate cut-off points (that is, not data-dependent) used? Yes for BMI. Are the prognostic factors measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as blind measurement and limited reliance on recall.) Interview-recall risk of bias. Are complete data for prognostic factors available for an adequate proportion of the study sample? Yes Are the method and setting of measurement the same for all study participants? Yes Are appropriate methods employed if imputation is used for missing data on prognostic factors? Not described. The outcome of interest is adequately measured in study participants, sufficient to limit potential bias Is a clear definition of the outcome of interest provided, including duration of follow-up? Yes definition of endometriosis and grading given Are the outcomes that were measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as 'blind' measurement and limited reliance on recall.) Unclear how many were visual/ biopsied and if surgeon was blinded to clinical history. Are the method and setting of measurement the same for all study participants? Yes for setting/ unclear who had biopsies. Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest Are all important confounders, including treatments (key variables in the conceptual model), measured? Are clear definitions of the important confounders measured (including dose, level and duration of exposures) provided? Yes for age. OC measured but not other hormonal contraceptives. Is measurement of all important confounders valid and reliable? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as 'blind' measurement and limited reliance on recall.)- interview, risk of recall bias. Are the method and setting of measurement of confounders the same for all study participants? Yes Are appropriate methods employed if imputation is used for missing data on confounders? Not described. Are important potential confounders accounted for in the study design (for example, matching for key variables, stratification or initial assembly of comparable groups)? Age and OC in MVA. Are important potential confounders accounted for in the analysis (that is, appropriate adjustment)? As above. The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results Is the presentation of data sufficient to assess the adequacy of the analysis? Yes. Where several prognostic factors are investigated, is the strategy for model building (that is, the inclusion of variables) appropriate and based on a conceptual framework or model? Yes					

<table>
<tr><th>Study details</th><th>Participants</th><th>Risk factor</th><th>Methods</th><th>Outcome and result</th><th>Comments</th></tr>
<tr><td colspan="6">Is the selected model adequate for the design of the study? Yes
Is there any selective reporting of results? No
Note: generalisability of results due to subfertile population (prevalence of endometriosis 45%). Inter-observer variability of grading of the endometriosis without biopsies.
Overall: moderate quality</td></tr>
<tr>
<td>Full citation
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</td>
<td>Sample size
N=495 women (operative cohort)
N=131 women (population cohort)- 'at risk of endometriosis'
Excluded: n=26 due to no diagnostic information, given cancellation of surgery (n=22), unreadable MRIs (n=4)

Characteristics
<table>
<tr><td>Characteristic</td><td colspan="2">Operative cohort</td><td colspan="2">Population cohort</td></tr>
<tr><td></td><td>Endometriosis n=190</td><td>No endometriosis n=283</td><td>Endometriosis n=14</td><td>No endometriosis n=113</td></tr>
<tr><td>Mean age (SD)</td><td>31.98 (6.75)</td><td>33.61 (7.09)</td><td>33.14 (8.33)</td><td>32.07 (7.76)</td></tr>
<tr><td>Ever sexually active (Y/N)</td><td>163/27</td><td>244/37</td><td>13/1</td><td>99/14</td></tr>
<tr><td>Ever use oral contraceptives (Y/N)</td><td>169/21</td><td>238/45</td><td>13/1</td><td>96/17</td></tr>
</table></td>
<td>Risk factor
Pelvic symptoms (pelvic pain, surgical indication for laparoscopy: pelvic pain vs other)
Uterus: pain (dysmenorrhea)
Infertility</td>
<td>Method of measurement of risk factor
Patients given a study packet introducing study
Research assistants screened and recruited women by telephone or in person
Standardized data collection protocol included a computer assisted interview administered at baseline, and anthropometric assessment (BMI and skin fold) and biospecimen collection for quantification of environmental chemicals
Women were queried on sociodemographic characteristics, medical and reproductive history, pain and lifestyle
Protocol done prior to surgery and at the</td>
<td>Outcome
<u>Logistic regression model results</u>
Adjusted for: age and site

<u>Risk factors for endometriosis by cohort:</u>
<table>
<tr><td rowspan="2">Risk factor</td><td colspan="2">Operative cohort n=473</td><td colspan="2">Population cohort n=127</td></tr>
<tr><td>Unadjusted OR (95% CI)</td><td>Adjusted OR (95% CI)</td><td>Unadjusted OR (95% CI)</td><td>Adjusted OR (95% CI)</td></tr>
<tr><td>Age, y</td><td>0.97 (0.94 - 0.99)</td><td>-</td><td>1.02 (0.95 - 1.09)</td><td>-</td></tr>
<tr><td>Infertility history (Y/N)</td><td>2.49 (1.61 - 3.83)</td><td>2.43 (1.57 - 3.76)</td><td>7.13 (1.72 - 29.6)</td><td>7.91 (1.69 - 37.2)</td></tr>
<tr><td>Surgical indication for laparoscopy</td><td>3.91 (2.65 - 5.76)</td><td>3.67 (2.44 - 5.50)</td><td>-</td><td>-</td></tr>
</table></td>
<td>Limitations
<u>NICE prognostic study checklist</u>
Overall moderate quality (see following row)</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Risk factor</th><th>Methods</th><th>Outcome and result</th><th>Comments</th></tr>
<tr>
<td>Country/ies where study was carried out
USA- Salt Lake City and San Francisco.

Study type
Prospective matched (with surgery being the exposure) cohort

Study dates
2007-2009

Aim of the study
To identify risk factors for endometriosis and their consistency across study populations int eh Endometriosis: Natural History, Diagnosis and Outcomes (ENDO) study.</td>
<td>
<table>
<tr><td>Gravidity, mean (SD)</td><td>1.65 (1.98)</td><td>2.28 (2.12)</td><td>2.21 (2.08)</td><td>1.65 (1.80)</td></tr>
<tr><td>History of STIs (Y/N)</td><td>30/160</td><td>64/219</td><td>1/13</td><td>22/91</td></tr>
<tr><td>Ever seek infertility treatment (Y/N)</td><td>64/126</td><td>48/235</td><td>4/10</td><td>6/107</td></tr>
<tr><td>Surgical indication
Pelvic pain
Pelvic mass
Menstrual irregularity
Fibroids
Tugal ligation
Infertility</td><td>

120
26
20
9
8
7</td><td>

86
48
40
40
40
28</td><td></td><td></td></tr>
<tr><td>Pelvic pain > 6 months affecting normal function (Y/N)</td><td>84/106</td><td>98/184</td><td>1/13</td><td>11/102</td></tr>
<tr><td>Painful menses (Y/N)</td><td>94/91</td><td>89/179</td><td>1/12</td><td>11/98</td></tr>
</table>
</td>
<td></td>
<td>earliest time for population cohort (approx 2 months prior to surgery or MRI)
Note: remuneration was given for time and travel

Outcome ascertainment measure
<u>Operative cohort:</u>
Definition of endometriosis: visualization by the surgeon
Histological endometriosis: presence of endometrial glands and/or stroma and/or hemosiderin laden macrophages
<u>Population cohort:</u>
Definition of endometriosis:MRI visualised endometriosis. Primarily ovarian endometriomas but also included nodular implants
MRI of the pelvis in those without prior surgery. To assess visceral fat</td>
<td>
<table>
<tr><td>(pelvic pain vs other)</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Dysmenorrhea (Y/N)</td><td>2.78 (1.46 - 5.29)</td><td>2.46 (1.28 - 4.72)</td><td>1.37 (0.28 - 6.58)</td><td>1.41 (0.28 - 7.14)</td></tr>
<tr><td>Pelvic pain (Y/N)</td><td>0.95 (0.93 - 0.98)</td><td>1.39 (0.95 - 2.04)</td><td>1.01 (0.93 - 1.09)</td><td>0.76 (0.09 - 6.54)</td></tr>
</table>
One consistent risk factor across the cohorts: a history of infertility.

<u>Risk factors for visually and histologically confirmed endometriosis</u>
<table>
<tr><td rowspan="2">Risk factor</td><td colspan="2">Operative cohort n=473</td></tr>
<tr><td>Unadjusted OR (95% CI)</td><td>Adjusted OR (95% CI)</td></tr>
<tr><td>Age, y</td><td>0.97 (0.93-1.00)</td><td>-</td></tr>
<tr><td>Infertility history (Y/N)</td><td>2.43 (1.40-4.20)</td><td>2.39 (1.38-4.16)</td></tr>
<tr><td>Surgical indication for laparoscopy (pelvic</td><td>3.01 (1.74-5.22)</td><td>2.82(1.59-4.99)</td></tr>
</table>
</td>
<td></td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Risk factor</th><th>Methods</th><th>Outcome and result</th><th>Comments</th></tr>
<tr>
<td>Source of funding
Funded by the Intramural Research Program, Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), National Institutes of Health. Ethicon Endo-Surgery LLC donated shears and scalpel blades though a signed Materials Transfer Agreement with the University of Utah and the NICHD.</td>
<td>Inclusion criteria
Surgical cohort:
• Menstruating women
• Aged 18-44 years
• Underwent a diagnostic and/or therapeutic laparoscopy or laparotomy at 1 of 5 participating centres in Salt Lake City area (n=432) or 1 of 9 sites int eh SanFrancisco area (n=63)
• Any surgical indication was acceptable: pelvic pain (n=206), pelvic mass (n=74), menstrual irregularities (n=60), fibroids (n=49), tubal ligation (n=48) and infertility (n=35)
Population cohort
• Matched (age and residence within a 50 mile geographic catchment area)
• Currently menstuating women
• No history of surgically confirmed endometriosis

Exclusion criteria
• Previous laparoscopic diagnosis of endometriosis
• Currently breastfeeding ≥6 months (because of its likely impact lowering concentrations of environmental chemicals)
• History of cancer other than nonmelanoma skin cancer
• Use of injectable hormonal therapy within the past 2 years that may affect somatic presentation
• Inability to communicate in Spanish or English</td>
<td></td>
<td>distribution and any gynecologic pathology including endometriosis. FDA approved protocol for imaging
1 radiologist supervised and evaluated all MRIs. Findings confirmed by second radiologist (specialist in gynae imaging)

Statistical method
Unadjusted odds ratio for all risk factors
Logistic regression model: included all significant ORs along with age (in years) and clinical site (Utah or California) to account for potential residual confounding
Separate models for each cohort
Sensitivity analyses: restricting endometriosis to visually and histologically confirmed disease, restricting to moderate or severe disease (stages 3</td>
<td>
<table>
<tr><td>pain vs other)</td><td></td><td></td></tr>
<tr><td>Dysmenorrhea (Y/N)</td><td>3.49 (1.06-11.5)</td><td>3.11(0.94-10.3)</td></tr>
<tr><td>Pelvic pain (Y/N)</td><td>1.72 (1.02-2.91)</td><td>1.63 (0.96-2.76)</td></tr>
</table>

<u>Risk factors for stages 3 and 4 endometriosis</u>
<table>
<tr><td rowspan="2">Risk factor</td><td colspan="2">Operative cohort n=473</td></tr>
<tr><td>Unadjusted OR (95% CI)</td><td>Adjusted OR (95% CI)</td></tr>
<tr><td>Age, y</td><td>0.99 (0.95-1.03)</td><td>-</td></tr>
<tr><td>Infertility history (Y/N)</td><td>4.90 (2.66-9.00)</td><td>4.74 (2.57-8.75)</td></tr>
<tr><td>Surgical indication for laparoscopy (pelvic pain vs other)</td><td>4.44 (2.42-8.16)</td><td>4.47 (2.39-8.38)</td></tr>
<tr><td>Dysmenorrhea (Y/N)</td><td>3.61 (1.08-12.0)</td><td>3.43(1.02-11.5)</td></tr>
</table>
</td>
<td></td>
</tr>
</table>

Study details	Participants	Risk factor	Methods	Outcome and result	Comments
			and 4) or restricitng the comparison group of women to those with a postoperative diagnosis of a a 'normal pelvis' **Confouders included in multivariate analysis model** • Risk factors included in the logistic regression model: • Infertility history • Surgical indication for laparoscopy (pelvic pain vs other) • Dysmenorrhea • Pelvic pain • age • above poverty level • college educated • gravid • parous • age at first consenting sex • age at menarche • mean no. of periods • mean cycle length	Pelvic pain (Y/N) 1.63 (0.91-2.91) 1.60 (0.89-2.87)	

Study details	Participants	Risk factor	Methods	Outcome and result	Comments
			• mean length shortest cycle • mean length longest cycle • BMI Hormonal contraception (OC) was recorded for the two groups. It is assumed that there was no significant difference between those with and without endometriosis for both groups as it was not included in the logistic regression model. **Length of follow-up** NA. The study went on for 2 years. Approximate time from protocol reviewing and surgery/MRI was 2 months.		
NICE prognostic study checklist for: 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 The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results Are the source population or the population of interest adequately described with respect to key characteristics? Yes					

Study details	Participants	Risk factor	Methods	Outcome and result	Comments
Are the sampling frame and recruitment adequately described, possibly including methods to identify the sample (number and type used; for example, referral patterns in healthcare), period of recruitment and place of recruitment (setting and geographical location)? Not in this study but the methods are referred to being in an additional paper Buck 2011. Are inclusion and exclusion criteria adequately described (for example, including explicit diagnostic criteria or a description of participants at the start of the follow-up period)? Yes Is participation in the study by eligible individuals adequate? Does not report how many did not want to participate Is the baseline study sample (that is, individuals entering the study) adequately described with respect to key characteristics? Yes Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias Is the response rate (that is, proportion of study sample completing the study and providing outcome data) adequate? 26 women did not have diagnostic data and were excluded (4% operative cohort n=22, 2% population cohort,n=4) Are attempts to collect information on participants who dropped out of the study described? No Are reasons for loss to follow-up provided? Yes Are the key characteristics of participants lost to follow-up adequately described? No Are there any important differences in key characteristics and outcomes between participants who completed the study and those who did not? Not described. Unclear The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias Is a clear definition or description of the prognostic factor(s) measured provided (including dose, level, duration of exposure, and clear specification of the method of measurement)? No details given as to the questions used to determine the risk factors Are continuous variables reported, or appropriate cut-off points (that is, not data-dependent) used? No Are the prognostic factors measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as blind measurement and limited reliance on recall.) No Are complete data for prognostic factors available for an adequate proportion of the study sample? Yes Are the method and setting of measurement the same for all study participants? Yes Are appropriate methods employed if imputation is used for missing data on prognostic factors? Not reported The outcome of interest is adequately measured in study participants, sufficient to limit potential bias Is a clear definition of the outcome of interest provided, including duration of follow-up? Yes. F/U NA. Are the outcomes that were measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as 'blind' measurement and limited reliance on recall.) Yes for surgery and histology. Are the method and setting of measurement the same for all study participants? Different centres. Unclear if laparoscopy or laparotomy. Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest Are all important confounders, including treatments (key variables in the conceptual model), measured? Only oral contraceptive was listed for hormonal contraceptives. Are clear definitions of the important confounders measured (including dose, level and duration of exposures) provided? No					

<table>
<tr><th>Study details</th><th>Participants</th><th>Risk factor</th><th>Methods</th><th>Outcome and result</th><th>Comments</th></tr>
<tr><td colspan="6">Is measurement of all important confounders valid and reliable? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as 'blind' measurement and limited reliance on recall.)No restricted to recall.
Are the method and setting of measurement of confounders the same for all study participants? Yes
Are appropriate methods employed if imputation is used for missing data on confounders? Not reported.
Are important potential confounders accounted for in the study design (for example, matching for key variables, stratification or initial assembly of comparable groups)? Age and site matched.
Are important potential confounders accounted for in the analysis (that is, appropriate adjustment)? Adjusted for age and site.
<u>The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results</u>
Is the presentation of data sufficient to assess the adequacy of the analysis? Yes
Where several prognostic factors are investigated, is the strategy for model building (that is, the inclusion of variables) appropriate and based on a conceptual framework or model? Yes
Is the selected model adequate for the design of the study? Yes
Is there any selective reporting of results? Unlikely
Are only pre-specified hypotheses investigated in the analyses? Yes
Overall moderate quality</td></tr>
<tr>
<td>Full citation
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: JOGC, 34, 552-7, 2012</td>
<td>Sample size
N=429 (n=168 endometriosis, n=261 no endometriosis)

Characteristics
<table>
<tr><td>Predictor variable</td><td>No endometriosis</td><td>Endometriosis</td><td>P value</td></tr>
<tr><td>Age, mean (SD), years</td><td>33.7 (4.7)</td><td>34.1 (4.1)</td><td>0.63</td></tr>
<tr><td>Primary infertility, n (%)</td><td>122 (47)</td><td>109 (65)</td><td><0.001</td></tr>
<tr><td>Duration of infertility, years, mean (SD)</td><td>2.9 (2.7)</td><td>2.4 (2.0)</td><td>0.21</td></tr>
</table></td>
<td>Risk factor
Pelvic symptoms (chronic pelvic pain)
Uterus (dysmenorrhea)
Vaginal pain (dyspareunia)
Infertility (type and duration of)
Pelvic signs (uterosacral /cul-de-sac tenderness</td>
<td>Method of measurement of risk factor
Standard questionnaire before the initial visit - severity of dysmenorrhea (absent, mild, moderate, severe), deep dyspareunia (present/absent) and chronic pelvic pain (present/absent)
Pelvic examination
Offered HSG and the majority of hysterosalpingograms performed at one radiology centre,</td>
<td>Outcome
<u>Logistic regression results</u>
<table>
<tr><td>Predictor variable</td><td>β-coefficient</td><td>Odds ratio</td><td>95% CI</td><td>P value</td></tr>
<tr><td>Primary infertility</td><td>0.68</td><td>1.98</td><td>1.29-3.04</td><td>0.002</td></tr>
<tr><td>Degree of dysmenorrhea</td><td>0.30</td><td>1.34</td><td>1.10-1.65</td><td>0.005</td></tr>
<tr><td>Uterosacral/ cul-de-sac nodularity</td><td>1.34</td><td>3.81</td><td>1.64-8.83</td><td>0.002</td></tr>
</table></td>
<td>Limitations
<u>NICE prognostic study checklist</u>
Overall moderate quality

(See following row)</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Risk factor</th><th>Methods</th><th>Outcome and result</th><th>Comments</th></tr>
<tr>
<td>Country/ies where study was carried out
Canada

Study type
Retrospective cohort

Study dates
2002-2005

Aim of the study
To determine which clinical factors including symptoms, signs, and HSG findings are independent predictors of finding endometriosis at laparoscopy in infertile women, using logistic regression.

Source of funding</td>
<td>
<table>
<tr><td>Dysmenorrhoea
None
Mild
Moderate
Severe</td><td>
90 (34)
82 (31)
60 (23)
29 (11)</td><td>
37 (22)
40 (24)
53 (32)
38 (23)</td><td>

<0.001</td></tr>
<tr><td>Deep dyspareunia</td><td>20 (8)</td><td>26 (15)</td><td>0.02</td></tr>
<tr><td>Chronic pelvic pain</td><td>33 (13)</td><td>31 (18)</td><td>0.13</td></tr>
<tr><td>Uterosacral/cul-de-sac tenderness</td><td>10 (4)</td><td>20 (12)</td><td>0.002</td></tr>
<tr><td>Utersacral/cul-de-sac nodularity</td><td>9 (3)</td><td>23 (14)</td><td><0.001</td></tr>
<tr><td>HSG
Intrauterine filling defect
Polypoid endometrium</td><td>
45 (17)
2 (1)</td><td>
27 (16)
5 (3)</td><td>
0.79
0.12</td></tr>
<tr><td>Physician specific Endometriosis-focused practice</td><td>56 (21)</td><td>78 (46)</td><td><0.001</td></tr>
</table>
Inclusion criteria
• Women with no prior laparoscopic diagnosis of endometriosis, having a laparoscopy performed (by gynaecologic infertility specialists at the British Columbia
</td>
<td>and nodularity)</td>
<td>read by same radiologist
Decision for laparoscopy for infertility made by individual clinician and patient

Outcome ascertainment measure
Laparoscopy: performed by gynae infertility specialists (n=3, biopsy suspected lesions typical or atypical and confirm with histology or make a visual diagnosis if typical in appearance) or gynae infertility specialists with an endometriosis-focused practice (n=2, uniformly excise all suspected lesions of endometriosis whether typical or atypical and confirm diagnosis on histology)

Statistical method
Multiple logistic regression modelling</td>
<td>
<table>
<tr><td>Endometriosis-focussed practice of gynaecologist</td><td>1.08</td><td>2.94</td><td>1.88-4.60</td><td><0.001</td></tr>
</table>
OR=Ex[β-coefficient]
For degree of dysmenorrhea: OR represents (1) odds of endometriosis in severe dysmenorrhea/ odds of endometriosis in moderate dysmenorrhea, (2) odds of endometriosis in moderate dysmenorrhea/ odds of endometriosis in mild dysmenorrhea and (3) the odds of endometriosis in mild dysmenorrhea/odds of endometriosis in absent dysmenorrhea.
There were no statistically significant squared or 2 x 2 interaction terms.
Also reports probabilities of endometriosis depending on infertility status, severity of dysmenorrhea and presence of uterosacra/ cul-de-sac nodularity.
</td>
<td></td>
</tr>
</table>

Study details	Participants	Risk factor	Methods	Outcome and result	Comments
None described.	Women's Centre fro Reproductive Health) between 2002-2005 • Medical records available on site **Exclusion criteria** • Not having HSG performed • Incomplete medical records (questionnaire not completed or pelvic examination findings not available)		performed using likelihood ratio modelling All squared terms (predictor variable squared) and 2 x 2 interaction terms (e.g. age x type of infertility, n=55) were test for with significance set at p<0.01 for multiple comparisons Final logistic regression model, the OR represents binary variables: equal to the odds with the variable present divided by the odd with variable absent scaled or ordinal variables: equal to the odds with the variable = n+1 divided by the odds with the variable=n (e.g. the odds with severe dysmenorrhea divided by the odds with moderate dysmenorrhea) **Confouders included in**		

Study details	Participants	Risk factor	Methods	Outcome and result	Comments
			multivariate analysis model Critical confounders: • Age Hormonal contraception was not included in the analysis. **Length of follow-up** NA		

NICE prognostic study checklist for: 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: JOGC, 34, 552-7, 2012

The study sample represents the population of interest with regard to key characteristics, sufficient to limit potential bias to the results

Are the source population or the population of interest adequately described with respect to key characteristics? Yes apart from no data on hormonal contraceptive use.

Are the sampling frame and recruitment adequately described, possibly including methods to identify the sample (number and type used; for example, referral patterns in healthcare), period of recruitment and place of recruitment (setting and geographical location)? Yes

Are inclusion and exclusion criteria adequately described (for example, including explicit diagnostic criteria or a description of participants at the start of the follow-up period)? Yes

Is participation in the study by eligible individuals adequate? Unclear who declined to participate.

Is the baseline study sample (that is, individuals entering the study) adequately described with respect to key characteristics? Yes apart from use of hormonal contraceptives.

Loss to follow-up is unrelated to key characteristics (that is, the study data adequately represent the sample), sufficient to limit potential bias

Is the response rate (that is, proportion of study sample completing the study and providing outcome data) adequate? Unclear who declined to participate (part of exclusion criteria if insufficient data etc.

Are attempts to collect information on participants who dropped out of the study described? NA as no drop outs.

Are reasons for loss to follow-up provided? NA

Are the key characteristics of participants lost to follow-up adequately described? NA

Are there any important differences in key characteristics and outcomes between participants who completed the study and those who did not? NA

The prognostic factor of interest is adequately measured in study participants, sufficient to limit potential bias

Study details	Participants	Risk factor	Methods	Outcome and result	Comments
Is a clear definition or description of the prognostic factor(s) measured provided (including dose, level, duration of exposure, and clear specification of the method of measurement)? No clear definitions given. Unclear/ inaccurate measurement of dysmenorrhea etc. Are continuous variables reported, or appropriate cut-off points (that is, not data-dependent) used? yes come continuous e.g. age, duration of infertility Are the prognostic factors measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as blind measurement and limited reliance on recall.) Reliance on recall and medical notes Are complete data for prognostic factors available for an adequate proportion of the study sample? Yes - part of exclusion criteria if inadequate. Are the method and setting of measurement the same for all study participants? Yes Are appropriate methods employed if imputation is used for missing data on prognostic factors? Not reported. <u>The outcome of interest is adequately measured in study participants, sufficient to limit potential bias</u> Is a clear definition of the outcome of interest provided, including duration of follow-up? Visual or histological confirmation of endometriosis at laparoscopy. Are the outcomes that were measured and the method of measurement valid and reliable enough to limit misclassification bias? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as 'blind' measurement and limited reliance on recall.) Yes. Risk of underdiagnosis in physicians without an endometriosis focussed practice. Are the method and setting of measurement the same for all study participants? Yes <u>Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest</u> Are all important confounders, including treatments (key variables in the conceptual model), measured? Are clear definitions of the important confounders measured (including dose, level and duration of exposures) provided? No information on hormonal contraceptive use. Is measurement of all important confounders valid and reliable? (This may include relevant outside sources of information on measurement properties, as well as characteristics such as 'blind' measurement and limited reliance on recall.) Yea. Are the method and setting of measurement of confounders the same for all study participants? Yes Are appropriate methods employed if imputation is used for missing data on confounders? NA Are important potential confounders accounted for in the study design (for example, matching for key variables, stratification or initial assembly of comparable groups)? Are important potential confounders accounted for in the analysis (that is, appropriate adjustment)? <u>The statistical analysis is appropriate for the design of the study, limiting potential for the presentation of invalid results</u> Is the presentation of data sufficient to assess the adequacy of the analysis? Yes Where several prognostic factors are investigated, is the strategy for model building (that is, the inclusion of variables) appropriate and based on a conceptual framework or model? Yes Is the selected model adequate for the design of the study? Yes Is there any selective reporting of results? Unlikely Are only pre-specified hypotheses investigated in the analyses? Yes **Overall moderate quality**					

AFS: American Fertility Society; AUC: Area under the curve; BMI: Body mass index; CI: Confidence Interval; FDA: Food and Drug Administration; F/U: Follow-up; HSG: hysterosalpingogram; MRI: Magnetic resonance imaging; MVA: Multivariable analysis; NICHD: National Institute of Child Health and Human Development; OAC: Oral contraceptive; OC: Oral contraceptive; OR: Odds ratio; SD: Standard deviation;

G.4 Review question: Information and support

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?

Study details	Participants	Methods	Findings/results	Limitations
Full citation 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 **Ref Id** 401041 **Aim(s)** To investigate possible reasons for a delayed diagnosis of endometriosis and examine the impact that this has on women's experiences of the condition. **Study type** Qualitative study.	**Sample size** 32 women **Characteristics** • Women were aged 16 to 47 years • Length of time of pelvic pain: median 15 years • Diagnostic delay: 2 years • 46% women experienced symptoms for over 10 years before diagnosis **Inclusion criteria** • Women with suspected or confirmed diagnosis of endometriosis **Exclusion criteria**	**Setting** Women attending a pelvic pain clinic **Data collection** • Data was collected by face-to-face in depth semi-structured interviews carried out in the woman's home, hospital or in the university. **Data analysis** • A thematic approach was applied to the analysis, and quotations were collated and organised by similarities and differences.	**Themes and categories** **Facilitators** • Relief of diagnosis • Sense of control over symptoms **Barriers** • Delayed diagnosis (at individual or medical level) • Unnecessary diagnostic investigations • Seeing many doctors before seeing a doctor who would be sympathetic to women's problems • Doctors not taking women seriously, and trivialising their concerns about symptoms	**Aims** Clearly reported. Aim of study clearly reported, research method was appropriate for answering the research question. **Sample selection** Sample selection was reported. The relationship between the researcher and the respondents was reported. **Data collection** Data was collected through interviews conducted by the researcher. Some discussion around identification of themes was discussed but there was no discussion on data saturation. **Data analysis**

<table>
<tr><th>Study details</th><th>Participants</th><th>Methods</th><th>Findings/results</th><th>Limitations</th></tr>
<tr><td>Study dates
May 2004 to April 2005.

Source of funding
Not reported</td><td></td><td></td><td></td><td>The analytical process was described in detail. The researchers did not critically review their own roles in the process.

Findings/results
Results were presented clearly (e.g., citation/data and the researchers' own input distinguished; the researchers' roles and potential influences in the analytical process were not critically reviewed).

Overall quality
Low

Other information
None</td></tr>
<tr><td>Full citation
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
Ref Id
403152

Aim(s):</td><td>Sample size
A survey was responsed by 670 women and 61 women participated in the focus group meetings.

Characteristics
Focus group demographics<table><tr><th>Age</th><th>Number</th></tr><tr><td>20-24</td><td>5</td></tr><tr><td>25-29</td><td>10</td></tr><tr><td>30-34</td><td>19</td></tr><tr><td>35-39</td><td>9</td></tr><tr><td>40-44</td><td>9</td></tr></table></td><td>Setting
Epworth hospital in Melbourne

Data collection
A survey and five focus groups designed to determine consumer needs for information related to day surgery for endometriosis-related problems.
In the focus groups, women were asked to give their opinions</td><td>Themes and categories
Facilitators
Documentation by personal diary
Relief of diagnosis, lifting burden from women's minds about their condition
Making lifestyle changes/self-help
Setting goals and being in control of own management of symptoms and treatment</td><td>Aims
Clearly reported. Aim of the study clearly reported, research method was appropriate for answering the question.

Sample selection
Sample selection was reported adequately. The relationship between the researcher and participants was reported.

Data collection</td></tr>
</table>

Study details	Participants	Methods	Findings/results	Limitations
To identify the information needs of women facing laparoscopy for endometriosis. **Study type** Qualitative study. **Study dates** 2000 **Source of funding** Department of Health and Aged Care	<table><tr><td>45-49</td><td>6</td></tr><tr><td>50-54</td><td>2</td></tr><tr><td>55-59</td><td>0</td></tr><tr><td>60-64</td><td>1</td></tr></table> **Inclusion criteria** • Women diagnosed with endometriosis through the Endometriosis Association (VIC) Inc. **Exclusion criteria** Not reported	regarding what information hey would like to receive or contribute about endometriosis including 1. the nature of the disease, 2.their experience living with endometriosis and 3. their experience with diagnosis and treatment. • all the focus groups were audio taped and were taken note by the study leader. **Data analysis** • Thematic analysis • Themes were identified and then checked to be sure that they had emerged from the data. • The data analysis was given to the other members of the study team who had attended the focus group. they could comment and they were sent to participants for validation.	**Barriers** • Delayed diagnosis • Trivialisation of symptoms (by doctor) • Lack of knowledge of health care professional about endometriosis • Refusal by doctor to refer to specialist/gynaecologist • going to see a number of doctors prior to one who would understand women's symptoms • Lack of understanding by family of symptoms • Breakdown of marriage/breakup with partner • Disruption of social activities/work and education • Fear of not being able to cope	Data collection relied on women's contribution to the focus groups in person or by telephone, no discussion on whether saturation was reached for any of the themes reported. **Data analysis** The analytical process was described, and description of how themes were identified were reported. The researchers did not critically review their own roles in the process. **Findings/results**: Results were presented clearly (e.g., citation/data and the researchers' own input distinguished; the researchers' role and potential influences in the analytical process not critically reviewed. **Overall quality** Low **Other information** None
Full citation	**Sample size** N=61	**Setting** Not reported	**Themes and categories** **Facilitators**	**Aims:**

Study details	Participants	Methods	Findings/results	Limitations
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 **Ref Id** 402175 **Aim(s)** The aim was to describe aspects of a study that was conducted to determine women's needs for information related to laparoscopy for endometriosis, to develop, implement and review an information pathway, which describes the process and content of care for this consumer group; and to develop and evaluate an integrated information delivery strategy targeted to this consumer group. **Study type**	**Characteristics** Age (years, n): 20-24: 5 25-29: 10 30-34: 19 35-39: 9 40-44: 9 45-49: 6 50-54: 2 55-59: 0 60-64: 1 **Inclusion criteria** • Women who were diagnosed with endometriosis attending focus groups (face-to-face) or telephone discussions. **Exclusion criteria** Not reported.	**Data collection** • A survey was mailed to women diagnosed with endometriosis and those women who responded (65%) attended focus groups or were interviewed by telephone. Focus group discussions were audiotaped and transcribed for analysis. **Data analysis** • Thematic analysis was undertaken.	• Personal diary; • self-help/lifestyle changes; • benefit of diagnosis **Barriers** • Delayed diagnosis at medical level; • unnecessary diagnostic investigations;	Clearly reported. The aim was clearly reported, research method was appropriate for answering the research question. **Sample selection** Sample selection was not clearly reported. The relationship between the researcher and the selected sample was not clearly reported. **Data collection** The data collection procedure was not clearly described and according to a theoretical framework **Data analysis** A thematic approach was used for data analysis by the project leader, but there was no indication of saturation of themes. **Findings/results** Results were presented as the researchers own input, and the researcher's role and potential influences in the analytical process were not critically reviewed.

Study details	Participants	Methods	Findings/results	Limitations
Qualitative study. **Study dates** 2003 **Source of funding** Department of Health and Aged Care as part of the Consumer and Provider Partnerships in Health.				**Overall quality** Low **Other information** None
Full citation 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 **Ref Id** 553545 **Aim(s)** To explore the impact of endometriosis on couples and to contribute to improving the wellbeing of people living with	**Sample size** N= 22 women with endometriosis and their partners **Characteristics** • Mean Age: 34.8 years. Age range: 25 - 50 years (women) • Mean Age: 36.3 years. Age range: 26 - 57 years (men) • Country: United Kingdom • length of time since onset of symptoms = 13.6 years (range: 2-37 years) • average length of time since diagnosis = 4.5 years (range: 1 month-20 years) **Inclusion criteria** • heterosexual couples • who were living together • in which the female partner had received a diagnosis of endometriosis following laparoscopy	**Setting** UK. Sample was recruited from support groups, hospital clinics and word of mouth **Data collection** • Face to face, semi-structured, in-depth interviews • Men and women were interviewed separately **Data analysis** • A thematic approach was applied to the analysis • The interview data were then analysed dyadic ally (taking each couple as a 'unit of analysis' and exploring similarities and differences in partners' accounts).	**Themes and categories** **Facilitators** • Supportive partner • Supportive workplace • "Being aware of the range of ways that endometriosis can affect a partner is likely to increase understanding, care and support within relationships • "Consultations should be on women, partners and the couple relationship" • "Healthcare practitioners should ask both women and partners how endometriosis is affecting them and how it is affecting the couple relationship" • "As endometriosis treatments often act as a contraceptive or create	**Aims** Aim of study clearly reported, research method was appropriate for answering the research question. **Sample selection** Sample selection was reported. The sample was recruited by many sources but was selected opportunistically. The relationship between the researcher and the respondents was not reported. **Data collection** Data was collected through interviews conducted by the researcher. Some discussion around identification of themes

Study details	Participants	Methods	Findings/results	Limitations
endometriosis by providing an evidence base for improving couple support. **Study type** Qualitative study (Scientific report – not peer-reviewed) **Study dates** Not reported Source of funding UK Economic and Social Research Council	• and had experienced symptoms for at least one years **Exclusion criteria** • gay couples and couples living apart		risks to fertility, some couples had to make a difficult choice to either accept treatment and reduce pain, or reject treatment to try to conceive" **Barriers** • Delayed diagnosis • Lack of understanding of health care professional; trivialisation of symptoms • Numerous operations and recurring symptoms • Impact on partners • Disruption of social relationships • Disruption of workplace performance	was discussed but there was no discussion on data saturation. **Data analysis** The analytical process was described in detail. The researchers did not critically review their own roles in the process. **Findings/results** Results were presented clearly **Overall quality** Low **Other information** Amongst the women, 14 were White British, six were South Asian and two identified themselves as coming from 'other' ethnic backgrounds. Amongst the men, 13 were White British, six were South Asian and three identified themselves as coming from 'other' ethnic backgrounds.
Full citation Denny, E., Women's experience of endometriosis, Journal	**Sample size** 15 women **Characteristics** Not reported.	**Setting** Self-help group, hospital setting. **Data collection**	**Themes and categories** **Facilitators** • Supportive partner	**Aims** Clearly reported. Aim of study clearly reported, research method was

Study details	Participants	Methods	Findings/results	Limitations
of Advanced Nursing, 46, 641-8, 2004 **Ref Id** 402889 **Aim(s)** To explore women's experiences of living with endometriosis. **Study type** Qualitative study. **Study dates** August 2001 and December 2002. **Source of funding** Not reported.	**Inclusion criteria** • Women with a confirmed diagnosis of endometriosis following laparascopic investigation. **Exclusion criteria**	• Data were collected through interviews in women's homes or in mutually convenient locations, such as participant's workplace. **Data analysis** • A thematic approach was applied to the analysis as in vivo quotations were collated and organised by categorising women's stories using the previously identified key areas.	• Supportive workplace • Improved health and reduction of symptoms after surgery (hysterectomy) **Barriers** • Delayed diagnosis • Lack of understanding of health care professional; trivialisation of symptoms • Numerous operations and recurring symptoms • Impact on partners • Disruption of social relationships • Disruption of workplace performance	appropriate for answering the research question. **Sample selection** Sample selection was not clearly reported; the relationship between the researcher and the respondents was not clearly reported. **Data collection** Data collection was not clearly reported, and there was no discussion on whether saturation had been reached for any of the themes reported. **Data analysis** The analytical process was reported but not in detail. The researchers did not critically review their own roles in the process. **Findings/results** Results were presented clearly (e.g. citation/data and the researchers' own input distinguished. The researchers' roles and potential influences in the analytical process not critically reviewed).

<table>
<tr><th>Study details</th><th>Participants</th><th>Methods</th><th>Findings/results</th><th>Limitations</th></tr>
<tr><td></td><td></td><td></td><td></td><td>Overall quality
Low

Other information
None</td></tr>
<tr><td>Full citation
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
Ref Id
403172

Aim(s):
The study assessed the impact of deep dyspareunia had on the quality of life in women with endometriosis.

Study type
Qualitative study

Study dates
Published 2007

Source of funding</td><td>Sample size
30 women

Characteristics<table><tr><th>Characteristic</th><th>Value</th></tr><tr><td>Average age in years (range)</td><td>31 (19–44)</td></tr><tr><td>Social class 1–3</td><td>27</td></tr><tr><td>Social class 4–5</td><td>3</td></tr><tr><td>Married/cohabiting</td><td>20</td></tr><tr><td>Single</td><td>10</td></tr><tr><td>Heterosexual</td><td>30</td></tr><tr><td>Women with children (n)</td><td>11 (plus 2 pregnant at interview)</td></tr><tr><td>Parity (range) 1–3</td><td>1-3</td></tr><tr><td>White British</td><td>27</td></tr><tr><td>Afro-Caribbean British</td><td>1</td></tr><tr><td>Indo-Caribbean</td><td>1</td></tr><tr><td>South American Indian</td><td>1</td></tr><tr><td>Average time from symptoms to 5.65 (1–18) diagnosis in years (range)</td><td>5.65 (1–18)</td></tr></table></td><td>Setting
Endometriosis outpatient clinic

Data collection
• A story-telling approach was used and Semi-structured interviews took place.
• All the interviews were taped-recorded with the permission of the participants.
• Follow-up questions were asked from women with painful sexual intercourse by the researcher expanded on the issues raised by participants, and introduced the concept of dyspareunia to those women who had not mentioned it originally.
• The transcript of the interview were sent to women and they were asked to confirm its veracity.</td><td>Themes and categories
Facilitator
• Supportive partners

Barriers
• Dyspareunia difficult to cope with, low self-esteem, feeling unfeminine and unattractive
• Relationships with partners strained
• Women feeling that partners may leave them</td><td>Aims
Clearly reported. Aims of the study clearly reported. Research method was adequate for answering the research question.

Sample selection
Sample selection was clearly reported, however, the relationship between the researcher and the respondents were not clearly reported.

Data collection
Data collected from women relied on a story-telling approach, there was some indication on saturation, and that recruitment was suspended when no new themes emerged from additional data collected.

Data analysis
The analytical process was described and how themes</td></tr>
</table>

Study details	Participants	Methods	Findings/results	Limitations
Birmingham Women's Hospital	**Inclusion criteria** • Laparoscopically diagnosed endometriosis **Exclusion criteria** • No laparoscopically diagnosed endometriosis	**Data analysis** • Narrative analysis • Thematic analysis • Rigour in the analytical process was achieved by both authors independently analysing the data and agreeing the emergent themes. • Rigour was increased by the involvement of the women in the sample in confirming the veracity of data from their own interview, and agreeing the relevance of themes.		were identified. Researchers did not critically review their own roles in the process. **Findings/results:** Results were presented clearly (e.g., citation/data and the researchers' own input distinguished; the researchers' roles and potential influences in the analytical process were not critically reviewed) **Overall quality:** Moderate **Other information** None
Full citation 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 **Ref Id** 415551 **Aim(s):**	**Sample size** 30 women **Characteristics** • Married (n): 23 • White British (n):27 • Afro-Caribbean British (n):1 • Indo-Caribbean (n):1 • South American Indian (n): 1 • Average time from experiencing symptoms to diagnosis (years): 5.65 (range <1 year to 18 years)	**Setting** The sample was recruited from a dedicated endometriosis clinic in a specialist women's hospital in the UK. **Data collection** • Data was collected through interviews with an open-ended invitation for women to answer a few simple questions about their experiences of living with endometriosis.	**Themes and categories** **Facilitators** • Diagnosis of endometriosis • Confirmation of pain visually on photographs/or visual image of endometriosis • Keeping a diary • Hope that laparoscopy would stop pain/symptoms of endometriosis	**Aims:** Clearly reported. Aim of the study clearly reported, research method was appropriate for answering the research question. **Sample selection** Sampel selection was reported. The relationship between the researcher and participants was clearly reported. **Data collection**

Study details	Participants	Methods	Findings/results	Limitations
To explore women's experiences of living with endometriosis. **Study type** Qualitative study **Study dates** Published 2009 **Source of funding** Birmingham Women's Hospital	**Inclusion criteria** • Women with endometriosis diagnosed by laparoscopy. **Exclusion criteria**	**Data analysis** • A story telling /thematic approach was applied to the analysis to enable women to have some control over the form and content of the interviews and communicate the complexities of their lives, while also enabling them to set parameters around what they were prepared to reveal.	• Realisation that surgery could make symptoms get better or worse • Having control of their symptoms, planning around 'bad days' of pain • Hope and faith in the medical system even with uncertainty about the future **Barriers** • Delay in diagnosis • Uncertainty about course of condition • Doctor's lack of sympathy and not understanding women's symptoms • Referral to a number of specialists before being referred to a gynaecologist • Numerous laparoscopies to manage symptoms • Staging: severity of pain not equating to extent of disease • Uncertainty of fertility	Data collection relied on interviews and by women's diaries which they were asked to keep. **Data analysis** The analytical process was described in detail, as well as description of how themes were identified. **Findings/results:** Results were reported clearly (e.g., citation/data and the researchers own input distinguished. The researchers roles and potential influences in the analytical process not critically reviewed). **Overall quality** Moderate **Other information** None
Full citation Fernandez, I., Reid, C., Dziurawiec, S., Living with endometriosis: the perspective of male partners, Journal of Psychosomatic	**Sample size** 16 male partners of women with endometriosis. **Characteristics** • **Age:** ranged from 24 to 67 years (mean age 40.6 years, SD 13.42).	**Setting** Not reported. **Data collection** • Data were collected by survey covering topics that were previously completed	**Themes and categories** **Facilitators** • Experience of their partners with endometriosis made couples stronger/closer	**Aims** Clearly reported. Aim of the study clearly reported, research method was appropriate for answering the research question.

Study details	Participants	Methods	Findings/results	Limitations
Research, 61, 433-8, 2006 **Ref Id** 403213 **Aim(s):** To explore the experiences of partners of women with endometriosis. **Study type** Qualitative study. **Study dates** Published 2006 **Source of funding** Not reported.	• Duration of relationship (mean years, SD): 11.5 (8.9). **Inclusion criteria** • Male partners involved in a relationship at the time of participation. **Exclusion criteria**	by their spouse. A forced-choice response method was used to improve response rate through minimising the time necessary to complete the survey. • The survey was distributed via post. Those who completed the survey were further invited to participate in a follow-up interview (by phone or e-mail) for 10-15 minutes. **Data analysis** • A thematic approach was applied to the analysis as in vivo quotations were collated and organised by common themes.	• Partners of women with endometriosis acknowledged that their spouse was resilient and were not letting endometriosis rule their lives **Barriers** • Shock and denial, and not knowing about endometriosis • Grief-like emotional impact when partners tell them of the diagnosis • Negativity towards the health care professional • Issues of fertility and hysterectomy • Powerlessness and not knowing how to help partners • Limited control of decision making related to management of endometriosis	**Sample selection** How the study sample was selected was reported. The relationship between the researcher and the respondents was not clearly reported. **Data collection** Data collection relied on the answers the partners responded to in the survey. No discussion on whether saturation had been reached for any of the themes reported. **Data analysis** The analytical process was not clearly described in detail, no description of how themes were identified; researchers did not critically review their own roles in the process. **Findings/results** Results were presented clearly (e.g., citation/data and the researchers' own input distinguished; the researchers roles and potential influences in the analytical process not critically reviewed).

Study details	Participants	Methods	Findings/results	Limitations
				Overall quality Low **Other information** None
Full citation 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 **Ref Id** 415554 4 **Aim(s)** To explore women's perceptions of living with endometriosis. **Study type** Qualitative study. **Study dates** Published 2008 **Source of funding** Not reported	**Sample size** 18 women **Characteristics** • Aged from 16 to 45 • Many of the women were educated at a tertiary level • All apart from the 16 year old, were currently, or had been, in paid employment **Inclusion criteria** • Women with endometriosis **Exclusion criteria** Not reported.	**Setting** New Zealand **Data collection** • The taped and transcribed interviews took an unstructured, interactive format commencing with the broad question: 'what impact has endometriosis had on your life?' **Data analysis** • A thematic approach was used to analyse the interview data. • The analytic process involves a process of reading and rereading texts, comparison of texts, grouping connected extracts and developing the groupings into themes. • The next step involved establishing the validity or 'trustworthiness' of the research data in representing the participants' stories. • The emerging themes were presented at two	**Themes and categories** **Facilitators** • Making nutritional changes, exercise, massage, meditation, behaviour changes to avoid fatigue, acupuncture, Chinese herbal treatments • Information from doctor • Support groups • Information provided by other women • Information from guest speakers, books, internet, chat rooms **Barriers** • Lack of formal diagnosis of endometriosis • Disruption to education, social relationships, barrier to full time employment • Pain and fatigue • Depressed, moody, angry, and irritable lacking enthusiasm • Non-provision of nurses	**Aims** Clearly reported. Aim of the study clearly reported, research method was appropriate for answering the research question. **Sample selection** Sample selection was clearly reported. The relationship between the researchers and participants was not clearly reported. **Data collection** Data was collected by taped and transcribed interviews. Interviews were unstructured, and there was no discussion on saturation of data. **Data analysis** A thematic approach was used to analyse the interview data. The analytical process was described in detail, and how the themes were identified. Researchers

Study details	Participants	Methods	Findings/results	Limitations
		endometriosis support group meetings. Participants in the group concurred that the research findings fitted with their experiences.	• Need for improved health care professional on preparation of surgery • Need for input from nurses on treatment benefits and harms to enable decision making	did not critically review their own roles in the process **Findings/results** Results were presented clearly (e.g., citation/data and the researchers' roles and potential influences in the analytical process not critically reviewed). **Overall quality** Low **Other information** None
Full citation 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 **Ref Id** 401465 **Aim(s):** To explore and describe the impact of	**Sample size** 24 women **Characteristics** • The mean age of the sample was 32.5 years (SD = 5.8, 21.5- 44). • 12 women were married, 3 were separated, 2 were co-habiting, 4 were in long-term relationships and 3 were single. • 14 were nulliparous. • 14 (58.3%) women were diagnosed with minimal to mild endometriosis, 8 (33.3%) with moderate to severe endometriosis and 2 (8.3%) with deeply infiltrating nodules.	**Setting** Gynecology outpatient clinic at the Women's Centre, John Radcliffe Hospital, Oxford **Data collection** • Twenty-four individual interviews were conducted. The interviews were in-depth and followed a semi-structured format. • Prompt questions concerning areas of HRQoL which may have been adversely affected by endometriosis were pre-prepared.	**Themes and categories** **Barriers** • delayed or incorrect diagnosis • lack of knowledge of HCP • trivialisation of symptoms by HCP, told that it is normal so have to cope with it • feeling frustrated that HCP did not do anything to help manage pain • negative feeling on physical appearance (feeling bloated, feeling unwell, weight gain)	**Aims** Clearly reported. Aim of the study clearly reported, research method was appropriate for answering the research question. **Sample selection** Sample selection was reported clearly. The relationship between the researcher and participants was not clearly reported. **Data collection** Data collection relied on in depth interviews in a semi structured format.

Study details	Participants	Methods	Findings/results	Limitations
endometriosis upon quality of life. **Study type** Qualitative study. **Study dates** Published 2004 **Source of funding** Pharmacia Corporation	**Inclusion criteria** • A laparoscopic diagnosis of endometriosis **Exclusion criteria** • Any woman without a laparoscopic diagnosis of endometriosis was excluded.	• All the interviews were tape-recorded, transcribed verbatim and ranged between 25 min and 2 h (mean = 55 min) in duration. **Data analysis** • The framework that was used for analyzing the qualitative interviews was grounded theory. • Starting with the first interview, the transcript was coded using 'open coding' which helped identify the concepts and enabled the categories of HRQoL affected by endometriosis to emerge. • On the basis of the emerging concepts and categories, a theoretical sampling technique was adopted. • After conducting 24 interviews 'theoretical saturation' of the data was reached. • From this analysis, 86 concepts were identified from the interviews. The 86 concepts were placed in 15 descriptive categories which are described below.	• negative impact on physical activity (walking, standing, sitting, exercising)/unable to carry out daily activities • disruption to social activities (not being able to attend social events, worry about pain starting in public, lack of energy) • powerlessness • emotional wellbeing (not being able to cope with pain, being moody and having short temper and taking it out on family, friends or children) • dyspareunia • employment • worry about infertility • trying to cope with over the counter drugs to manage pain • discontinuation of prescription drugs /further surgery due to side effects	**Data analysis** The analytical process was described in detail. To reduce interviewer bias, a research nurse went through some of the transcripts. **Findings/results:** Results were presented clearly (e.g., citation/data and the researchers own input distinguished; interviewer bias (research nurse went through some of the transcripts) **Overall quality** Moderate **Other information** None
Full citation	**Sample size**	**Setting**	**Themes and categories**	**Aims**

Study details	Participants	Methods	Findings/results	Limitations
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 **Ref Id** 403416 **Aim(s):** To understand the relationship between socio-demographic background and health related phenomena between women with endometriosis. **Study type** Qualitative study **Study dates** Published 2008 **Source of funding** Australian Research Council Victorian Department of Innovation, Industry	30 women **Characteristics** **Sociodemographic profile of women** Age, years (n): 20-29 years: 4 30-39 years:7 40-49 years:12 50-59 years: 3 60+ years:4 Country of birth (n): Australia: 25 Overseas:5 Occupation (n): Managers/professionals/associate professionals: 16 Clerical: 4 No occupation:10 Marital status (n): Married: 19 Separated/divorced:5 Single/never married:6 **Inclusion criteria** • Women with endometriosis **Exclusion criteria** Not reported	Women with endometriosis were invited to participate in the study in Victoria who were recruited as part of a larger study. **Data collection** • Data was collected by in depth interviews lasting for approximately 60 minutes, conducted at a woman's home or other place of choice. • A **story telling approach** was taken to gather data, and were conducted concurrently, allowing for the refinement of interview guidelines and cessation of further recruitment upon achieving data saturation. **Data analysis** • A **grounded-theory approach** was applied in the analysis of the narratives, an iterative process in which all authors read the transcripts and developed a coding book. Themes were identified by careful reading of the interview data, but also searching from themes identified in prior research in the area of	**Facilitators** • Women recalling some support from teachers at school being helpful • Few mothers concerned about daughter's painful periods and were encouraged by them to see the general practitioner • Women with severe pain due to dyspareunia seek medical advice • Seeing a doctor who was sympathetic to women's symptoms resulted in OC to reduce pain and gave women 'control over their body' even though the diagnosis had not been made • Symptoms resolving after hysterectomy • Diary keeping was positive approach • Persistence of some women to be referred to a specialist • Diagnosis • Reading about the condition • Seeking alternative information about managing pain by themselves	Clearly reported. Aim of the study clearly reported, research method was appropriate for answering the research question. **Sample selection** Sample selection was reported. Relationship between researcher and participants not clearly reported. **Data collection** Data collection relied on story telling by women until data saturation of themes was achieved. **Data analysis** The analytical process was described in detail, and how the authors identified the themes. Researchers did not critically review their own roles in this process. **Findings/results:** Results were presented clearly (e.g., citation/data and the researchers' own input distinguished; the researchers' roles and potential influences in the

Study details	Participants	Methods	Findings/results	Limitations
and Regional Development Monash University University of Melbourne		women's reproductive health. • Themes were included only if a significant number of women (50%) spoke about them. • Narratives of illness were explored (interrelationship of themes and how they led to emerging patterns in illness narratives: endurance and contest.	• Taking control and making decisions about further treatment/surgery • Changes in lifestyle (information from article in newspaper) to manage pain **Barriers** • Women believed that symptoms were normal, from experiences of relatives or friends • Not given information or opportunity to discuss period pain or other discomfort at school, or no discussion by teachers about their pain or any advice on obtaining professional help from the doctor • Doctors trivialise women's symptoms and lack of recognition from doctor • "shopping around" for a doctor would would provide medication for relief of symptoms or referral to specialist • Numerous laparoscopies before formal diagnosis of endometriosis • Relationship breakdown after diagnosis • Uncertainty about fertility (e.g., lack of information	analytical process were not critically reviewed. **Overall quality** Moderate Other information None

Study details	Participants	Methods	Findings/results	Limitations
			about timing of conception)	
Full citation 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 **Ref Id** 402321 **Aim(s)** To understand how bloggers present information sources and make cases for and against the authority of those sources. **Study type** Discourse analysis. **Study dates** Published 2011. **Source of funding** Not reported.	**Sample size** 11 blogs were selected. **Characteristics** Blogs varied in the number, length of posts, scope and content. Some were very broad, describing endometriosis symptoms and treatments and personal and family happenings. Others were more focused on the illness. There was also substantial variation in the kinds of things happening in bloggers' lives during the data collection period. **Inclusion criteria** Blogs which are authored by women living with endometriosis and focused exclusively or primarily on their authors' experiences of endometriosis. **Exclusion criteria** • Bloggers who incorporated experience with multiple chronic illnesses • Bloggers with endometriosis who mainly posted about infertility	**Setting** **Data collection** • Beginning with one prominent chronic illness blog, successive links were searched until all known endometriosis blogs had been identified. • Posts from each blog for the same 2-month period were captured. • The data set consisted of 87 posts, comprising nearly 27,500 words. **Data analysis** • Potter's discourse analytic approach was used to analyze how bloggers described, supported, or challenged the authority of information sources. • First, each author read the entire corpus and individually identified instances in which the bloggers discussed information sources. • Next, the authors individually analyzed the rhetorical strategies that bloggers used to present	**Themes and categories** **Facilitators** Blogs by other women with endometriosis share their experience with other women	**Aims** Clearly reported. Aim of the study clearly reported, research method was appropriate for answering the research question. **Sample selection** Not applicable **Data collection** Not applicable **Data analysis** The analysis was clearly reported. **Findings/results** The results were presented clearly (e.g., citation/data and the researchers' own input distinguished). **Overall quality** Moderate **Other information** None

Study details	Participants	Methods	Findings/results	Limitations
		or challenge the authority of information sources. • They met regularly to compare their individual analyses, to look for confirming and disconfirming examples, and to analyze the functions performed by bloggers' accounts until they had identified and agreed on the major techniques.		
Full citation Seear, Kate, The third shift: Health, work and expertise among women with endometriosis, Health Sociology Review, 18, 194-206, 2009 **Ref Id** 415706 **Aim(s)** To explore the experiences of women living with chronic and incurable endometriosis, and how women become experts in their own care and ramifications of these processes for women. **Study type**	**Sample size** 20 women **Characteristics** • Women were mainly Anglo-Celtic, aged between 24 and 55 years (mean age 34 years) • Average length of diagnostic delay: 9 years. • 9 women were married, one woman was in a same-sex relationship, 10 women were either single or partnered. • 5 women had children, one was pregnant with her first child. • 4 women had undergone hysterectomy. • 15 women had tertiary education, and several worked in allied health and medical areas (e.g., trained scientist, medical secretary, nurse, psychotherapist)	**Setting** • Unclear setting. Women were recruited by snowball sampling (information about the study was was passed on to potential participants via friends, family and colleagues and potentially interested participants were invited to contact the author). • An advertisement was also placed in the newsletter of an Australian support group for sufferers, inviting them to contact the author if interested in the study. **Data collection** • Data was collected through semi-structured interviews, with questions exploring diagnosis, treatment, doctor-patient relationship,	**Themes and categories** **Facilitators** • Joining support groups • Searching the internet and reading about the condition • Acquiring technical knowledge of the condition, drug therapies, natural therapies and management options • Changes in lifestyle • Becoming an expert patient **Barriers** • Shock of diagnosis • Internet searching bringing up overwhelming information that was complex, conflicting and confusing.	**Aims** Clearly reported. Aim of the study clearly reported, research method was appropriate for answering the research question. **Sample selection** Sample selection was reported. The relationship between the researcher and respondents was not clearly reported. **Data collection** Data collection was reported. **Data analysis** The analytical process was not described fully. Researchers did not

Study details	Participants	Methods	Findings/results	Limitations
Qualitative study. **Study dates** Published 2009 **Source of funding** Not reported.	**Inclusion criteria** • Women diagnosed with endometriosis. **Exclusion criteria**	self-help, causation and reflections on the illness experience. **Data analysis** • A thematic approach was applied to the analysis: data was organised into major themes and concepts. After identification, data was checked to ensure they were supported by the data.	• Being knowledgeable about endometriosis did not reduce the level of anxiety • Giving up full time work to manage their condition	critically review their own roles in the process. **Findings/results** Results were presented clearly (e.g., citation/data and the researchers' own input distinguished; the researchers' roles and potential influences in the analytical process not critically reviewed. **Overall quality** Moderate **Other information** None
Full citation Shoebotham, A., Coulson, N. S., Therapeutic Affordances of Online Support Group Use in Women With Endometriosis, Journal of Medical Internet Research, 18, e109, 2016 **Ref Id** 496837 **Aim(s)** To examine the presence of therapeutic	**Sample size** N=69 women Of the overall sample, 66 (95.7%) women had received a confirmed diagnosis of endometriosis **Characteristics** • Mean **Age:** 34.2 years. Age range: 19 - 50 years • Country: ○ United Kingdom (65.2% 45/69) ○ United States (21.7% 15/69). • Mean time since diagnosis = 4 years, 1 month (range: between 1 month and 20 years before survey completion)	**Setting** ○ The recruitment happened on 3 online support groups, more than half of respondents (62.3% 43/69) were recruited from 1 group, the one hosted by Facebook **Data collection** • Web-based survey with open-ended questions: - 1. a series of short answer questions relating to their background and use of online support groups	**Themes and categories** **Facilitators** • connection, that is, the ability to connect in order to support each other, exchange advice, and to try to overcome feelings of loneliness;" • exploration," that is, the ability to look for information, learn, and bolster their knowledge"; • narration," that is, the ability to share their experiences, as well as read about the experiences of others;"	**Aims** Aim of the study was clearly reported, research method was appropriate for answering the research question. **Sample selection** Sample selection was self-selected. The relationship between the researcher and the respondents was not clearly reported. **Data collection** Data collection was clearly reported.

Study details	Participants	Methods	Findings/results	Limitations
affordances as perceived by women who use endometriosis online support groups **Study type** Qualitative study. **Study dates** June to July 2015 **Source of funding** Not reported	• Participants had been using online support groups for endometriosis for between 1 month and 14 years, 9 months (mean use period = 2 years, 4 months) **Inclusion criteria** women (aged 16 years or older) who use online support groups for endometriosis **Exclusion criteria** Not reported	- 2. open-ended questions that explored their motives and experiences of using online support groups and whether their use has any effect on how they cope with or manage the condition. **Data analysis** • the responses to the open-ended questions were qualitatively analysed using deductive-inductive semantic thematic analysis • QSR's NVivo 10 software was used to maintain an audit trail • an independent researcher read through some of the transcripts and agreement was reached on the final themes.	• "self-presentation," that is, the ability to manage how they present themselves online. The associated outcomes of use were predominantly positive, such as reassurance and improved coping" **Barriers** • concerns about the accuracy of information • arguments between members • overreliance on the group • becoming upset by negative experiences or good news items • confidentiality of personal information.	**Data analysis** The analytical process was described in detail. There was description of how themes were identified, researchers did critically review their roles in the process. **Findings/results** Results were presented clearly **Overall quality** Moderate **Other information** None
Full citation Strzempko Butt, F., Chesla, C., Relational patterns of couples living with chronic pelvic pain from endometriosis, Qualitative Health Research, 17, 571-85, 2007 **Ref Id** 415663	**Sample size** 13 women in a partnered or marital relationship. **Characteristics** • Partners: male • Length of time couples had lived together ranged from 1 to 23 years (mean=6 years) • All participants were childless except for two couples	**Setting** • Public and private treatment providers and clinics, as well as endometriosis support and informational groups. **Data collection** • Data was collected through responses of participants to informal flyers via telephone who were	**Themes and categories** **Facilitators** • Self help, lifestyle changes **Barriers** • Partner not understanding condition • Worries about fertility • Psychosexual problems/ dyspareunia	**Aims** Clearly reported. Aim of the study clearly reported, research method was appropriate for answering the research question **Sample selection** Sample selection was reported clearly and how women with endometriosis

Study details	Participants	Methods	Findings/results	Limitations
Aim(s) To investigate responses in the couple's relationship to living with chronic pelvic pain from endometriosis. **Study type** Qualitative study. **Study dates** Published 2007 **Source of funding** National Institute of Nursing Research American Legion Auxillary Award UCSF Graduate Student Research award UCSF School of Nursing Century Club award 2002 Sigma Theta Tau Research award	• Age range of women was 23 to 48 years (sample mean=34 years) • Age range of partners was 24 to 50 years (sample mean=38) • 92% women were in paid employment • 84% of partners were in paid employment • 85% of partners had health insurance • 60 % of both men and women were European American, remainder were Hispanic, Asian, Pacific Islander, multiracial or other. **Inclusion criteria** • English-speaking women who had received a diagnosis of endometriosis and experienced pelvic pain for at least 6 months. • At least 18 years of age and living with their intimate partner for at least one year. **Exclusion criteria**	interested in participating. Individual interviews were conducted with each participant followed by a conjoint interview approximately 4 weeks later. • Data comprised of 39 in depth interviews, including 13 individual interviews with the women, 13 with their partners and 13 couple interviews. • The decision to stop recruiting was based on theoretical criteria, as considerable about of data had been collected and repetitive patterns and themes were noted. • All interviews lasted up to 2 hours, followed by an interview schedule and were conducted in a conversational manner by the first author. **Data analysis** • The analytical process included thematic analysis across cases to clarify distinctions and similarities until a pattern of meaning or common situation had been identified.	• disruption to social activities, work or education	and their partners were recruited. **Data collection** Data collection was clearly reported. **Data analysis** The analytical process was described in detail. **Findings/results** Results were presented clearly (e.g., citation/data and the researchers' own input distinguished. **Overall quality** Moderate **Other information** None
Full citation	**Sample size**	**Setting**	**Themes and categories**	**Aims**

Study details	Participants	Methods	Findings/results	Limitations
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 **Ref Id** 402342 **Aim(s)** To investigate motivations and reflections of participant who had provided epidemiological information, blood samples and access to clinical records and data in a large genetic epidemiological study of endometriosis. **Study type** Qualitative study. **Study dates** Not reported	16 **Characteristics** • 15 females and 1 male, aged between 23 and 58 years. • These individuals were among participants in GBE who had previously expressed interest in participating in further endometriosis research. • Of the 15 female participants, 2 were unaffected family members who had not been diagnosed with endometriosis but had had hysterectomies, 5 had been diagnosed with endometriosis and had had hysterectomies and the remaining 8 had been diagnosed but had not had hysterectomies. • 2 participants (a mother and daughter) came from a family in which the daughter was the only affected family member. 1 participant had been adopted at birth. All other participants came from families with at least 2 affected members. **Inclusion criteria** • A sub-group of the large Australian Genes Behind Endometriosis (GBE) study • Aged 18 years or over **Exclusion criteria** Not reported.	Australia **Data collection** • In keeping with a breadth-maximizing approach to exploratory qualitative research, diversity and heterogeneity in sampling was sought from the participants of the large Australian GBE study. • Semi-structured interviews were conducted via telephone • To explore the experiences of participants in GBE with regard to their recruitment and participation in the research, the perceived benefi ts and disadvantages associated with their research participation, and the perceived impact of their participation upon their understanding of both endometriosis and the concept of complex aetiology. • Interviews were later transcribed verbatim and prepared for analysis. **Data analysis** • Qualitative thematic analysis of the interview	**Facilitators** • Being part of a research study increased women's knowledge about endometriosis • Improved psychological wellbeing • Brought family closer together and being aware of the condition	Clearly reported. Aim of the study was clearly reported, research method was appropriate for answering the research question. **Sample selection** Sample selection was clearly reported. The relationship between the researcher and the respondents was not clearly reported. **Data collection** Data collection was clearly reported. **Data analysis** The analytical process was not described in detail. There was no description of how themes were identified, researchers did not critically review their own roles in the process. **Findings/results** Results were presented clearly (e.g., citation/data and the researchers' own input distinguished; the researchers role and potential influences in the

Study details	Participants	Methods	Findings/results	Limitations
Source of funding University of Queensland.		transcripts between April and August 2003. • While themes were identifi ed from the data according to the direction of questions asked, the researcher, in keeping with a qualitative research approach, took an open-ended approach to the interview.		analytical process was not critically reviewed). **Overall quality** Moderate **Other information** None
Full citation 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 **Ref Id** 402345 **Aim(s)** To investigate women's strategies and views about knowledge surrounding endometriosis. **Study type** Qualitative study. **Study dates**	**Sample size** 24 women **Characteristics** The women who participated in this research were all members of endometriosis patient venues, often driven to them after highly negative experiences with medical treatment. **Inclusion criteria** • Member of endometriosis patient venues **Exclusion criteria**	**Setting** Endometriosis support group in Winnipeg, Canada **Data collection** <u>First stage 1994</u> • 20 hours of focus group meetings with six women recruited from an endometriosis support group • The focus of the sessions was GnRH agonists, o understand how women gathered, evaluated, and used information about a specific element of the endometriosis experience, a medical treatment. <u>Second stage 2000</u> • An open-ended survey on an electronic mailing list for women with endometriosis in different countries	**Themes and categories** **Faciltators** • Health care professional was a starting point to obtain information about endometriosis • Self-education and 'doing homework' by means of internet searching, WITSENDO list, Endometriosis Association, books for lay audience, medical publication, local support/patient group and sufferers, drug reference manual, leaflets, videotapes from doctors **Barriers** • Delay in diagnosis • Variation in expert opinion in terms of treatment	**Aims** Clearly reported. Aims of study were not clearly reported, research method was appropriate to answer the research question. **Sample selection** Sample selection was not clearly reported. The relationship between the researcher and respondents was reported. **Data collection** There was no discussion on whether saturation had been reached for any themes reported. **Data analysis** The analytical process was not described in detail, no description of how themes were identified; the

Study details	Participants	Methods	Findings/results	Limitations
1994 **Source of funding** Social Sciences and Humanities Research Council.		• While a few broad questions about their views on endometriosis information were included, they were encouraged to frame their narratives as they saw fit Both focus group transcripts and the electronic responses of survey participants were coded using Atlas TI™. **Data analysis** • The data were searched for knowledge-related keywords, and coded to reflect key themes. • Codes were modified throughout according to the inductive, constant comparative method of grounded theory. • The formal readings for this analysis focused on three elements: • (1) the narrators' presentation of knowledge claims; • (2) the narrators' presentations of themselves and physicians as knowledgeable agents (or not); • (3) the relational aspects of the narrators' accounts,	• Health care professional not taking symptoms seriously • Concerns about side effects of GnRHa treatment (may cause depression, irritability, confusion, anxiety, and memory loss)	researchers did not critically review their own roles in the process. **Findings/results** Results were reported clearly (e.g., citation/data and the researchers' own input distinguished; the researchers role and potential influences in the analytical process were not critically reviewed. **Overall quality** Moderate **Other information** None

Study details	Participants	Methods	Findings/results	Limitations
		focusing on the focus group interaction and the participants' representations of the endometriosis patient community in the survey.		

GBE: Genes behind endometriosis; HCP: Healthcare professional; HRQoL: Health-related quality of life; OC: Oral contraceptive; SD: Standard deviation;

G.5 Review question: Risk of reproductive cancer

- Do women with endometriosis have an increased risk of reproductive cancer and do they need to be monitored or referred accordingly?

Study details	Participants	Diagnosis	Outcomes	Comments
Full citation 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 **Ref Id** 428576 **Country/ies where the study was carried out** Canada **Study dates**	**Sample size** 2854 identified patients. n=2521 women with endometriosis n=292 women with ovarian cancer n=41 women with endometriosis and ovarian cancer Total population size - unclear **Characteristics** The only baseline characteristics provided were the age and type of ovarian cancer. Women with endometriosis: age 40.0 (9.6 SD) Women with ovarian cancer: age 53.8 (11.4 SD) Women with endometriosis and ovarian cancer: age 41.6 (10.9 SD)	**Details** Sherbrooke University Hospital Centre the Centre Informatise de Recherche Evaluative en Services et Soins de Sante system manages all the clinical and pathological data of all residents in the Estrie region of Quebec (300383 individuals). Cancer incidence: ICD coding for	**Results** Adjustment for confounders: age, number of pregnancies, family history of ovarian cancer, race, oral contraceptive use, tubal ligation, hysterectomy, and breast feeding. Increased risk of ovarian cancer in those with endometriosis: RR 1.6 95% CI 1.12-2.09 (adjusted for the above confounders) Women with ovarian cancer and endometriosis: 41/2521 Women with ovarian cancer and no endometriosis: 251/24,693* (the denominator has been taken from SR Kim2014) Census data from 2001 in the Estrie Region: Prevalence 10.7% endometriosis, 0.11% for endometriosis with ovarian cancer. In those with ovarian cancer 14% had endometriosis. Incidence of ovarian cancer was 24%. Types of cancer Type of ovarian cancer: Clear-cell type; EAOC n: 9; EAOC %: 21.95; OC n: 22; OC %: 7.53; P value: 0.0029	**Limitations** Prevalence study critical appraisal Was the sample representative of the target population? Unclear. No baseline characteristics apart from age were given in the paper. Were the study participants recruited in an appropriate way? Yes Was the sample size adequate? Yes Were the study subjects and setting described in detail?

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr>
<td>1997-2006

Source of funding
None described.</td>
<td>p<0.0001 between the groups. After Tukey adjustment:
mean difference (SE) of Age:
EAOC and ENDO: 8.2 (1.6), p<0.0001
EAOC and OC: -5.5 (1.7), p<0.0001
ENDO and OC:-13.8 (0.6), p<0.0001

Inclusion criteria
• Women with endometriosis, ovarian cancer or both, registered between 1997-2006.

Exclusion criteria
None described.</td>
<td>oncology (ICD-O-2)
Endometriosis: International Classification of Diseases ninth edition, clinical modification (ICD-9-CM), 617.00-617.99.

Medical and pathological data were analysed including their reports to confirm the diagnosis.Histology was also obtained.</td>
<td>
<table>
<tr><td>Endometroid</td><td>10</td><td>24.39</td><td>29</td><td>9.93</td><td>0.0070</td></tr>
<tr><td>Mucinous type</td><td>2</td><td>4.88</td><td>6</td><td>2.05</td><td>0.2571</td></tr>
<tr><td>Serous type</td><td>8</td><td>19.51</td><td>130</td><td>44.52</td><td>0.0023</td></tr>
<tr><td>Other types</td><td>15</td><td>36.58</td><td>112</td><td>38.36</td><td>0.8270</td></tr>
</table>
</td>
<td>No baseline characteristics described.
Is the data analysis conducted with sufficient coverage of the identified sample? Yes.
Were objective, standard criteria used for measurement of the condition? Yes ICD codes. ?risk of misclassification bias/ undiagnosed endometriosis.
Was the condition measured reliably? Yes ICD codes, confirmed by medical and pathology reports.
Was there appropriate statistical analysis? No description of how they adjusted for the confounders.
Are all confounding factors/ subgroups/ differences identified and accounted for? No: only age and family history out of the GDG listed confounders.</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td>Additional confounders controlled for: number of pregnancies, race, oral contraceptive use, tubal ligation, hysterectomy and breast feeding.
Were subpopulations identified using objective criteria? No subpopulations were identified.

Other information
None</td></tr>
<tr><td>Full citation
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
Ref Id
428516

Country/ies where the study was carried out</td><td>Sample size
n=22,207 unique national registration numbers with at least one discharge diagnosis of endometriosis between 1969-1983.
n=20,686 women included in the analysis (see below for exclusions)

Characteristics
• Total follow up 216,851 person years.
• Mean follow up of 11.4 years (range 1-21)
• Average age at entry 38.8 (range 12-82)</td><td>Details
Swedish National Board of Health and Welfare register started in 1969 collected information on surgical procedures, hospital department, and up to 8 discharge diagnoses (ICD 8). 60% coverage in 1969 to 85% in 1983.</td><td>Results
Excluded 19,751 person years and 54 cancer cases that occurred during the first year of follow up to reduce selection bias.

Cancers involving gynecologic organs person years and events were truncated at the time of the first recorded gynae operation as it was unclear as to the ovarian status of the women i.e. whether the ovaries were removed at the same time as a hysterectomy.
<table>
<tr><td>Cancer type or site and ICD 7 code</td><td>Observed</td><td>Expected</td><td>Ratio of observed to expected</td><td>95% CI</td></tr>
<tr><td>Cervix (171)</td><td>11</td><td>15.24</td><td>0.72</td><td>0.4-1.3</td></tr>
<tr><td>Endometrium (172)</td><td>12</td><td>10.97</td><td>1.09</td><td>0.6-1.9</td></tr>
</table></td><td>Limitations
<u>Prevalence study critical appraisal</u>
Was the sample representative of the target population? Unclear. Very limited baseline characteristics given. Population is hospitalized women with endometriosis. Does not include those that have not been hospitalized for endometriosis.
Were the study participants</td></tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr>
<td>Sweden

Study dates
1969-1983

Source of funding
Unclear if financial-supported in part by United States Public Health Service contract N01-CP-85636.</td>
<td>• Average age at cancer diagnosis 52.3 (range 24-82)

Inclusion criteria
• Women diagnosed with endometriosis on the Swedish National Board of Health and Welfare register 1969-1983

Exclusion criteria
• People whose national registration number was not found in the population register/any other register listed as linked to this study (n=809, 3.6%).
• Death during hospital stay (n=181, 0.8%)
• Malignancy before the diagnosis of endometriosis (n=514, 2.4%)
• Record linkage showed incorrect/inconsistent dates (n=17, 0.1%)</td>
<td>Endometriosis ICD code for diagnosis: 625.3
Linkage to national register for population to check individual registration numbers. Record linkage to National Registry of Causes of Death to 1989 ICD 7 classification.
Observation time: time of first endometriosis hospitalization until occurrence of a cancer diagnosis, emigration, death or end of the observation period (Dec 31 1989).
Expected figures: Derived from the entire Swedish population. Done for each calendar ear</td>
<td>
<table>
<tr><td>Uterus not otherwise specified (174)</td><td>1</td><td>1.69</td><td>0.59</td><td>0.0-3.3</td></tr>
<tr><td>Other female genital (176)</td><td>0</td><td>1.25</td><td>0.00</td><td>0.0-2.9</td></tr>
<tr><td>Ovary (183)</td><td>29</td><td>15.11</td><td>1.92</td><td>1.3-2.8</td></tr>
</table>
Total person years for the above cancers: 95,873 (as person years were truncated at time of first gynae operation).

SIR by endometriosis site (Note: was not prespecified in the methods):
<table>
<tr><td rowspan="2">Cancer type or site</td><td colspan="3">Ovary endometriosis (99,092 person yr)</td><td colspan="3">Pelvis endometriosis (21,698 person yr)</td></tr>
<tr><td>Observed</td><td>SIR</td><td>95% CI</td><td>Observed</td><td>SIR</td><td>95% CI</td></tr>
<tr><td>Cervix</td><td>3</td><td>0.48</td><td>0.1-1.4</td><td>4</td><td>1.47</td><td>0.4-3.8</td></tr>
<tr><td>Endometrium</td><td>6</td><td>1.69</td><td>0.6-3.7</td><td>0</td><td>0.00</td><td>0.0-2.7</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Ovary</td><td>17</td><td>3.08</td><td>1.8-4.9</td><td>3</td><td>1.37</td><td>0.3-4.0</td></tr>
</table>
Uterus endometriosis (46,480)
<table>
<tr><td>Observed</td><td>SIR</td><td>95%CI</td></tr>
<tr><td>2</td><td>1.30</td><td>0.2-4.7</td></tr>
<tr><td>2</td><td>0.71</td><td>0.1-2.6</td></tr>
</table>
</td>
<td>recruited in an appropriate way? Yes- National Database. Note: coverage varied from 60-85% of the country's population.
Was the sample size adequate? Yes
Were the study subjects and setting described in detail? Very limited baseline characteristics described.
Is the data analysis conducted with sufficient coverage of the identified sample? 55.6% women had data truncated due to gynae operations as it was unclear if their ovaries were removed or not reducing the at risk population.
Were objective, standard criteria used for measurement of the condition? ICD code- but only one was used. Unclear accuracy of</td>
</tr>
</table>

Study details	Participants	Diagnosis	Outcomes	Comments
		and in a 5 year age group. Method of first diagnosis of endometriosis: laparoscopy 34.9%, laparotomy 54.1%, other 11.0%.	<table><tr><td>0</td><td>0.00</td><td>0.0-1.3</td></tr></table> Results also stratified by follow up year, age on admission, calendar time.	capturing all of those diagnosed with endometriosis. Was the condition measured reliably? Yes ICD codes. Around 90% were by laparoscopy/ laparotomy (visual). No mention of histology samples. Was there appropriate statistical analysis? Yes. Are all confounding factors/ subgroups/ differences identified and accounted for? No: only age and calendar year. Stratified by follow up period and site of endometriosis (not pre-specified in methods).No other confounders were reviewed. Were subpopulations identified using objective criteria? No- location of endometriosis and follow up

Study details	Participants	Diagnosis	Outcomes	Comments
				peroid was presented but not described in the methods. **Other information** Uses some of the same population as Melin 2006 and Melin 2007.
Full citation 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 **Ref Id** 428657 **Country/ies where the study was carried out** USA **Study dates** 1965-1988	**Sample size** n=12,193 women evaluated for infertility between 1965-1988. n=8,429 in the SIR analysis n=8,369 in the RR analysis (excluded were n=2,442 lost to follow up, n=1,319 refused access to medical data, n=3 ovarian cancer diagnosed within 1 year of clinic visit from both analyses and n=60 ovaries removed within 1 year of clinic visit was also excluded from the second analysis) n=1,919 women with endometriosis **Characteristics** Median age of the women at first evaluation: 30 years Nearly 80% are white Median length of follow up was 18.8 years with over 80% followed for 15+ years. **Inclusion criteria**	**Details** Data sources: Clinic records, telephone directories, credit bureaus, postmasters and motor vehicle administration records. Questionnaires sent through linkage with the cancer registries and the National Death Index. Questionnaires (info on health status, lifestyle factors including menstrual, pregnancy, breast feeding history, use of exogenous	**Results** Two analyses: 1 comparing to the US population, 2nd comparing to an infertile population with MVA. N=45 ovarian cancers (21 medical records/cancer registry, 10 death certificates, 14 (31%) self reported) Total follow up 148,318 person years Results are adjusted for age and calendar year. **1st analysis:** against the US population n=13 ovarian cancer events in the endometriosis group n=5.2 expects events SIR (95%CI): 2.48 (1.3-4.2) **2nd analysis:** compared to patients with no evidence of the specified cause of infertility and adjusting for wormen who were not medically evaluated. Adjusted for age at follow up, calendar time, study site, gravidity at entry, causes of infertility no of ovarian cancers in endometriosis patients: n=13 RR (95% CI): 1.26 (0.6-2.6)	**Limitations** <u>Prevalence study critical appraisal</u> Was the sample representative of the target population? Only women who were seeking treatment for infertility. Does not include those with endometriosis who were not seeking infertility treatment. Very limited baseline characteristics given. Were the study participants recruited in an appropriate way? From five large reproductive centres in the US. Was the sample size adequate? Yes

Study details	Participants	Diagnosis	Outcomes	Comments
Source of funding Supported by National Cancer Institute intramural funds.	• Women who sort advice for infertility at 1 of 5 large reproductive endocrinology practices; Boston, New York City, Chicago, Detroit, and San Francisco Bay area between 1965 and 1988. • US address at time of evaluation Seen >1 time or been referred by another physician who provided relevant medial information Primary or secondary infertility **Exclusion criteria** • Those who were evaluated for reversal of tubal ligation	hormones, anthropometric factors, cigarette smoking, alcohol consumption and breast and ovarian disease screening history) were sent out and followed up with a telephone call. N=5,597 responded to the questionnaire. Note 6 self reported ovarian cancers were found to be benign (medical records) and so were excluded. Person years were accrued beginning 1 year after first clinic registration and continuing through the earliest date of cancer diagnosis,		Were the study subjects and setting described in detail? Very limited baseline characteristics described. Is the data analysis conducted with sufficient coverage of the identified sample? 20% were lost to follow up. Were objective, standard criteria used for measurement of the condition? Trained abstractors retrieved the data from medical records, telephone directories, credit bureaus, postmasters, and motor vehicle administration records. Questionnaire. Linkage with registries. Was the condition measured reliably? Unclear how reliable data extraction was and if ICD coding was used. Also unclear

Study details	Participants	Diagnosis	Outcomes	Comments
		death or date last known alive and free of cancer Endometriosis definition: women who had a pelvic laparoscopy, culdoscopy, or laparotomy at which endometriosis was was found. Those categorized as having no endometriosis had one or more of these procedures and did not have endometriosis as a finding.		coverage of the databases. Was there appropriate statistical analysis? Yes. Are all confounding factors/ subgroups/ differences identified and accounted for? No only age and calendar year for population comparison. Age at follow up, calendar time, study site, gravidity at entry, and causes of infertility were controlled for in the secondary analysis. Were subpopulations identified using objective criteria? No- primar y and secondary infertility was explored but not described in the methods. **Other information** 20% was lost to follow up.

Study details	Participants	Diagnosis	Outcomes	Comments
				31% self reported ovarian cancer
Full citation 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 **Ref Id** 403718 **Country/ies where the study was carried out** Denmark **Study dates** 1st January 1978-December 31 1998 **Source of funding** Intramural Research Program of the	**Sample size** See Brinton 2004. **Characteristics** See Brinton 2004. **Inclusion criteria** See Brinton 2004. **Exclusion criteria** See Brinton 2004.	**Details** See Brinton 2004.	**Results** See Brinton 2004. **Additional results:** • N= 39 uterine cancers (only reported overall, no n figures given for women with and without endometriosis). Comparison group is infertile women as described in Brinton 2004. • RR (95% CI): 0.82 (0.3-1.9) • Adjusted for age at follow up, calendar time, study sites, gravidity at entry and all causes of infertility. • It does state that other risk factors e.g. age at first birth, family history of cancer, hysterectomy/ovarian status at follow up, obesity, or use of estrogen replacement therapy, oral contraceptives or ovulation stimulating drugs did not appreciably change risk estimates (no data was given).	**Limitations** Prevalence study critical appraisal Was the sample representative of the target population? Only women who were seeking treatment for infertility. Does not include those with endometriosis who were not seeking infertility treatment. Very limited baseline characteristics given. Were the study participants recruited in an appropriate way? From five large reproductive centres in the US. Was the sample size adequate? Yes Were the study subjects and setting described in detail? Very limited baseline characteristics described.

Study details	Participants	Diagnosis	Outcomes	Comments
NIH, National Cancer Institute.				Is the data analysis conducted with sufficient coverage of the identified sample? 20% were lost to follow up. Were objective, standard criteria used for measurement of the condition? Trained abstractors retrieved the data from medical records, telephone directories, credit bureaus, postmasters, and motor vehicle administration records. Questionnaire. Linkage with registries. Was the condition measured reliably? Unclear how reliable data extraction was and if ICD coding was used. Also unclear coverage of the databases. Was there appropriate statistical analysis? Yes.

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td></td><td>Are all confounding factors/ subgroups/ differences identified and accounted for? Age at follow up, calendar time, study site, gravidity at entry, and causes of infertility were controlled for in the secondary analysis.
Were subpopulations identified using objective criteria? No- primary and secondary infertility was explored but not described in the methods.

Other information
20% lost to follow up.</td></tr>
<tr><td>Full citation
Brinton, L. A., Sakoda, L. C., Sherman, M. E., Frederiksen, K., Kjaer, S. K., Graubard, B. I., Olsen, J. H., Mellemkjaer, L., Relationship of</td><td>Sample size
Ovarian cancer analysis: n=101,912
Borderline ovarian tumor analysis: n= 100,498
Uterine cancer analysis:n= 100,570

Characteristics</td><td>Details
Case group selection: ICD codes (see inclusion criteria).
Control group selection: Two stage sample design.</td><td>Results<table><tr><td></td><td colspan="2">Ovarian cancers</td><td colspan="2">BOT</td><td colspan="2">Uterine cancers</td></tr><tr><td></td><td>n</td><td>RR* (95% CI)</td><td>n</td><td>RR* (95% CI)</td><td>n</td><td>RR* (95% CI)</td></tr></table></td><td>Limitations
<u>Prevalence study critical appraisal</u>
Was the sample representative of the target population? Unclear. No baseline characteristics</td></tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr>
<td>benign gynecologic diseases to subsequent risk of ovarian and uterine tumors, Cancer Epidemiology Biomarkers and Prevention, 14, 2929-2935, 2005
Ref Id
428705

Country/ies where the study was carried out
Netherlands

Study dates
Hospital admissions from 1978-1998 and outpatient visits from 1995-1998.

Source of funding
Intramural Research Program of the NIH, National Cancer Institute.</td>
<td>see table in the following row

Inclusion criteria
• Cases: Women with incident invasive ovarian cancers (ICD for oncology codes183.0, behaviour code 3), borderline ovarian tumours (ICD-O 183.0, behaviour code 1) and uterine cancers (ICD-O 182.0, behaviour code 3) diagnosed between January1 1978 and December 31 1998 among female residents of Denmark who were born after 1936 (Source Danish Cancer Registry)
• Controls: Subgroup of the population, randomly chosen from the Central Population Register.

Exclusion criteria
• Women who were not at risk of developing uterine cancer, invasive ovarian cancers or borderline ovarian tumors at study entry (undergone hysterectomy n=385, bilateral oophorectomy n=41, or diagnosed with uterine n=7 or ovarian n=31 cancer before 1 January 1978) where excluded as appropriate.</td>
<td>1st stage:99,812 women born after 1936 and living in Denmark at study entry (1 January 1978). Random sample based on birth year and the 9th digit of the CPR number, with digit values of 1,2,3 selected for birth years 1937 to 1951, 5 and 6 for birth years 1952-1977 and 7 and 8 for birth years 1978-1991.
2nd stage: Selection of women into the subsample was further narrowed according to the birth years of all the breast, ovarian and endometrial cancers and borderline ovarian tumors</td>
<td>
<table>
<tr><td>No Endometriosis</td><td>2,441</td><td>1.00 (Reference)</td><td>848</td><td>1.00 (Reference)</td><td>1,389</td><td>1.00 (Reference)</td></tr>
<tr><td>Yes Endomtriosis</td><td>50</td><td>1.69 (1.27-22.25)</td><td>12</td><td>1.22 (0.69-2.17)</td><td>9</td><td>1.23 (0.63-2.38)</td></tr>
<tr><td><1y</td><td>5</td><td>3.01 (1.25-7.25)</td><td>5</td><td>7.51 (3.10-18.18)</td><td>5</td><td>13.97 (5.76-33.93)</td></tr>
<tr><td>1-4yrs</td><td>14</td><td>1.95 (1.15-3.31)</td><td>2</td><td>0.75 (0.19-3.01)</td><td>1</td><td>0.71 (0.10-5.07)</td></tr>
<tr><td>≥5 years</td><td>31</td><td>1.49 (1.04-2.14)</td><td>5</td><td>0.77 (0.32-1.86)</td><td>3</td><td>0.54 (0.17-1.68)</td></tr>
</table>
*RR adjusted for calendar time (per 5 years), parity (yes/no), number of births (continuous), and age at first birth (per 5 years) as time dependent variables (with age used as a time metric in the regression models). Additional adjustment for obesity tubal ligation, hysterectomy (for ovarian analysis), unilateral oophorectomy and bilateral oophorectomy (for uterine analysis) did not result in substantial changes it the risk estimates.
The type of ovarian cancer was also recorded: serous (n=932), mucinous (n=344), endometrioid (n=300), germ cell (n=126), clear cell (n=123) and carcinosarcoma (n=19).
Borderline ovarian cancer: serous (n=363) or mucinous (n=391).
Uterine cancer:
a) common indolent types (including adenocarcinoma not otherwise specified, papillary adenocarcinoma,</td>
<td>apart from age and parity were given in the paper.
Were the study participants recruited in an appropriate way? Yes
Was the sample size adequate? Yes
Were the study subjects and setting described in detail? Limited baseline characteristics described.
Is the data analysis conducted with sufficient coverage of the identified sample? Yes.
Were objective, standard criteria used for measurement of the condition? Yes ICD codes. ?risk of misclassification bias/ undiagnosed endometriosis.
Was the condition measured reliably? Yes ICD codes, hospital admissions and discharge diagnoses.</td>
</tr>
</table>

Study details	Participants	Diagnosis	Outcomes	Comments
		diagnosed during the study period. 4 women/case were selected for each birth year between 1937-1951 and 6 women/case between 1952-1991. Record linkage from the cases identified through the Danish Cancer Registry with hospital admissions from 1978-1998 and to outpatient visits from 1995-1998 (Hospital Discharge Register). Each admission record has information on personal ID no. date of admission/outpt visit, date of discharge surgical procedures and up to 20	endometrioid carcinoma, mucinous adenocarcinoma, adenocarcinoma with squamous metaplasia, n=1,178) b) sarcoma, including leiomyosarcoma, endometrial stromal sarcoma, sarcoma not otherwise specified, epithelioid leiomyosarcoma, adenosarcoma, rhabdomyosarcoma, n=137 c) carcinosarcoma, n=19 d) aggressive types including clear cell adenocarcinoma, serous cystadenocarcinoma and papillary serous cystadenocarcinoma, n=18 Tumours not classified into the above categories were excluded (647 ovarian cancers, 106 borderline ovarian tumours, 46 uterine cancers). The number of women with endometriosis is not reported. Kim2014 has reported the proportion of those with ovarian cancer in those with endometriosis and those without endometriosis to be 50/2491 and 1181/99,421 respectively.	Was there appropriate statistical analysis? Unclear weighting system. Are all confounding factors/ subgroups/ differences identified and accounted for? No: only age out of the GDG listed confounders. Additional confounders controlled for: calendar time, parity, no. of births, age at first birth. Additional adjustment for obesity tubal ligation, hysterectomy (for ovarian analysis), unilateral oophrectomy and bilateral oophrectomy (for uterine analysis). Were subpopulations identified using objective criteria? Cancer sub types by ICD codes. Follow up time was split into time

Study details	Participants	Diagnosis	Outcomes	Comments
		discharge diagnoses. Endometriosis (ICD-8, 625.30-625.39; ICD 10 DN80) and uterine leiomyoma were identified. Diagnoses of obesity was also noted. Additional information retrieved: relevant surgical procedures (hysterectomy, bilateral/unilateral oophorectomy and tubal ligation), with the date of surgery defined as the first of the month following the date of admission. Records then linked to CPR to determine the number of children born by each woman. Note:		intervals (not stated in the methods). **Other information** No information given on the total number of women who were diagnosed with endometriosis and unable to calculate. Figures are given in Kim2014 but it is unclear how they were obtained, likely to have been from contacting the authors.

Study details	Participants	Diagnosis	Outcomes	Comments
		CPR has the birth dates of all the children that a woman may have and does not specify if any of them are adopted. If 2 birth dates <10 months, the first child was defined as being adopted in the study. Censoring: diagnosis of a medical condition if diagnosis was before the censoring date. Censoring occurred at death, emigration from Denmark or surgical removal of the uterus/ both ovaries depending on the outcome of interest. Women were followed until cancer diagnosis, any censoring event		

Study details	Participants	Diagnosis	Outcomes	Comments
		or the end of the study. Confounders: calendar time (per 5 years), parity (yes/no), number of births and age at first birth (per 5 years).		

Patient characteristic table for Brinton 2005

Characteristic	Ovarian cancer analysis		Borderline ovarian tumour analysis		Uterine cancer analysis	
	Cases (n=2,391)	Non cases (n=99,421)	Cases (n=860)	Non cases (n=99,638)	Cases (n=1,398)	Non cases (n=99,172)
Birth year						
1937-1941	34.1	30.7	19.8	30.7	47.7	30.7
1942-1946	28.9	29.0	24.9	29.0	33.2	29.0
1947-1951	15.1	17.6	18.1	17.6	12.0	17.6
1952-1956	9.0	12.8	12.5	12.8	5.0	12.8
1957-1961	5.4	5.9	11.2	5.9	1.2	5.9
1962 or later	7.5	4.0	13.5	4.0	0.9	4.0
Parity (%)						
0	22.2	10.8	27.2	10.8	18.4	10.8
1	18.2	16.0	19.1	16.0	17.7	16.0
2	38.3	45.5	33.1	45.5	41.7	45.5
3	16.0	20.8	15.7	20.8	16.1	20.8
≥4	5.3	6.8	4.9	6.8	6.1	6.8
Mean (SD)	1.7 (1.2)	2.0 (1.1)	1.5 (1.2)	2.0 (1.1)	1.8 (1.2)	2.0 (1.1)
Age at first birth (%)						
<20	14.9	15.7	17.4	15.7	14.1	15.6

Study details	Participants		Diagnosis	Outcomes		Comments
20-24	36.5	42.7	34.8	42.7	41.7	42.7
25-29	19.9	22.8	15.4	22.8	19.9	22.8
≥30	6.6	8.0	5.2	8.0	5.9	8.0
Mean (SD)	23.3 (4.3)	23.4 (4.3)	22.8 (4.3)	23.4 (4.3)	23.2 (4.2)	23.4 (4.3)

Full citation

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

Ref Id

381247

Country/ies where the study was carried out

Netherlands

Study dates

January 1989 and June 2007

Source of funding

Sample size:

Total in OMEGA study n=26465

Endometriosis group n=3657

Comparison group n=5247

Characteristics

Year of birth

Characteristic	Endometriosis group		Comparison group	
	N	%	N	%
Year of birth				
≤1955	778	21.3	836	15.9
1955-9	1382	37.8	1819	34.7
1960-4	1125	30.8	1882	35.9
≥1965	372	10.2	710	13.5
Age (years) at diagnosis of endometriosis or first visit				
<25	351	9.6	182	3.5
25-29	1314	35.9	1258	24.0
30-34	1300	35.5	2301	43.9
35-39	527	14.4	1326	25.3
≥40	165	4.5	180	3.4
Time since diagnosis of endometriosis or first visit (years)				
<5	75	2.1	150	2.9

Details

OMEGA study: initiated in 1995, nationwide cohort study of 26465 women with subfertility problems (unable to concieve after 1 or more years of frequent unprotected intercourse). Looked at the effect of hormone stimulation in IVF treated women who had completed at least one IVF treatment cycle. Women were treated in 1 of 2 IVF clinics and a comparison group of non IVF women from 4 clinics who were subfertile (had

Results

Two analyses: 1st: included events in women diagnosed with OC or BOT on the same date or after date of first diagnosis of endometriosis. 2nd (Main analysis): included events in women diagnosed with OC or BOT after the date of first diagnosis of endometriosis.

Also analysed by self reported endometriosis and medical record.

Confounder adjustment: age, oral contraceptive use, IVF treatment and parity.

Median follow up time: 15.2 years (whole population), 10.9 years to ovarian cancer diagnosis, 9.5 years to BOT diagnosis.

78% of diagnoses of endometriosis was confirmed by pathology report (surgery/histology), 22% self reported.

Time intervals between diagnosis of endometriosis and OC or BOT: 3-12 months n=3, 1-10 years n=7, 10-20 years n=13, 20 years + n=3.

	All case n=34		Ovarian cancer (n=19)		BOT n=15	
	HR	95% CI	HR	95% CI	HR	95% CI
First analytic approa						
No endometriosis (n=5247)	1.0	Ref.	1.0	Ref.	1.0	Ref.

Limitations

Prevalence study critical appraisal

Was the sample representative of the target population? Unclear. Subfertile population - unclear if the results would differ/apply to a fertile population.

Were the study participants recruited in an appropriate way? Yes through the OMEGA cohort study.

Was the sample size adequate? Yes

Were the study subjects and setting described in detail? Yes.

Is the data analysis conducted with sufficient coverage of the identified sample? 4% refused linkage with PALGA and were

Study details	Participants	Diagnosis	Outcomes	Comments
Grants from the Health Research and Development Counsel and the Dutch Ministry of Health.	5-9 209 5.7 238 4.5 10-14 934 25.5 2725 51.9 15-19 1554 42.5 1962 37.4 ≥20 885 24.2 172 3.3 Oral Contraceptive use (years) No OC use 426 11.6 708 13.5 1-4 775 21.1 1059 20.2 5-9 1075 29.4 1583 30.2 ≥10 475 13.0 721 13.7 unknown 906 24.8 1176 22.4 Number of children 0 1510 41.3 2060 39.3 1-2 1775 48.5 2873 54.8 ≥3 160 4.4 226 4.3 Unknown 212 5.8 88 1.7 Main cause of subfertility Tubal 711 19.4 Male Unexplained 579 15.8 3413 65.0 Endometriosis 696 19.1 1834 35.0 Ovarian 468 12.8 Cervical 49 1.3 Mixed 19 0.5 831 22.7 304 8.4	other treatments e.g. tubal surgery/ hormonal treatments) were evaluated (n=6604). Diagnosis of endometriosis: Cohort linked with PALGA (all records of histological and cytological diagnoses made in the Netherlands). Trained research assistants extracted data from medical files on gynae history, diagnoses, treatments. NOTE: due to limited funding only 9/12 centres had the data extracted (76%). 968 women with endometriosis (PALGA confirmed)	Any endometriosis (n=3657) Crude 7.9 3.0-20.3 11.6 2.7-50.2 5.4 1.5-19.1 Age adjusted 9.7 3.7-25.1 13.4 3.1-58.4 7.3 2.0-26.3 **Second analytical approach** n=31 n=18 n=13 Any endometriosis Crude 7.0 2.7-18.3 10.9 2.5-47.4 4.4 1.2-16.1 Age adjusted 8.2 3.1-21.6 12.4 2.8-54.2 5.5 1.5-20.2 Adjusted for all confounders * 8.4 3.2-22.1 12.7 2.9-55.5 5.5 1.5-20.4 Ovarian endometriosis 11.3 4.0-31.8 15.0 3.1-72.4 8.9 2.2-35.7 Extraovarian endometriosis 7.7 2.1-28.7 19.1 3.5-104.5 - - Unknown location of endometriosis 6.0 2.0-18.1 8.1 1.6-41.8 4.7 1.0-21.5	excluded (n=1017). 24% medical records were not extracted due to limited funding and used results from questionnaire. Were objective, standard criteria used for measurement of the condition? Mixed methods.ICD codes linked with the National Cancer Institute and PALGA and/or medical records and/or self reported in risk questionnaire. Was the condition measured reliably? Yes for ICD codes, and medical records. Unclear validation of the questionnaire. Was there appropriate statistical analysis? Yes Are all confounding factors/ subgroups/ differences identified and accounted for? No: only age out of the

Study details	Participants	Diagnosis	Outcomes	Comments
	Unknown IVF No 592 16.2 478 9.1 Yes 3065 83.8 4769 90.9 **Inclusion criteria** • Women diagnosed with endometriosis • Comparison group: women with subfertility (not due to endometriosis. it is unexplained or a male factor) • See Diagnosis for further information. **Exclusion criteria** None described.	2270 women with endometriosis (medical records) of which 387 were on PALGA 806 reported endometriosis in the questionnaire (medical records could not be retrieved) Total included: 3657 women with endometriosis Comparison group selection: Subfertile women whose cause was not endometriosis e.g male fertility issue, unexplained cause (no abnormalities found in work up), in their medical records. Also included women who reported a male	*age (2.d.p), OC use (<5 and ≥5years), child (y/n), IVF (y/n). Note: OC use had missing data (24.8% and 22.4% respectively). Parity missing data (5.8% and 1.7% respectively) which may have biased the data. **First analysis:** Ovarian cancer: 17/3657 endo, 2/5247 non endo BOT: 12/3657 endo, 3/5247 non endo **Second analysis:** Ovarian cancer: 16/3657 endo, 2/5247 non endo BOT: 10/3657 endo, 3/5247 non endo Also report results restricted to: only self reported endometriosis diagnoses	GDG listed confounders. Additional confounders controlled for: parity, oral contraceptive use, IVF Were subpopulations identified using objective criteria? No subpopulation analysis was described in the methods but location of the endometriosis and the risk of ovarian cancer results were presented. **Other information** Note: prevalent and incident cases of endometriosis. All cancer cases are included from after the index date in main analysis.

Study details	Participants	Diagnosis	Outcomes	Comments
		cause in the questionnaire but it was not in their medical records (n=794) as it had a 71% positive predictive value. Total included: 5247 Risk factor information:23 page questionnaire sent to 25353. 16,343 returned it (65.2% response). 4% refused linkage with NCR or PALGA. Cancer diagnosis: Linked the cohort to the Dutch Pathology Database (PALGA) and the Netherlands Cancer Registry (96% complete data of the Netherlands) to assess the occurrence of ovarian cancer		

Study details	Participants	Diagnosis	Outcomes	Comments
		and borderline ovarian tumours. January 1989-June 2007 cancer incidence retrieved. Only those who explicitly declined linkage to the databases were excluded (n=1017) Observation time: time from diagnosis of endometriosis or 1 January 1989 (if diagnosed before then). N=2 excluded due to being diagnosed with ovarian cancer prior to this date. Comparison group: time from first IVF/first clinic visit for subfertility evaluation/1 January 1989,		

Study details	Participants	Diagnosis	Outcomes	Comments
		whichever came last. Observation stopped: June 2007/ date of first cancer diagnosis/ date of bilateral oophorectomy (n=32)/ death (n=42), whichever came first.		
Full citation 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 **Ref Id** 428570 **Country/ies where the study was carried out**	**Sample size** N= 7,537 endometriosis patients (5,468 with surgical confirmation) N=15,074 control group (matched by age, index year, obstetric history, SES, work and urbanisation), two controls per case. **Characteristics** Total follow up: 136,643 person years. **Inclusion criteria** • Women aged 20-51 years **Exclusion criteria** • Women with a diagnosis of EOC, endometriosis or with a total hysterectomy prior to their diagnosis of endometriosis and without a visit to an obstetrician	**Details** Note: only women with 3 or more visits and with a primary diagnosis of endometriosis within 1 year or with one surgically confirmed diagnosis of endometriosis during the study period were classed as the exposure group. Index date: date of the first visit/admission to between 2000-2009 that	**Results** 72.5% of all women with endometriosis had a surgical confirmation of their diagnosis. Risk of invasive epithelial ovarian cancer: Endometriosis patients with EOC: 15/7537 Control group with EOC: 9/15,074 Adjusted HR (95% CI): 3.28 (1.37-7.85) Adjusted for age, SES, work, urbanization, PID, infertility, CVD, DM, chronic liver disease, rheumatic disease and Charlson Comorbidity Index. Results by type of diagnosis (Post hoc analysis): Surgical confirmation adjusted HR (95% CI): 3.87 (1.58-9.47), n=13 EOC in 5,468 women. No surgical confirmation adjusted HR (95% CI): 1.64 (0.35-7.80), n=2 EOC in 2069 women.	**Limitations** Prevalence study critical appraisal Was the sample representative of the target population? Yes Were the study participants recruited in an appropriate way? Yes through the national database Was the sample size adequate? Yes Were the study subjects and setting described in detail? Yes. Is the data analysis conducted with sufficient coverage of the identified sample? Unclear

Study details	Participants	Diagnosis	Outcomes	Comments
Taiwan **Study dates** 2000-2009 **Source of funding** Grants from the Ministry of Science and Technology, Executive Yuan, Taipei, Taiwan, Taipei Veterans General Hospital, Taipei, Taiwan and the Foundation of Cheng-Hsin General Hospital, Taipei, Taiwan.	or gynaecologist during the study period • Patients with synchronous EOC and endometriosis • Patients with a diagnosis of EOC within the 1st year after their first diagnosis of endometriosis or the first visit/ admission to an obstetric/gynae provider.	resulted in the diagnosis of endometriosis in the endometriosis group, first visit/ admission to an obstetric/gynae provider during the study period for the control group. Validation of cancer diagnosis with the Registry of Catastrophic Illness Patients database. Follow up: until hospital admission for EOC, death, or end of the study. Does not describe any censoring.		the number of drop outs/ lost to follow up. No description of censoring. Were objective, standard criteria used for measurement of the condition? ICD coding. Note: women who had less than 3 outpt apts within the year of initial endometriosis diagnosis and without a surgical confirmation were not included in the exposure group. Potentially milder cases were excluded or put in the control group. Was the condition measured reliably? See comment above. Was there appropriate statistical analysis? Yes. Are all confounding factors/ subgroups/ differences identified and accounted for? Age and infertility were

Study details	Participants	Diagnosis	Outcomes	Comments
				controlled for. No information on severity, FHx, smoking or hormone treatment use. Additional confounders controlled for: SES, work, urbanization, PID, CVD, DM, chronic liver disease, rheumatic disease and Charlson Comorbidity Index. Were subpopulations identified using objective criteria? No subpopulation analysis was described in the methods but surgical confirmation of diagnosis of endometriosis was explored. **Other information** Note: population overlap with Chang 2014, Kok 2015, and Lee 2015.
Full citation	**Sample size**	**Details**	**Results** Observed: 46 incident ovarian cancers	**Limitations**

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr>
<td>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
Ref Id
403349

Country/ies where the study was carried out
Japan

Study dates
1985-1995 recruitment with follow up to 2002.

Source of funding
Grant-in-aid for Scientific</td>
<td>N=70,251 enrolled in the Shizuoka Cohort Study of Ovarian Cancer Screening Programme.
N=7,563 women with ovarian endometrioma detected by US.
n=6398 women with a clinically documented ovarian endometrioma and successful tracing (study population)

Characteristics
Mean age at diagnosis of ovarian endometrioma: 38.4 years
Average age at ovarian cancer diagnosis 51.4 (range 24-59) years.
Average follow up time of 12.8 years, with a total of 79, 102 person years.
Total number of women according to duration of follow up: <8 years n=995, 8-12 years n=1,991, >12 years n=3,412
Age at cohort entry: 20-29 years n=926, 30-39 years n=2,019, 40-49 years n=1,892, >50 years n=1,561.
For other baseline characteristics see Kobayashi 2008.

• Inclusion criteria
Women from the Shizuoka Cohort Study of Ovarian Cancer Screening Programme who on ultrasound revealed an ovarian</td>
<td>The Shizuoka Cohort study on Endometriosis and Ovarian Cancer Programme started in 1985 as part of the Shizuoka Cohort Study of Ovarian Cancer Screening Programme and the Shizuoka Cancer Registry System (established 1980). 212 hospitals, with participants from 35 townships.
Diagnosis: ultrasound ovarian endometrioma (transabdominal and/or transvaginal ultrasound). Sonographic criteria: cystic structure with round-shaped homogeneous hypoechoic</td>
<td>Expected: 5.14 (taken from the general population)
Overall SIR: 8.95 (95% CI 4.12-15.3)
<table>
<tr><th>Variable</th><th>Observed</th><th>SIR</th><th>95% CI</th></tr>
<tr><td>Ovarian cancer</td><td>46</td><td>8.95</td><td>4.12-15.3</td></tr>
<tr><td>Years of follow up
<8
8-12
>13
p value for trend</td><td>
9
12
25
</td><td>
19.3
6.42
8.92
0.021</td><td>
6.94-30.6
4.79-8.01
7.56-11.5
</td></tr>
<tr><td>Year of diagnosis
1985-1987
1988-1990
1991-1993
1994-1995
P value for trend</td><td>
10
15
8
13
</td><td>
7.14
10.7
5.71
13.9
0.341</td><td>
3.07-11.6
4.11-17.0
2.18-9.19
6.01-20.7
</td></tr>
<tr><td>Age at diagnosis, year
20-29
30-39
40-49
50-59
P value for trend</td><td>
2
5
13
26
</td><td>
3.88
4.85
8.03
13.2
0.014</td><td>
1.28-4.61
2.09-7.74
4.78-11.9
8.87-18.5
</td></tr>
</table>
For other results see Kobayashi 2008.</td>
<td><u>Prevalence study critical appraisal</u>
Was the sample representative of the target population? Only for ovarian endometrioma population.
Were the study participants recruited in an appropriate way? Yes
Was the sample size adequate? Yes
Were the study subjects and setting described in detail? Yes
Is the data analysis conducted with sufficient coverage of the identified sample? Yes.
Were objective, standard criteria used for measurement of the condition? USS. Risk of misclassification bias.
Was the condition measured reliably? USS. Risk</td>
</tr>
</table>

Study details	Participants	Diagnosis	Outcomes	Comments
Research from the Ministry of Education, Science, and Culture of Japan (H.K.).	endometrioma at a study hospital during the recruitment period. Age 20-59 years. **Exclusion criteria** • Those who did not want to participate (n=743, 9.8%) Entry ultrasounds were lost (n=108, 1.4%) Records were deleted due to inconsistencies uncovered during record linkage (n=66, 0.87%) Known ovarian cancer at time of enrollment (n=6, 0.1%) Prevalent cancer before entry (n=41, 0.5%) Unilateral oophorectomy or cystectomy for reasons other than ovarian endometrioma (n=201, 2.7%) Women >60 years	tissue of low level echoes within the ovary and thick cystic wall with regular margins. Pelvic examination was also carried out. Repeat US every 3-6 months (carried out by a gynaecologist at a regional hosptial). Follow up: stopped at the date of emmigration or gynaecological surgery, diagnosis of ovarian cancer, death, or end of follow up on December 31 2002, which ever occurred first. Info taken from hospital medical chart and location information (clinic records, telephone		of misclassification bias. Was there appropriate statistical analysis? Model based on age, year of follow up and age at diagnosis (for prevalence data). Logistic regression was only used for risk factor analysis. (longitudinal length of the tumors, menopausal status, age, parity, marital status, use of hormones, family history of cancer and current or previous smoking history. Dependent variable: endometrioma associated ovarian cancer). Are all confounding factors/ subgroups/ differences identified and accounted for? Not for prevalence data.Only for risk factor analysis (severity of endometriosis not looked at).

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr><td></td><td></td><td>directory, postmasters).
Questionnaires sent out to cohort who were living, linkage with Cancer registries.</td><td></td><td>Were subpopulations identified using objective criteria?
No subpopulations were identified.

Other information
Risk of misdiagnosis of the ovarian endometrioma with only using US
Selection bias- symptoms and US findings of ovarian cancer may be misinterpreted as endometriosis disease
Unknown if pelvic endometriosis</td></tr>
<tr><td>Full citation
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</td><td>Sample size:
See Kobayashi 2007

Characteristics
<table>
<tr><td>Variable</td><td>46 with ovarian cancer</td><td>6352 without ovarian cancer</td><td>P</td></tr>
<tr><td colspan="4">Age, years</td></tr>
<tr><td>Mean
20-44
45-9</td><td>50 +/-9
10 (22)
36 (78)</td><td>39 +/- 7
4281 (67)
2071 (23)</td><td>0.027</td></tr>
</table></td><td>Details
See Kobayashi 2007</td><td>Results
For other results see Kobayashi 2007.
Univariate analysis:
<table>
<tr><td rowspan="2">Variable</td><td colspan="3">Prediction of development of ovarian cancer</td></tr>
<tr><td>HR</td><td>95% CI</td><td>P</td></tr>
<tr><td>Tumor size (cm)
<9
≥9</td><td>
1.00
13.5</td><td>
8.98-19.3</td><td>
0.010</td></tr>
<tr><td>Menopausal status
No</td><td>
1.00</td><td>5.01-12.8</td><td>0.011</td></tr>
</table></td><td>Limitations
<u>Prevalence study critical appraisal</u>
Was the sample representative of the target population? Only for ovarian endometrioma population.
Were the study participants recruited in an appropriate way? Yes</td></tr>
</table>

Study details	Participants	Diagnosis	Outcomes	Comments

ovarian cancer development, European Journal of Obstetrics, Gynecology, & Reproductive Biology, 138, 187-93, 2008

Ref Id

428663

Country/ies where the study was carried out

Japan

Study dates:

See Kobayashi 2007

Source of funding

See Kobayashi 2007

Menopausal status			
Yes	35 (76)	731 (12)	0.011
No	11 (24)	5558 (87)	
Unknow	0 (0)	63 (1)	
Parity (No. of full term pregnancies)			
0	8 (61)	2147 (34)	0.212
1	16 (35)	1903 (30)	
2	1 (2)	1343 (21)	
≥3	1 (2)	639 (10)	
Unknow	0 (0)	320 (5)	
Marital status			
Yes	35 (76)	4159 (65)	0.674
No	11 (24)	1791 (28)	
Unknow	0 (0)	448 (7)	
Use of hormones			
None	12 (26)	5054 (79)	0.739
Unopposed E	0 (0)	192 (3)	
P	0 (0)	64 (1)	
E-P combination	7 (15)	129 (2)	
Others/unknown	27 (59)	959 (15)	
Current or previous smoking history			
Current	2 (4)	177 (3)	0.661
Former	1 (2)	197 (3)	

Yes	8.68		
Age <44 ≥45	 1.00 8.12	 5.21-11.7	 0.027
Parity	2.17	1.28-3.49	0.212
Marital status	1.13	0.89-1.42	0.674
Use of hormones	0.91	0.79-1.12	0.739
Family history of cancer	1.04	0.93-1.25	0.661
Current or previous smoking history	0.96	0.87-1.09	0.708

Multivariate analyses for the prediction of ovarian cancer

Variable	Prediction of development of ovarian cancer		
	HR	95% CI	P
Tumor size (cm) <9 ≥9	 1.00 5.51	 2.09-9.22	 0.031
Menopause No Yes	 1.00 3.21	 1.79-4.69	 0.039

Prevalence of ovarian cancer in tumors <6cm 0%, 16 (35%) in women with an endometrioma that was 6-9 cm, and 30 (65%) if ≥9cm diametre at the time of discovery.

At surgery for ovarian cancer, 32 (69.6%) of patients also had pelvic endometriosis.

Was the sample size adequate? Yes

Were the study subjects and setting described in detail? Yes

Is the data analysis conducted with sufficient coverage of the identified sample? Yes.

Were objective, standard criteria used for measurement of the condition? USS. Risk of misclassification bias.

Was the condition measured reliably? USS. Risk of misclassification bias.

Was there appropriate statistical analysis? Model based on age, year of follow up and age at diagnosis (for prevalence data). Logistic regression was only used for risk factor analysis. (longitudinal length of the tumors, menopausal status,

Study details	Participants	Diagnosis	Outcomes	Comments
	<table><tr><td>Never Unknow</td><td>43 (93) 0 (0)</td><td>5466 (86) 512 (8)</td><td></td></tr><tr><td colspan="4">Family history of cancer</td></tr><tr><td>Yes No Unknow</td><td>4 (9) 42 (91) 0 (0)</td><td>315 (5) 5716 (90) 321 (5)</td><td>0.708</td></tr><tr><td colspan="4">Diametre of endometrioma (cm)</td></tr><tr><td>≥9 <9 Unknow</td><td>30 (65) 16 (35) 0 (0)</td><td>512 (8) 5529 (87) 311 (5)</td><td>0.010</td></tr></table> Mean +/- SD. E: oestrogen, P: progesterone, others contain androgen (n=7), or GnRHa (n=20) for treatment of endometrioma. For other baseline characteristics see Kobayashi 2007 **Inclusion criteria** See Kobayashi 2007 **Exclusion criteria** See Kobayashi 2007		Clear cell in 18 (39%) and endometroid 16 (35%) of 46 women with ovarian cancer. Serous 5 (11%) and mucinous 4 (9%).	age, parity, marital status, use of hormones, family history of cancer and current or previous smoking history. Dependent variable: endometrioma associated ovarian cancer). Are all confounding factors/ subgroups/ differences identified and accounted for? Not for prevalence data.Only for risk factor analysis (severity of endometriosis not looked at). Were subpopulations identified using objective criteria? No subpopulations were identified. **Other information** None
Full citation Kok, V. C., Tsai, H. J., Su, C. F., Lee, C. K., The Risks for Ovarian, Endometrial,	**Sample size** n= 2266 endometriosis cohort (note includes 768 cases of pure adenomyosis) n= 9064 comparison cohort (1: 4 matching)	**Details** Data source: Taiwan National Health Insurance Research	**Results** Median time from the index date to cancer occurrence (all cancers) in endometriosis group: 34.3 months (IQR 18.7-46.8 months) and in the comparison group: 33 months (15.5-44.3 months).	**Limitations** Prevalence study critical appraisal Was the sample representative of

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr>
<td>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
Ref Id
370671

Country/ies where the study was carried out
Taiwan

Study dates
2003-2005 claims data followed up until December 31 2008

Source of funding
None reported.</td>
<td>Characteristics
<table>
<tr><th>Variable</th><th>Endometriosis cohort n=2266</th><th>Comparison cohort n=9064</th></tr>
<tr><td>Age group
20-30
31-40
41-50
>50</td><td>
551 (24.3%)
847 (37.4%)
788 (34.8%)
80 (3.5%)</td><td>
2204 (24.3%)
3388 (37.4%)
3152 (34.8%)
320 (3.5%)</td></tr>
<tr><td colspan="3">Site of endometriosis</td></tr>
<tr><td>Ovarian only</td><td>165 (7.3%)</td><td>0</td></tr>
<tr><td>Ovarian coexistent with other site</td><td>221 (9.8%)</td><td>0</td></tr>
<tr><td>Ovarian coexistent with adneomyosis</td><td>172 (7.6%)</td><td>0</td></tr>
<tr><td>Adenomyosis alone</td><td>768 (33.9%)</td><td>0</td></tr>
<tr><td>Adenomyosis coexistent with other site</td><td>401 (17.7%)</td><td>0</td></tr>
</table></td>
<td>Database (NHIRD)
Endometriosis: Newly diagnosed endometriosis or adenomyosis who had preserved uterus and ovaries and had no preexisting cancer and had an adequately lengthy follow up period (not defined). At least 3 outpatient claims, with at least 2 months between the first and third claims using ICD code 9th edition 617.
Comparison group: matched in a 1:4 ratio by age and index date.
Follow up: until they received a cancer diagnosis (3 claims using ICD code of</td>
<td>
<table>
<tr><td>Study cohort</td><td>Ovary cancer (13 endo/ 9 comparison groups)</td><td>Endometrial cancer (12 end o/ 5 comparison group)</td></tr>
<tr><td>Comparison cohort</td><td>Reference</td><td>Reference</td></tr>
<tr><td>Endometriosis cohort</td><td>4.56 (1.72-12.11)</td><td>4.05 (1.20-13.66)</td></tr>
<tr><td>Ovarian endometriosis group</td><td>4.37 (1.07-17.83)</td><td>3.23 (0.54-19.27)</td></tr>
<tr><td>Pure ovarian endometriosis</td><td>5.59 (0.67-46.48)</td><td>-</td></tr>
</table>
HR adjusted for: age, diabetes, chronic kidney disease, liver cirrhosis, rheumatoid arthritis and medication (medroxyprogesterone acetate, norethindrone acetate, danazol and gonadotropin-releasing hormone agonist (GnRH) for endometriosis.
Note: 34% of the endometriosis group had isolated adenomyosis.
</td>
<td>the target population? Yes
Were the study participants recruited in an appropriate way? Yes through the national database
Was the sample size adequate? Yes
Were the study subjects and setting described in detail? Yes.
Is the data analysis conducted with sufficient coverage of the identified sample? Unclear the number of drop outs/ lost to follow up but censoring was carried out.
Were objective, standard criteria used for measurement of the condition? ICD coding. Note: women who were evaluated less than 3 times or for a follow up period less than 2 months were excluded (n=3099). Potentially milder</td>
</tr>
</table>

Study details	Participants	Diagnosis	Outcomes	Comments
	All other sites, extragonadal, nonadenomyosis: 539 (23.8); 0 Medication Medroxyprogesterone acetate: 902 (39.8%); 713 (7.9%) Norethindrone acetate: 789 (34.8%); 972 (10.7%) Danazol: 377 (16.6%); 13 (0.1%) GnRH agonist: 2 (0.1%); 0 (0%) Comorbidity Diabetes Mellitus: 194 (8.6%); 344 (3.8%) Chronic Kidney disease: 2 (0.1%); 6 (0.1%) Liver cirrhosis: 413 (18.2%); 609 (6.7%) Rheumatoid arthritis: 60 (2.6%); 76 (0.8%) Follow up time, patient years: 9842; 36,274 **Inclusion criteria**	140-208, 9th edition or 1 inpatient claim), the last date of claims recorded or December 31, 2008. Endometriosis group: 9842 person years Comparison group: 36,274 person years Censoring: death, drop out of the National Health Insurance program or end of the observation period.		cases were excluded. Was the condition measured reliably? See comment above. No histological or surgical confirmation data was given. Was there appropriate statistical analysis? Yes. Are all confounding factors/ subgroups/ differences identified and accounted for? Age was controlled for. No information on severity, FHx, infertility, smoking or hormone treatment use. Additional confounders controlled for: DM, chronic kidney disease, liver cirrhosis, rheumatoid arthritis, and medication (medroxyprogesterone acetate, norethindrone acetate, danazol and gonadotropin-

Study details	Participants	Diagnosis	Outcomes	Comments
	• Women >20 years old with claims data from 2003-2005 **Exclusion criteria** • Women with preexisting malignancies, hysterectomy or oophorectomy • Women with preexisting endometriosis • Cases evaluated less than 3 times or for a follow up period less than 2 months			releasing hormone agonist (GnRH). Were subpopulations identified using objective criteria? Type of endometriosis. **Other information** Note: Cases evaluated less than 3 times or for a follow up period less than 2 months were excluded(n=3099) No censoring for women who have hysterectomy etc. after their index date.
Full citation 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	**Sample size** N=239,385 women were analyzed n=73,724 endometriosis (recall) to n=3782 tissue proved ovarian endometrioma (various diagnostic criteria explored) n=165,661 comparison control group **Characteristics** Median age of endometriosis patients with ≥1 medical record at outpatients or during hospitalization of endometriosis:	**Details** Data taken from the National Health Insurance Research Institute database (NHIRD) and was based on ICD codes. Endometriosis diagnosis: explored 13 different criteria	**Results** In total 348 of the 239,385 participants had EOC between 2001-2010. Recall endometriosis: n=73,724, EOC n=166, 874108.5996 person years compared to the control group n=165,661, EOC 182, 2354690.47 person years with a HR of 1.90 (1.51-2.37) Tissue proved endometriosis: n=3782, EOC n=47, 25138.4695 person years compared to the control group n=235,703, EOC 301, 3384200.4330 person years with a HR of 18.57 (13.37-25.79) The above were adjusted for: PID, infertility, Charlson co-morbidity index and age.	**Limitations** Prevalence study critical appraisal Was the sample representative of the target population? Yes Were the study participants recruited in an appropriate way? Yes through the national database Was the sample size adequate? Yes

Study details	Participants	Diagnosis	Outcomes	Comments
are different: A nationwide population-based cohort study, Medicine (United States), 94, e1633, 2015 **Ref Id** 428719 **Country/ies where the study was carried out** Taiwan **Study dates** 1996-2010 **Source of funding** Partly supported by grants from the Ministry of Science and Technology, Executive Yuan and Taipei Veterans General Hospital. No additional external funding was received.	34.0 (15-61) and for the control group 29.0 (15-60). Median age of endometriosis patients with medical records on surgically confirmed procedures limited by ICD9-CM 65.1X and 65.2X (tissue proven endo) 38.0 (18-59) and for the control group 30.0 (15-60). **Inclusion criteria** • Women aged 20-51 years with at least 1 gynaecologic visit after 2000 **Exclusion criteria** • Men • 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	from: at least 1 medical record of endometriosis at outpatient clinics or during hospitalization (recalled and or/ self reported endometriosis) to medical record based on surgically confirmed procedures limited by ICD9-CM 65.1 and 65.2X (tissue proved ovarian endometrioma). Index date endometriosis group: date of the first visit/admission from 2000-2010 Index date comparison control group: date of the first visit to an obstetric/ gynaecological provider or admission during the study period.		Were the study subjects and setting described in detail? Yes. Is the data analysis conducted with sufficient coverage of the identified sample? Unclear the number with inadequate basic data and the number of drop outs/ lost to follow up but censoring was carried out. Were objective, standard criteria used for measurement of the condition? ICD coding, medical records. Was the condition measured reliably? various diagnostic criteria were explored. Was there appropriate statistical analysis? Yes. Are all confounding factors/ subgroups/ differences identified and accounted for? Age and infertility were

Study details	Participants	Diagnosis	Outcomes	Comments
		Follow up:hospitalization with EOC or death, whichever came first, or the end of the study. Censored patients: lost to follow up, no diagnosis of EOC EOC was confirmed in inpatients with tissue approval and validated using the major disease files (Registry for Catastrophic Illness patients)		controlled for. No information on severity, FHx, smoking or hormone treatment use. Additional confounders controlled for:PID, Charlson co-morbidity index. Were subpopulations identified using objective criteria? No. **Other information** Note: 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 Presume 1st year of EOC was excluded as the paper only presents EOC values from 2001-2010.
Full citation Melin, A., Sparen, P., Persson, I.,	**Sample size** N=67339 cases idenitifed	**Details** National Swedish	**Results** Accuracy of ICD coding: 42/326 randomly selected medical records of patients in the cohort treated at	**Limitations** Prevalence study critical appraisal

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr>
<td>Bergqvist, A., Endometriosis and the risk of cancer with special emphasis on ovarian cancer, Human Reproduction, 21, 1237-1242, 2006
Ref Id
370912

Country/ies where the study was carried out
Sweden

Study dates
1969-2000

Source of funding
None described.</td>
<td>N=66187 with complete data/ eligible for follow up
N=64492 women entered the study (1691 had cancer diagnosis before/ same time as hospitalization and 4 had incomplete date of diagnosis).

Characteristics
• Average time of follow up: 12.7 years
• Average age at the first hospitalization with a diagnosis coded for endometriosis: 39.4 years (SD 10.4) - over whole study period, 42.1 (SD 11.7, p<0.001) between 1994-2000.
• Average age at cancer diagnosis was 55.1years (SD 10.2).

Inclusion criteria
• Women discharged from hospital with a first diagnosis of endometriosis from 1969-2000 (National Swedish Inpatient Register data).

Exclusion criteria
• First year of follow up was excluded.
• 3622 incident cases of cancer recorded (5.6%) and 264 had ≥1 type of cancer during follow up. 1968 (37%) were</td>
<td>Inpatient Register (covered 60% of the Swedish population in 1969, 85% in 1983, close to 100% from 1987): to identify women with endometriosis for the first time who had been discharged from a Swedish hospital. Note: previous diagnosis made clinically or day laparoscopic surgery is not covered by the register. Used ICD codes; ICD 8 625.30-625.33, 625.38 and 625.39, ICD 9; 617A-617G and 617X, ICD 10; N80.0-N80.9.
National Swedish Cancer Register: to identify women</td>
<td>Huddinge University Hospital were reviewed- 100% accuracy.
Histological verification: 47/326 randomly selected medical records of patients in the cohort treated at Huddinge University Hospital were reviewed- 81%, n=38 had histological confirmation of endometriosis.
Total number of person years: 766,556
Total of 3349 cancer cases included in the cohort.

<table>
<tr><th>Cancer type or site (ICD 7 code)</th><th>Number of person years</th><th>Observed number</th><th>Expected number</th><th>Ratio of observed to expected</th><th>95% CI</th></tr>
<tr><td>Cervical (170)</td><td>528441</td><td>51</td><td>80.18</td><td>0.64</td><td>0.47-0.84</td></tr>
<tr><td>CIS of the cervix (not included in 170)</td><td>508447</td><td>523</td><td>584.5</td><td>0.89</td><td>0.82-0.97</td></tr>
<tr><td>Endometrial (172)</td><td>427114</td><td>92</td><td>77.37</td><td>1.19</td><td>0.96-1.46</td></tr>
<tr><td>Uterine not otherwise specified (174)</td><td>427220</td><td>11</td><td>10.33</td><td>1.06</td><td>0.53-1.90</td></tr>
<tr><td>Ovarian (1750)</td><td>444931</td><td>122</td><td>85.09</td><td>1.43</td><td>1.19-1.71</td></tr>
<tr><td>Fallopian tube (1751,</td><td>766498</td><td>10</td><td>8.32</td><td>1.20</td><td>0.58-2.21</td></tr>
</table>
</td>
<td>Was the sample representative of the target population? Unclear. Very limited baseline characteristics given. Population is hospitalized women with endometriosis. Does not include those that have not been hospitalized for endometriosis.
Were the study participants recruited in an appropriate way? Yes- National Database.
Was the sample size adequate? Yes
Were the study subjects and setting described in detail? Very limited baseline characteristics described.
Is the data analysis conducted with sufficient coverage of the identified sample? Yes.
Were objective, standard criteria used for measurement of the</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr>
<td></td>
<td>excluded from the analysis due to having cancer before or at the time of diagnosis of endometriosis, or diagnosed within the first year of follow up (14 of these were ovarian cancer).
• Cancer specific exclusions:
• Uterine cancer: 26,334 had a hysterectomy before or at the same time as the diagnosis for endometriosis
• Ovarian cancer: 22633 had both ovaries removed before at the same time as the diagnosis for endometriosis.
• Cervical cancer: Total but not supravaginal hysterectomy- censored from follow up at that point in time for risk of cervical cancer.</td>
<td>with cancer ICD 7.
Start of follow up: 1 year after the year the woman was diagnosed with endometriosis (to exclude cancer prevalent already). Follow up continued until death, or emigration or until the end of the year 2000.
Censoring: women were censored at supravaginal or total hysterectomy (uterine cancer), total hysterectomy (cervical cancer) or when both ovaries had been removed (ovarian cancer)</td>
<td>
<table>
<tr><td>1758,1759)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Other female genital (176)</td><td>766409</td><td>25</td><td>24.72</td><td>1.01</td><td>0.65-1.49</td></tr>
</table>
Expected values: According to the cancer incidence in the female Swedish population by calendar year and 5 year age class (Breslow and Day 1987)
Ovarian cancer by location of endometriosis:
Ovarian endometriosis: SIR 1.77 (95% CI 1.38-2.24)
Non ovarian endometriosis: SIR 1.47 (95% CI 1.05-1.99)

<u>Ovarian cancer SIR by year of follow up, age and ovarian endometriosis by Age:</u>
<table>
<tr><th>Variable</th><th>Person years</th><th>Observed cases</th><th>SIR</th><th>95% CI</th></tr>
<tr><td>Years of follow up</td><td></td><td></td><td></td><td></td></tr>
<tr><td>1-2</td><td>29786.82</td><td>4</td><td>1.25</td><td>0.34-3.20</td></tr>
<tr><td>3-4</td><td>27350.48</td><td>9</td><td>2.64</td><td>1.20-5.00</td></tr>
<tr><td>5-10</td><td>57202.66</td><td>18</td><td>1.99</td><td>1.18-3.14</td></tr>
<tr><td>10-15</td><td>41182.81</td><td>20</td><td>2.23</td><td>1.36-3.44</td></tr>
<tr><td>15-20</td><td>26774.34</td><td>10</td><td>1.33</td><td>0.64-2.45</td></tr>
<tr><td>20-25</td><td>14909.87</td><td>8</td><td>1.58</td><td>0.68-3.10</td></tr>
<tr><td>Age</td><td></td><td></td><td></td><td></td></tr>
<tr><td>0-20</td><td>8582</td><td>0</td><td>0</td><td>0.00-10.26</td></tr>
<tr><td>20-30</td><td>143081</td><td>22</td><td>2.01</td><td>1.26-3.05</td></tr>
<tr><td>30-40</td><td>167155</td><td>52</td><td>1.76</td><td>1.32-2.31</td></tr>
<tr><td>40-50</td><td>108681</td><td>37</td><td>1.02</td><td>0.72-1.40</td></tr>
<tr><td>50-60</td><td>15000</td><td>9</td><td>1.32</td><td>0.61-2.52</td></tr>
</table>
</td>
<td>condition? Yes ICD codes.
Was the condition measured reliably? Yes ICD codes. Histology on a random sample was found on 81% of the cases.
Was there appropriate statistical analysis? No adjustment for the confounders. Stratification by age and year of follow up.
Are all confounding factors/ subgroups/ differences identified and accounted for? No: only age out of the GDG listed confounders.
Were subpopulations identified using objective criteria? No- location of endometriosis (ovarian) was presented but not described in the methods.

Other information</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td>
<table>
<tr><td>60-70
70+</td><td>1520
911</td><td>2
0</td><td>2.47
0</td><td>0.30-8.94
0.00-7.27</td></tr>
<tr><td>Ovarian endometriosis</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Age
20-30
30-40</td><td>
67622
82897</td><td>
12
37</td><td>
2.02
2.36</td><td>
1.04-3.52
1.66-3.25</td></tr>
</table>
</td><td>Limited to women who were hospitalized for endometriosis.
Note: uses some of the same population as Brinton 1997, Melin 2007.</td></tr>
<tr><td>Full citation
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
Ref Id
401660

Country/ies where the study was carried out
Sweden

Study dates
1969-2002

Source of funding
None described.</td><td>Sample size
n=3822 cases of cancer

Characteristics
Average time of follow up: 13.4 years
Average age at the first hospitalization with a diagnosis for endometriosis: 39.5 years (SD 10.5) for whole population.
Average age at cancer diagnosis in women with endometriosis: 55.9 years (SD 10.4)

Inclusion criteria
• Swedish Multi Generation Registered women (register from 1961 and born since 1932) who had been discharged from a Swedish hospital with the diagnosis of endometriosis for the first time from 1969-2002. Discharge diagnoses: ICD 8; 625.30-625.33, 625.38 and 625.39, ICD 9; 617A-617G, 617X and ICD; N80.0-N80.9.</td><td>Details
Endometriosis diagnosis by ICD code from the National Swedish Inpatient Register with linkage to the Multi-Generation Register.
Cancer diagnosis: National Swedish Cancer Register from 1958-2022 (ICD 7).
Follow up: until death, emigration or until the end of year 2002.
Censoring: when both</td><td>Results
4125 incident cases of cancer recorded (6.5%) and 567 women had ≥1 type of cancer during the follow up period.
3882 incident cases after the first year of follow up.
Expected values are taken from the population comparison cancer incidence created from the MGR by calendar year and 5 year age class.
Total person years in the cohort 792 013.
<table>
<tr><th rowspan="2">Type of cancer ICD 7 code</th><th colspan="2">All women</th><th colspan="2">Non parous women</th><th colspan="2">Parous women</th><th>P value for homogeneity</th></tr>
<tr><th>Observed</th><th>SIR (95%CI)</th><th>Observed</th><th>SIR (95% CI)</th><th>Observed</th><th>SIR (95%CI)</th><th></th></tr>
<tr><td>Ovarian (1750)</td><td>134</td><td>1.37 (1.14-1.62)</td><td>48</td><td>1.48 (1.11-1.96)</td><td>86</td><td>1.30 (1.05-1.61)</td><td>0.49</td></tr>
<tr><td>Endometrial (172)</td><td>97</td><td>1.14 (0.93-1.39)</td><td>28</td><td>0.93 (0.64-1.35)</td><td>69</td><td>1.04 (0.82-1.32)</td><td>0.62</td></tr>
</table>
</td><td>Limitations
<u>Prevalence study critical appraisal</u>
Was the sample representative of the target population? Unclear. Very limited baseline characteristics given. Population is hospitalized women with endometriosis. Does not include those that have not been hospitalized for endometriosis.
Were the study participants recruited in an appropriate way? Yes- National Database.
Was the sample size adequate? Yes
Were the study subjects and setting</td></tr>
</table>

Study details	Participants	Diagnosis	Outcomes	Comments
	Exclusion criteria • Patients clinically diagnosed within an open ward system, in private practice or as a day surgery procedure (as they are not covered by the register). Patients diagnosed with cancer before or at the same time as the first hospitalization and diagnosis of endometriosis (n=1719, 2.7%). Patients diagnosed with cancer within the first year of follow up (n=303, 7.3%)	ovaries were removed for ovarian cancer, supravaginal or total hysterectomy for endometrial cancer and total hysterectomy for cervical cancer. Parity: data does not cover stillbirths.	<table><tr><td>Cervical (171)</td><td>49</td><td>0.71 (0.53-0.94)</td><td>13</td><td>0.70 (0.40-1.21)</td><td>36</td><td>0.64 (0.46-0.90)</td><td>0.80</td></tr></table> Paper also reports ovarian cancer by parity SIR. Endometriosis location (Note: not specified as a subgroup in the methods): Ovarian endometriosis (n=24955 women, 39.2%) risk of ovarian cancer: SIR 1.59 (95%CI 1.26-1.98)	described in detail? Very limited baseline characteristics described. Is the data analysis conducted with sufficient coverage of the identified sample? Yes. Were objective, standard criteria used for measurement of the condition? Yes ICD codes. Was the condition measured reliably? Yes ICD codes. Was there appropriate statistical analysis? Yes. Are all confounding factors/ subgroups/ differences identified and accounted for? No: Adjustment for calendar year and 5 year age class. Stratification for parity. No other confounders adjusted for out of the GDG listed confounders.

Study details	Participants	Diagnosis	Outcomes	Comments
				Were subpopulations identified using objective criteria? No- location of endometriosis (ovarian) was presented but not described in the methods. **Other information** Adjusted by calendar year and 5 year age classes. Difference to Melin2006: access to MGR for parity information. Population: only hospitalized diagnoses of endometriosis. Uses some of the same data as Melin 2006 and Brinton 1997.
Full citation Mogensen, J. B., Kjaer, S. K., Mellemkjaer, L., Jensen, A., Endometriosis and risks for ovarian, endometrial and breast cancers: A	**Sample size** Ovarian cancer: N=45356 Endometrial cancer: N=43784 **Characteristics** Median age at ovarian cancer diagnosis was 55.4 years, at	**Details** The Danish National Patient Register - a nationwide register that comprises all hospital admissions for	**Results** <u>Endometrial cancer:</u> Subgroup analysis by age at first endometriosis (years) <30: SIR = 0.62 (0.17 - 1.59) 30-39: SIR = 1.81 (1.26 - 2.53) 40-49:	**Limitations** <u>Prevalence study critical appraisal</u> Was the sample representative of the target population? Unclear. Very limited baseline

Study details	Participants	Diagnosis	Outcomes	Comments
nationwide cohort study, Gynecologic Oncology, 143, 87-92, 2016 **Ref Id** 496724 **Country/ies where the study was carried out** Denmark **Study dates** 1977-2012 **Source of funding** This research was supported by an internal grant from the Danish Cancer Society (R121-A7558). The funding source was not involved in the study design, data collection, analysis, interpretation, writing or decision to submit this manuscript.	endometrial cancer diagnosis - 59 years. Median follow-up: ovarian cancer: 10.75, endometrial cancer: 4.1 **Inclusion criteria** • Women with a diagnosis of endometriosis in Denmark (a register-based cohort) **Exclusion criteria** • Women with an invalid personal identification number (n = 107) and women who had emigrated before a diagnosis of endometriosis (n = 37) were excluded. • For the analysis of ovarian cancer, further 434 women, who had undergone bilateral oophorectomy (operation codes 60,120 and 60,320 during 1977–1995 and KLAE20-21 and KLAF10-11 during 1996–2012) on the same date or before the date of diagnosis of endometriosis, were excluded. • For the analysis of endometrial cancer, 2006 women, who had a hysterectomy (operation codes 61000, 61020, 61040-050 and 61100 during 1977–1995 and KLCC10-11, KLCC20, KLCD00-01, KLCD04, KLCD10-11, KLCD30-31, KLCD40, KLCD96-97, KLEF13 and KMCA33 during 1996–2012) on	somatic conditions in Denmark since January 1977 and outpatient and emergency services since 1995: to identify women with a diagnosis of endometriosis. All first diagnoses of endometriosis (Danish version of the International Classification of Diseases (ICD), ICD-8 625.3, during 1977–1993 and ICD-10 N80 during 1994–2012) in both hospitalised patients and outpatients and identified a total of 45,934 women during the study period, were included. Ovarian cancer diagnosis: ICD-7=175; ICD-	SIR = 1.23 (0.80 - 1.80) ≥50: SIR = 1.75 (0.93 - 2.99) <u>Ovarian cancer:</u> Subgroup analysis by age at first endometriosis (years) <30: SIR 1.27 (0.71 – 2.10) 30-39: SIR 1.44 (1.10 – 1.85) 40-49: SIR 1.06 (0.83 - 1.34) ≥50: SIR 2.27 (1.61 – 3.10) SIR, standardised incidence ratio	characteristics given. Population is hospitalized women with endometriosis. Does not include those that have not been hospitalized for endometriosis. Were the study participants recruited in an appropriate way? Yes- National Database. Was the sample size adequate? Yes Were the study subjects and setting described in detail? Very limited baseline characteristics described. Is the data analysis conducted with sufficient coverage of the identified sample? Yes. Were objective, standard criteria used for measurement of the condition? Yes ICD codes. Was the condition measured reliably? Yes ICD codes.

Study details	Participants	Diagnosis	Outcomes	Comments
	the same date or before the date of diagnosis of endometriosis, were excluded.	10=C56, C570-C574 Endometrial cancer diagnosis: ICD-7=172-174; ICD-10=C54-C55, C58		Was there appropriate statistical analysis? Yes Are all confounding factors/ subgroups/ differences identified and accounted for? No, only age Were subpopulations identified using objective criteria? No - location of endometriosis (ovarian/endometrial) was presented but not described in the methods. Other information Limited to women who were hospitalized for endometriosis. **Other information** None
Full citation Stewart, L. M., Holman, C. D. J., Aboagye-Sarfo, P., Finn, J. C., Preen, D. B., Hart, R., In vitro fertilization, endometriosis,	**Sample size** n=22,045 women with a first diagnosis of either infertility or procreative management between 1982-2002 n=21,646 included in the study n=2,978 women with endometriosis	**Details** Women were included if they had at least one hospital diagnosis of infertility or procreative	**Results** Total duration of follow up: 366,041 person years with a mean of 17 years Ovarian cancer was diagnosed in women between 33 and 61 years of age, mean age at diagnosis: 46 years. Out of the women with endometriosis (n=2,978), 1,914 were undergoing infertility treatment but not IVF and 1,064 were undergoing IVF.	**Limitations** <u>Prevalence study critical appraisal</u> Was the sample representative of the target population? Subfertile population comparison so may

Study details	Participants	Diagnosis	Outcomes	Comments
nulliparity and ovarian cancer risk, Gynecologic Oncology, 128, 260-264, 2013 **Ref Id** 371465 **Country/ies where the study was carried out** Western Australia **Study dates** 1982-2002 **Source of funding** Supported in part by a capacity building grant form the National Health and Medical Research Council, Australia.	**Characteristics** Mean age at the start of follow up: 31 years (also the median age) Mean age at the end of follow up: 48 years (also the median age) **Inclusion criteria** • Women aged 20-44 years • First diagnosis of infertility or procreative management between 1982-2002 **Exclusion criteria** • Interstate address or having moved out of the State (WA) • Started infertility treatment (classed as not at risk of ovarian cancer; n=13 BSO before 1st interferon admission, n=7 had ovarian cancer prior to or within 6 months of first infertility admission).	managment (ICD coding). WA Data Linkage System was used: retrieved exposure data from 1980-2010. Information was also extracted from the Hospital Morbidity Data System (inpatient admissions at all hospitals in WA) to identify cohort, diagnoses and surgical procedures. IVF treatment data was identified using the Hospital Morbidity Data System and the Reproductive Technoogy Register. Linkage to Midwives Notifications System to identify births, Death Register - deaths, WA	Risk of ovarian cancer in endometriosis patients, HR (95% CI): 2.23 (0.97-5.12) MVA: risk of ovarian cancer in endometriosis patients, HR (95% CI): 2.33 (1.02-5.35) adjusted for age at the start of follow up, SES, birth and IVF. In total there were 38 cases of ovarian cancer in the cohort (16 undergoing IVF and 22 not undergoing IVF). Figures specifically for endometriosis were not published so it is unclear how many of the women got ovarian cancer.	have a different risk to the general population. Were the study participants recruited in an appropriate way? Yes- National Databases, covers the state of Western Australia. Was the sample size adequate? Yes Were the study subjects and setting described in detail? Very limited baseline characteristics described. Is the data analysis conducted with sufficient coverage of the identified sample? Yes. Were objective, standard criteria used for measurement of the condition? ICD coding from different registries/ databases. Was the condition measured reliably? Yes ICD codes. Does not mention any pathology

Study details	Participants	Diagnosis	Outcomes	Comments
		Cancer Registry- cancers. Endometriosis: diagnosis recorded in hospital records at or before the start of follow up. Censoring: women diagnosed with Borderline Ovarian Cancer only if they underwent a BSO. Follow up: from date of first infertility admission and continued until the date of epithelial ovarian cancer diagnosis, date of BSO, date of death or censor date (15 August 2010)		confirmation of diseases. Was there appropriate statistical analysis? Yes. Are all confounding factors/ subgroups/ differences identified and accounted for? No: only age at the start of follow up, birth, IVF and socioeconomic status. Were subpopulations identified using objective criteria? No subpopulations. **Other information** Generalisability of results- subfertile population
Full citation Wang, K. C., Chang, W. H., Lee, W. L., Huang, N., Huang, H. Y., Yen, M. S.,	**Sample size** N=5,945 women with a new surgico-pathological diagnosis of endometriosis from 2000-2010	**Details** Surgico-pathological diagnosis of endometriosis: ICD 9th edition	**Results** Total person year follow up for endometriosis patients; 33,519 and controls; 135,408. Median f/u (range) for endometriosis patients; 2059 days (3-4019) and controls; 2080 days (1-5243 days)	**Limitations** Prevalence study critical appraisal Was the sample representative of

<table>
<tr><th>Study details</th><th>Participants</th><th>Diagnosis</th><th>Outcomes</th><th>Comments</th></tr>
<tr>
<td>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
Ref Id
417395

Country/ies where the study was carried out
Taiwan
Study dates:
2000-2010

Source of funding
Grants from the Ministry of Science and Technology, Executive Yuan,Taipei Veterans General Hospital, and the Foundation of Cheng-Hsin General Hospital.</td>
<td>N=23,780 controls (multivariable matched;age, year, SES, work, obstetric history, frequency of gynaecological/ obstetric providers' outpatient visits and urbanization) 4 per case.

Characteristics
Age of endometriosis patients (≤41, >41): 49.02%, 50.98%
Age of control patients (≤41, >41): 50.31%, 49.69%

Other factors listed in baseline characteristics are controlled for in the HR calculation.

Inclusion criteria
• Women with newly diagnosed endometriosis (after year 2000) ICD code 617 (9th edition)

Exclusion criteria
• Male
• Age <20 or >51 years old in 2000
• Subjects without OPD (outpt apt) >2000
• Subjects with a diagnosis of ovary cancer year<2000
• Subjects with a diagnosis of endometriosis year <2000
• Subjects with a hysterectomy year <2000</td>
<td>coding of 617. Surgical treatment coding was also retrieved limited to the ovary tube and peritoneal cavity eg. laparoscopy etc.
Index date for endometriosis patients: date of a new surgico-pathological diagnosis of endometriosis
Index date for controls: first visit to an obstetric/ gynae provider or admission during the study period
Cancer diagnosis validated using files from the Registry for Catastrophic Illness Patients with histologic subtype found from the National</td>
<td>Epithelial ovarian cancer:
Endometriosis patients: 39/5945
Control patients: 36/23780
Adjusted HR (95% CI): 5.62 (3.46-9.14) - adjusted for PID, infertility status, CVD, DM, chronic liver disease and rheumatic disease.
Post hoc subgroup analysis by age group (not described in methods):
<table>
<tr><td>Variable</td><td>Age<30 years (n=3148)</td><td>Age 30-39 years (n=9310)</td><td>Age 40-49 years (n=13747)</td><td>Age ≥50 years</td></tr>
<tr><td>Diagnosis of EOC (endo/control)</td><td>2/3</td><td>10/4</td><td>18/22</td><td>9/7</td></tr>
<tr><td>Adjusted HR* (95% CI)</td><td>3.34 (0.54-20.60)</td><td>19.41 (5.02-75.10)</td><td>3.41 (1.76-6.61)</td><td>9.63 (3.27-28.37)</td></tr>
</table>
*adjusted for the same factors as listed above</td>
<td>the target population? Yes
Were the study participants recruited in an appropriate way? Yes through the national database
Was the sample size adequate? Yes
Were the study subjects and setting described in detail? Yes.
Is the data analysis conducted with sufficient coverage of the identified sample? Unclear the number of drop outs/ lost to follow up. Patients were censored at this point.
Were objective, standard criteria used for measurement of the condition? ICD coding.
Was the condition measured reliably? Yes.
Was there appropriate statistical analysis? Yes.</td>
</tr>
</table>

Study details	Participants	Diagnosis	Outcomes	Comments
	• Bilateral salpingo oophorectomy and tubal ligation patients	Cancer Registration System. Patients followed until hospitalization with EOC or end of the study (Dec 31, 2010). Censoring: drop outs/ lost to follow up/ patients without an EOC event		Are all confounding factors/ subgroups/ differences identified and accounted for? No, only age and infertility. No information on severity, FHx, smoking or hormone treatment us. Additional confounders controlled for: PID, CVD, DM, chronic liver disease and rheumatic disease. Were subpopulations identified using objective criteria? No subpopulation analysis was described in the methods but age of patients and risk of invasive epithelial ovarian cancer was presented. **Other information** 1st year of cancer and endometriosis diagnoses were not excluded (29/39 EOC in endo pts were diagnosed in the first year of

Study details	Participants	Diagnosis	Outcomes	Comments
				follow up, 22/36 in the control group). Note: population overlap with Chang 2014, Kok 2015, and Lee 2015.
Full citation 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 **Ref Id** 428616 **Country/ies where the study was carried out** Taiwan **Study dates**	**Sample size** n=15,488 women with a diagnosis of endometriosis n=123,904 control cohort (8 to each case of endometriosis, age, sex and index year matched) **Characteristics** • Age 40-49 years: endometriosis group 12,656/15,488, and control group 101,248/123,904 • Age 50-59 years: endometriosis group 2304/15,488, and control group 18432/123,904 • Age ≥60 years: endometriosis group 528/15,488, and control group 4224/123,904 **Inclusion criteria** • Women with a diagnosis of endometriosis and cases which were matched (age, sex and index year) **Exclusion criteria** • Women with a diagnosis of cancer before the diagnosis of endometriosis	**Details** Used Longitudinal Health Insurance Database (part of the National Health Insurance Research Databases (NHIRDs) Selected patients with a diagnosis of endometriosis (ICD 9th edition code 617.X). Date of diagnosis was the baseline date for the patient. Women with ICD code for endometriosis assigned by a gynaecologist and the patients must have the	**Results** Endometrial cancer: Endometriosis group: 104/15488 Control group: 288/123,904 Adjusted HR (95% CI): 2.83 (1.49-5.35) Adjusted for age, urbanization level, monthly income, geographic region, hypertension, hyperlipidemia, obesity and diabetes mellitus. Age at first diagnosis subgroup analysis: ≤40 years: n=48 (endometriosis group) and n=224 (control group); adjusted HR (95% CI) 1.42 (0.55-3.70) >40 years: n=56 (endometriosis group) and n=64 (control group); adjusted HR (95% CI) 7.08 (2.33-21.55)	**Limitations** Prevalence study critical appraisal Was the sample representative of the target population? Yes Were the study participants recruited in an appropriate way? Yes through the national database Was the sample size adequate? Yes Were the study subjects and setting described in detail? Yes. Is the data analysis conducted with sufficient coverage of the identified sample? Unclear the number of drop outs/ lost to follow up. No description of censoring. Were objective, standard criteria

Study details	Participants	Diagnosis	Outcomes	Comments
January 1 1997-December 31 2000. Patients tracked for 10 years from study entry. **Source of funding** Supported by the National Science Council, Taiwan.		diagnosis for at least 2 times in the same year in outpatient clinic records. Endometrial cancer diagnosis: received 2 or more endometrial cancer diagnoses for ambulatory care visit or 2 or more diagnoses for inpatient care. Follow-up: from the endometriosis diagnosis until the occurrence of endometrial cancer or the end of the study,which ever came first. Censoring was not described.		used for measurement of the condition? ICD coding. Note: women who had less than 2 outpt apts within a year assigning the diagnosis code of endometriosis by a gynaecologist were not included. Potentially milder cases were excluded. Was the condition measured reliably? See comment above. No histological or surgical confirmation data was given. Was there appropriate statistical analysis? Yes. Are all confounding factors/ subgroups/ differences identified and accounted for? Age was controlled for. No information on severity, FHx, infertility, smoking or hormone treatment use.

Study details	Participants	Diagnosis	Outcomes	Comments
				Additional confounders controlled for: urbanization level, monthly income, resident region, and comorbidities. Were subpopulations identified using objective criteria? Age stratification. **Other information** Note: Censoring was not described. Unclear how many were lost to follow up/ inadequate data etc. No description of any exclusions for women with hysterectomy etc. Unclear if just new or includes old diagnoses of endometriosis prior to study start date.

BSO: Bilateral Salpingo-oophorectomy; BOT: Borderline ovarian tumour; CI: Confidence Interval; CPR: to add; CVD: Cardiovascular disease; DM: Diabetes mellitus; E: Estrogen; E-P: Estrogen-progesterone pill; EAOC: Endometriosis-associated ovarian carcinoma; ENDO: to add; EOC: Epithelial ovarian carcinoma; FHx: Family history; GDG: Guideline development group; GnRHa: gonadotropin-releasing hormone agonist; HR: Hazard ratio; ICD: International classification of disease;IQR: Interquartile range; IVF: In vitro fertilisation; MGR: to add; MVA: Multivariable analysis; NCR: to add; NIH: National Institute of Health; NHIRD: National Health Insurance Research Institute database; OC: Oral contraceptive; OPD: Outpatient data; OR: Odds ratio; P: progesterone; PALGA: Dutch public pathology database; PID: Pelvic inflammatory disease; RR: Risk ratio; SD: Standard deviation; SE: Standard error; SES: Socioeconomic status; SIR: Standardised incidence ratio; SR: to add; US: Ultrasound; USS: to add; WA: Western Australia;

G.6 Review question: Diagnosis – Ultrasound

2 **What is the accuracy of ultrasound in diagnosing endometriosis?**

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation 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 **Ref Id** 416861 **Country/ies where the study was carried out** United Kingdom **Study type**	**Condition** Women referred because of suspected or confirmed pelvic mass observed on ultrasound examination in primary care **Sample size** Total patients who had TVS n=1279 - scheduled for surgery n=364 excluded n=34 suspected or histologically confirmed ovarian torsion n=17 Included n=313 **Characteristics** Mean age 47 (95%CI 45-49) premenopausal 62% malignancy prevalence 31% **Inclusion Criteria** • Women had to have undergone at least one TVS examination for an adnexal mass at a maximum of 120 days before surgical excision of the mass.	**Tests** TVS Surgery and histology	**Methods** Defined Level II ultrasound examiners as non consultant examiners who could recognise and diagnose correctly almost all pathologies affecting female genital tract. All ultrasound examiners involved in this study were considered to be at Level II for performing ultrasound examinations (2D gray-scale and color Doppler) of the ovary. 37 ultrasound examiners did the ultrasounds Examiners were asked to give their primary subjective assessment of ultrasound findings to classify the mass as malignant or benign and to give a secondary	**Results** Diagnostic performance of subjective assessment of adnexal masses: Endometrioma: TP 41 TN244 FP 2 FN 14 sensitivity 0.75 (0.61-0.85) specificity 0.99 (0.97-1) LR+ 92 (23-368) LR- 0.26(0.16-0.40)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: low concern Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? unclear risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Some other intervention type **Aim of the study** To assess the diagnostic performance of subjective assessment by level II ultrasound examiners in predicting the specific histology of adnexal masses **Study dates** September 2010 to May 2013 at QCH February 2012 to December 2012 at WMUH May 2012 to December 2012 at PAH **Source of funding** Not reported	• Inclusion criteria published previously in Sayasneh et al 2013 Br J Cancer 108:2448-2454 **Exclusion Criteria** • patients referred to level III ultrasound		subjective assessment to predict final specific histology. Outcomes of subjective assessment were grouped into 16 categories corresponding to 16 histological subtypes. The ultrasound report was reviewed by the patients' clinician and further management was based on clinical assessment and ultrasound findings as well as further tests and imaging Histological examination: examination of excised tissue was carried out at each local center. Surgery: laparoscopy or laparotomy		B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Were all patients included in the analysis? No Could the patient flow have introduced bias? high risk
Full citation 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 **Ref Id** 401037 **Country/ies where the study was carried out** France **Study type** Prospective cohort study **Aim of the study**	**Condition** patients suspected of having deep pelvic endometriosis **Sample size** n=37 **Characteristics** Mean age 35.8 (range 24-46) 22 patients had never had surgery for endometriosis (15 had). 25 patients had hormonal therapy before surgery. **Inclusion Criteria** • Suspicion of 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. **Exclusion Criteria** None	**Tests** Endorectal ultrasonography surgery (laparoscopy [n=26] and laparotomy [n=11])	**Methods** Endorectal ultrasonography was performed by the same investigator in each case thereby avoiding interobserver variability. Patients had a rectal enema before the examination and were placed in the dorsal position. The examination was conducted without sedation with an axial rotating rigid probe.The 7.5MHz to 10MHz transducer was covered with a balloon filled with degassed water producing a 360 degrees view of the rectal wall and adjacent areas (posterior vaginal wall, uterine cervix, pouch of Douglas, and the region of	**Results** The time between endorectal ultrasonography and surgery ranged from 4 to 529 days. Sensitivity: 88% (47 to 100) Specificity: 97% (82 to 100)	**Limitations** QUADAS 2 Patient sampling: A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? NA

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Aim to evaluate the validity of endorectal ultrasonography in predicting rectal infiltration in patients with deep pelvic endometriosis **Study dates** April 1996 to July 2003 **Source of funding** Not reported			the uterosacral ligaments). The principal objective of ultrasonography was to visualize any infiltration of the rectal wall by slowly moving the probe up and down along its longitudinal axis. The examination focused particularly on the anterior and lateral sides of the rectum. Surgeons were informed of the results of the endorectal ultrasonography before the intervention. They were particularly requested to evaluate endometriosis infiltration of the rectal wall. The results of the endorectal ultrasonography were compared with the surgical and histopathologic findings. The diagnosis of endometriosis was		Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? high risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
			confirmed by histopathological means in all patients		Was there an appropriate interval between index test and reference standard? unclear Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk
Full citation Nisenblat, Vicki, Farquhar, Cindy, Akoum, Ali, Fraser, Ian, Bossuyt, M. M. Patrick, Hull, Louise M., Non-invasive tests for the diagnosis of endometriosis, Cochrane Database of Systematic Reviews, 2012 **Ref Id** 359883 **Country/ies where the study was carried out** New Zealand **Study type** Cochrane Review **Aim of the study**	**Condition** Study participants included women of reproductive age (puberty to menopause) with suspected endometriosis based on clinical symptoms and/or pelvic examination, who undertook both the index test and the reference standard. **Sample size** N=49 studies involving 4807 women (for both transvaginal ultrasound and MRI) **Characteristics** **Abrao 2007** **Clinical presentation:** dysmenorrhoea 53/104, deep dyspareunia 66/104, acyclical pelvic pain 17/104, infertility 55/104, cyclical bowel symptoms (pain/bleeding) 59/104,	**Tests** **Abrao 2007** **Index test:** TVUS **Reference test:** laparoscopy 104/104 (100%) + histopathology **Bazot 2009** **Index test:** TVUS (TVS); TRUS (RES) **Reference test:** laparoscopy 79/92 (85.9%), laparotomy 13/92 (14.1%) + histopathology **Bergamini 2010** Index tests: TRUS (TRS); TVUS (RWC-TVS) **Reference test:** laparoscopy 57/61 (93.4%), laparotomy 4/61 (6.6%) + histopathology	**Methods** **Abrao 2007** TVUS: deep retrocervical endometriosis defined as thick blocks of tissue, nodular formations or irregular shaped, hypoechoic, retractable masses in USL, POD and/or vagina; bowel involvement established as a long, nodular, predominantly solid, hypoechogenic lesion adhered to the wall of the intestinal loop; each examination interpreted in real time; **Bazot 2009** TVUS: all scans performed by a	**Results** **Abrao 2007** RVS (rectovaginal septum) endometriosis: Sensitivity (95% CI): 95% (83 to 99) Specificity (95% CI): 98% (91 to 100) Rectosigmoid endometriosis: Sensitivity (95% CI): 98% (90 to 100) Specificity (95% CI): 100% (93 to 100) **Bazot 2009** RVS (rectovaginal septum) endometriosis (TVUS): Sensitivity (95% CI): 9% (0 to 41) Specificity (95% CI): 99% (91 to 100) RVS (rectovaginal septum) endometriosis (TRUS):	**Limitations** AMSTAR Checklist 1. Was an 'a priori' design provided? Y 2. Was there duplicate study selection and data extraction? Y 3. Was a comprehensive literature search performed? Y 4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? No 5. Was a list of studies (included and excluded) provided? Y 6. Were the characteristics of the included studies provided? Y 7. Was the scientific quality of the included studies assessed and documented? Y 8. Was the scientific quality of the included studies used

Study details	Participants	Tests	Methods	Outcomes and results	Comments
To provide estimates of the diagnostic accuracy of imaging modalities for the diagnosis of pelvic endometriosis, ovarian endometriosis and deeply infiltrating endometriosis (DIE) versus surgical diagnosis as a reference standard. To describe performance of imaging tests for mapping of deep endometriotic lesions in the pelvis at specific anatomical sites. **Study dates** 2016 **Source of funding** Internal sources Cochrane Menstrual Disorders and Subfertility Group, University of Auckland, New Zealand. Technical support	cyclical urinary symptoms 14/104 **Age:** mean 33.8 ± 6.1 years, range 18 to 45 years **Number enrolled:** 104 women **Number available for analysis:** 104 women **Setting:** tertiary university hospital, referral centre for endometriosis, São Paulo University **Place of study:** São Paolo, Brazil **Period of study:** August 2004 to October 2006 **Bazot 2009** **Clinical presentation:** dysmenorrhoea 79/92, dyspareunia 63/92, dyschezia 32/92, dysuria 3/92, infertility 21/92; history of surgery for endometriosis 31/92 **Age:** median age 31.8 years, range 20 to 50 years **Number enrolled:** 92 women **Number available for analysis:** 92 women **Setting:** tertiary care Tenon Hospital, referral centre for endometriosis and Surgical Centre Trocadero	**Dessole 2003** **Index test:** TVUS (transvaginal ultrasonography); sonovaginography **Reference test:** laparoscopy 20/46 (43.5%), laparotomy 26/46 (56.5%) + histopathology **Eskenazi 2001** **Index test:** TVUS (transvaginal ultrasound) **Reference test:** laparoscopy 72/90 (80%), laparotomy 18/90 (20%) + histopathology **Falco 2011** **Index test:** TVUS (TVS) **Reference test:** laparoscopy 96/96 (100%) + histopathology **Fedele 1998** **Index test:** TRUS (transrectal ultrasonography) **Reference test:** laparoscopy 114	single radiologist with extensive experience in gynaecological imaging. TRUS: each examination interpreted in real time by the same gastroenterologist with 5 years' experience in endometriosis. **Bergamini 2010** TVUS, TRUS: all scans performed by the same operator (gynaecologist), who had extensive experience in ultrasonographic diagnosis of endometriosis. Operator blinded with respect to other diagnostic findings; unclear whether operator was aware of the results of an additional index test (same operator, different test times) **Dessole 2003** TVUS: operator obtained longitudinal and transversal scans of the uterus, with	Sensitivity (95% CI): 18% (2 to 52) Specificity (95% CI): 95% (88 to 99) <u>Rectosigmoid endometriosis (TVUS):</u> Sensitivity (95% CI): 94% (85 to 98) Specificity (95% CI): 100% (88 to 100) <u>Rectosigmoid endometriosis (TRUS):</u> Sensitivity (95% CI): 89% (78 to 95) Specificity (95% CI): 93% (77 to 99) <u>USL (TVUS):</u> Sensitivity (95% CI): 78% (68 to 87) Specificity (95% CI): 67% (30 to 93) <u>USL (TRUS):</u> Sensitivity (95% CI): 48% (37 to 59) Specificity (95% CI): 44% (14 to 79) <u>Vaginal wall involvement (TVUS):</u> Sensitivity (95% CI): 47% (28 to 66) Specificity (95% CI): 95% (87 to 99) <u>Vaginal wall involvement (TRUS):</u>	appropriately in formulating conclusions? Y 9. Were the methods used to combine the findings of studies appropriate? Y 10. Was the likelihood of publication bias assessed? No 11. Was the conflict of interest included? Y Where there is a high risk regarding applicability it is due to a two-gate design: according to Nisenblat et al. 2016 these are studies with two sets of inclusion criteria with respect to **Clinical presentation:** and one set of inclusion criteria with respect to reference standard (participants with or without a clinical suspicion of endometriosis scheduled for abdominal surgery). **<u>Quadas 2</u>** **Abrao 2007** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the

Study details	Participants	Tests	Methods	Outcomes and results	Comments
The Robinson Institute, University of Adelaide, Other. Access to academic resources External sources No sources of support supplied	**Place of study:** Paris, France **Period of study:** April 2000 to May 2005 **Bergamini 2010** **Clinical presentation:** dyspareunia and/or catamenial rectal pain 61/61, history of intermittent bowel obstruction 4/61, nulliparous 11/61, history of surgery for endometriosis 19/61 **Age:** mean age 33.1 years, range 28 to 37 years **Number enrolled:** 61 women **Number available for analysis:** 61 women **Setting:** University Hospitals of Verona and Varese, referral centres for endometriosis treatment **Place of study:** Verona and Varese, Italy **Period of study:** January 2008 to February 2009 **Dessole 2003** **Clinical presentation:** chronic pelvic pain, dysmenorrhoea or dyspareunia 38/46, infertility 20/46, gastrointestinal disorders 7/46, urinary disorders 6/46;	(81.4%), laparotomy 26 (18.6%) + histopathology **Ferrero 2011** **Index test:** TVUS (RWC-TVS) **Reference test:** laparoscopy 96/96 (100%) + histopathology **Ghezzi 2005** **Index test:** TVUS (transvaginal ultrasound, sign of 'kissing ovaries') **Reference test:** laparoscopy 710/710 (100%) + histopathology **Goncalves 2010** **Index test:** TVUS (TVUS-BP, with bowel preparation) **Reference test:** laparoscopy 194/194 (100%) + histopathology **Grasso 2010** **Index test:** TVUS (3D-TVUS) **Reference test:** laparoscopy 33/33	particular attention given to rectovaginal septum for detection of endometriotic lesions - criteria not specified **Eskenazi 2001** TVUS: all pelvic examinations and transvaginal ultrasounds conducted by a single gynaecologist who was not blinded to clinical information and to results of pelvic examination; level of expertise not reported **Falco 2011** TVUS: Operator not unaware of results of bimanual clinical examination but could ask questions about symptoms present; number of operators and level of expertise not provided **Fedele 1998** TRUS: ultrasonographer not aware of clinical findings or patient history; knew only	Sensitivity (95% CI): 7% (1 to 22) Specificity (95% CI): 100% (94 to 100) Ovarian endometriosis: Sensitivity (95% CI): 94% (81 to 99) Specificity (95% CI): 86% (74 to 94) **Bergamini 2010** Rectosigmoid endometriosis (RWS-TVUS): Sensitivity (95% CI): 96% (87 to 100) Specificity (95% CI): 90% (55 to 100) Rectosigmoid endometriosis (TRUS): Sensitivity (95% CI): 88% (76 to 96) Specificity (95% CI): 80% (44 to 97) **Dessole 2003** Posterior DIE (TVUS): Sensitivity (95% CI): 44% (26 to 62) Specificity (95% CI): 50% (23 to 77) Posterior DIE (SVG): Sensitivity (95% CI): 91% (75 to 98)	CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s)

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	endometriotic lesion detected on gynaecological examination 8/46; no patients had undergone surgical pelvic procedure before entering the study **Age:** mean 30.3 ± 4.2 years **Number enrolled:** 46 women **Number available for analysis:** 46 women **Setting:** University Hospital, University of Sassari **Place of study:** Sassari, Italy **Period of study:** January 2000 to October 2001 **Eskenazi 2001** **Clinical presentation:** dysmenorrhoea 40/90, pelvic pain 20/90, dyspareunia 20/90, infertility 12/90, abnormal pelvic examination 42/90; indications for surgery including pelvic pain 21%, infertility 13%, ovarian cysts 30%, fibroids 28%, suspected endometriosis 16%, tubal ligation 6.7%; nulliparous 42/90, nulligravid 33/90, current oral contraceptive users 4/90 **Age:** mean 35.7 ± 7.2 years, range 20 to 49 years	(100%) + histopathology **Guerriero 1996a** **Index test:** TVUS (transvaginal ultrasonography) **Reference test:** laparoscopy 99/118 (84%), laparotomy 19/118 (16%) + histopathology **Guerriero 1996b** **Index test:** TVUS (transvaginal ultrasonography) **Reference test:** laparoscopy, laparotomy (number for each group not reported) + histopathology **Guerriero 2007** **Index test:** TVUS (TVUS tenderness-guided approach) **Reference test:** laparoscopy 50/50 (100%) + histopathology **Guerriero 2008** **Index test:** TVUS (tg-TVUS)	that endometriosis was suspected; numbers of examiners and level of expertise not reported **Ferrero 2011** TVUS: bowel endometriosis appears ultrasonographically as a nodular, solid, hypoechoic lesion, adjacent to and/or penetrating the intestinal wall; unclear whether prespecified criteria or description of findings **Ghezzi 2005** TVUS: all ultrasound examinations performed by 3 examiners; level of expertise and blinding to clinical data not reported **Goncalves 2010** TVUS: all exams performed by the same radiologist, who was blinded with respect to clinical data and results of other exams to which the	Specificity (95% CI): 86% (57 to 98) **Eskenazi 2001** <u>Pelvic endometriosis:</u> Sensitivity (95% CI): 57% (39 to 73) Specificity (95% CI): 98% (90 to 100) **Falco 2011** <u>Pelvic endometriosis:</u> Sensitivity (95% CI): 96% (89 to 99) Specificity (95% CI): 80% (56 to 94) <u>Posterior DIE:</u> Sensitivity (95% CI): 74% (58 to 87) Specificity (95% CI): 96% (88 to 100) <u>RVS (rectovaginal septum) endometriosis:</u> Sensitivity (95% CI): 27% (6 to 61) Specificity (95% CI): 100% (96 to 100) <u>Rectosigmoid endometriosis</u>: Sensitivity (95% CI): 84% (64 to 95) Specificity (95% CI): 99% (92 to 100) <u>USL endometriosis:</u>	Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Bazot 2009** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number enrolled: 90 women (study sample); 120 women (test sample) **Number available for analysis:** 90 women – only 'study sample' arm included in current analysis; 'test sample' excluded for retrospective design **Setting:** Hospital of Desio (study sample) and University Hospital, Mangiagalli Hospital, University of Milan (test sample) **Place of study:** Desio (study sample) and Mangiagalli (test sample), Italy **Period of study:** July 1998 to December 1999 **Falco 2011** **Clinical presentation:** dysmenorrhoea 65/128, chronic pelvic pain 52/128, infertility 49/128, dyspareunia 41/128, dyschezia 23/128, palpable peritoneal nodules 33/128, ovarian cyst 18/128; previously diagnosed endometriosis 9/128 **Age:** mean 33.6 years, range 18 to 48 years **Number enrolled:** 128 women	**Reference test:** laparoscopy 88/88 (100%) + histopathology **Guerriero 2014** **Index test:** TVUS 2 types (2D-US (tg-TVUS) and 3D-US) **Reference test:** laparoscopy 194/202 (96%), laparotomy 8/202 (4%) + histopathology **Holland 2010** **Index test:** TVUS (TVS) **Reference test:** laparoscopy 201/201 (100%) **Hudelist 2011** **Index test:** TVUS (TVS) **Reference test:** laparoscopy 129/129 (100%) + histopathology **Hudelist 2013** **Index test:** TVUS (TVS) **Reference test:** laparoscopy 117/117	patient had been submitted; level of expertise not stated **Grasso 2010** TVUS: diagnosis of pelvic endometriosis based on different morphological criteria, which varied for each anatomical location of the disease and included thickening or echogenic nodules or masses with regular or irregular outlines, as described for each site (ovary, USL, posterior vaginal fornix, RVS, sigmoid colon, bladder, POD); **Guerriero 1996a** TVUS: all scans performed by the same physician; level of expertise and blinding to clinical data not reported **Guerriero 1996b** TVUS: all scans performed by the same physician; level of expertise and blinding to	Sensitivity (95% CI): 74% (57 to 88) Specificity (95% CI): 98% (91 to 100) <u>Vaginal wall involvement:</u> Sensitivity (95% CI): 31% (9 to 61) Specificity (95% CI): 100% (96 to 100) **Fedele 1998** <u>RVS (rectovaginal septum) endometriosis:</u> Sensitivity (95% CI): 97% (85 to 100) Specificity (95% CI): 96% (91 to 99) <u>Rectosigmoid endometriosis:</u> Sensitivity (95% CI): 100% (66 to 100) Specificity (95% CI): 98% (93 to 100) <u>USL:</u> Sensitivity (95% CI): 80% (44 to 97) Specificity (95% CI): 98% (93 to 100) <u>Vaginal wall involvement :</u> Sensitivity (95% CI): 100% (79 to 100) Specificity (95% CI): 100% (97 to 100)	Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number available for analysis: 96 women **Setting:** University Hospital "Federico II" **Place of study:** Naples, Italy **Period of study:** December 2008 to May 2010 **Fedele 1998** **Clinical presentation:** infertility 67/140, pelvic pain 52/140; clinical findings 21/140 **Age:** mean 30.2 ± 5.7 years **Number enrolled:** 140 women **Number available for analysis:** 140 women **Setting:** University Hospital, The University of Verona **Place of study:** Verona, Italy **Period of study:** November 1995 to April 1997 **Ferrero 2011** **Clinical presentation:** dysmenorrhoea 72/96, deep dyspareunia 49/96, chronic pelvic pain 61/96, dyschezia 39/96, infertility 32/96, diarrhoea 28/96, constipation 39/96, intestinal cramping 40/96, abdominal bloating 53/96, mucus in the	(100%) + histopathology **Leon 2014** **Index test:** TVUS (extended method: combination of bowel preparation with transvaginal gel instillation and use of 'sliding sign' for diagnosis) **Reference test:** laparoscopy 51/51 (100%) + histopathology **Mangler 2013** **Index test:** TVUS(vaginal ultrasound) **Reference test:** surgery (vaginal approach + laparoscopy ± laparotomy) 79/79 (100%) + histopathology **Menada 2008** **Index test:** TVUS 2 types (TVS; RWC-TVS) **Reference test:** laparoscopy, laparotomy (number in each group not	clinical data not reported **Guerriero 2007** TVUS: all scans performed by 1 investigator, who has had more than 15 years of experience with TVUS; unclear whether blinded to clinical data **Guerriero 2008** TVUS: all scans performed by 1 investigator who had more than 15 years' experience with transvaginal ultrasonography at the outset of the study; unclear whether blinded to clinical data **Guerriero 2014** TVUS: ll scans performed by 1 investigator who had more than 20 years' experience with transvaginal ultrasonography. Unclear whether operator was blinded to clinical data **Holland 2010**	**Ferrero 2011** <u>Bowel endometriosis:</u> Sensitivity (95% CI): 88% (76 to 96) Specificity (95% CI): 98% (88 to 100) <u>Rectosigmoid endometriosis:</u> Sensitivity (95% CI): 94% (83 to 99) Specificity (95% CI): 98% (89 to 100) **Ghezzi 2005** <u>Pelvic endometriosis:</u> Sensitivity (95% CI): 9% (6 to 12) Specificity (95% CI): 99% (97 to 100) **Goncalves 2010** <u>Rectosigmoid endometriosis:</u> Sensitivity (95% CI): 98% (91 to 100) Specificity (95% CI): 100% (97 to 100) **Grasso 2010** <u>DIE:</u> Sensitivity (95% CI): 79% (54 to 94) Specificity (95% CI): 60% (15 to 95)	A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Bergamini 2010** A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	stools 13/96, rectal bleeding 2/96; previous live birth 27/96, previous surgery for endometriosis 39/96, hormonal therapy at time of study 34/96 **Age:** mean 33.4 ± 5.2 years **Number enrolled:** 96 women **Number available for analysis:** 96 women **Setting:** University Hospital: San Martino University Hospital, endometriosis referral centre, Galliera Hospital **Place of study:** Genoa, Italy **Period of study:** January 2008 to November 2009 **Ghezzi 2005** **Clinical presentation:** chronic pelvic pain, dyspareunia, dysmenorrhoea 309/722, infertility 145/722, adnexal mass not suggestive of endometriosis 413/722 **Age:** premenopausal, mean age and age range not reported **Number enrolled:** 722 women **Number available for analysis:** 710 women	specified) 90/90 (100%) + histopathology **Pascual 2010** **Index test:** TVUS (Introital 3D-US) **Reference test:** laparoscopy 38/38 (100%) + histopathology **Piessens 2014** **Index test:** TVUS-BP (DIE-TVUS) **Reference test:** laparoscopy 85/85 (100%) + histopathology **Piketty 2009** **Index test:** TVUS; TRUS **Reference test:** laparoscopy, laparotomy (numbers for each procedure not specified) + histopathology **Reid 2013** **Index test:** TVUS, sliding sign (TVS) **Reference test:** laparoscopy 100/100	TVUS: TVS examination performed by 4 ultrasound operators who were all gynaecologists with a high level of expertise in gynaecological ultrasonography. Ultrasound operators blinded to previous surgical findings. Examiner A performed 104 (51.7%), examiner B performed 68 (33.8%), examiner C performed 18 (9%) and examiner D performed 11 (5.5%) examinations **Hudelist 2011** TVUS: all TVS scans performed by 1 experienced examiner who was blinded to results of the vaginal examinations but was aware that women were being investigated for chronic pelvic pain; therefore, endometriosis was suspected	<u>Bladder endometriosis*:</u> Sensitivity (95% CI): 25% (5 to 57) Specificity (95% CI): 100% (77 to 100) **Guerriero 1996a** <u>Ovarian endometriosis:</u> Sensitivity (95% CI): 85% (69 to 94) Specificity (95% CI): 97% (91 to 100) **Guerriero 1996b** <u>Ovarian endometriosis:</u> Sensitivity (95% CI): 83% (64 to 94) Specificity (95% CI): 93% (85 to 98) **Guerriero 2007** <u>Posterior DIE:</u> Sensitivity (95% CI): 90% (74 to 98) Specificity (95% CI): 95% (74 to 100) <u>Ovarian endometriosis:</u> Sensitivity (95% CI): 100% (66 to 100) Specificity (95% CI): 100% (91 to 100) **Guerriero 2008**	Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Setting: 2 university hospitals: University of Insubria Del Ponte Hospital and University of Berne Hospital **Place of study:** Varese, Italy, and Berne, Switzerland **Period of study:** January 2000 to November 2003 **Goncalves 2010** **Clinical presentation:** severe dysmenorrhoea 109/194, deep dyspareunia 120/194, cyclical bowel complaints 112/194, chronic pelvic pain 39/194, infertility 97/194, cyclical urinary complaints 18/194; mean time between onset of symptoms and diagnosis 5.2 years (range 0.4 to 10 years) **Age:** mean 34.2 ± 4.9 years **Number enrolled:** 194 women **Number available for analysis:** 194 women **Setting:** University Hospital, Sirio Libanes Hospital, University of São Paulo Medical School **Place of study:** São Paulo, Brazil **Period of study:** October 2006 to September 2008	(100%) + histopathology **Reid 2014** **Index test:** Sonovaginography (SVG) **Reference test:** laparoscopy 189/189 (100%) + histopathology **Ribeiro 2008** **Index test:** TRUS (Tr EUS) **Reference test:** laparoscopy 37/37 (100%) + histopathology **Said 2014** **Index test:** TVUS (TVS) **Reference test:** laparoscopy 125/125 (100%) + histopathology **Savelli 2011** **Index test:** TVUS (TVS) **Reference test:** laparoscopy 69/69 (100%) + histopathology	**Hudelist 2013** TVUS: all TVS scans performed by 1 experienced examiner who was not blinded to clinical data **Leon 2014** TVUS: all extended transvaginal sonographic examinations performed by 1 operator who had more than 10 years' experience in gynaecological sonography and 3 years' experience in assessment of deep infiltrating endometriosis; unclear whether operator was blinded to clinical data **Mangler 2013** TVUS: consultants who were not aware of results of the other tests and of the reference procedure **Menada 2008**a TVUS: 2 different experienced ultrasonographers	<u>RVS (rectovaginal septum) endometriosis:</u> Sensitivity (95% CI): 74% (59 to 86) Specificity (95% CI): 88% (74 to 96) <u>Anterior DIE:</u> Sensitivity (95% CI): 33% (13 to 59) Specificity (95% CI): 100% (95 to 100) <u>Rectosigmoid endometriosis:</u> Sensitivity (95% CI): 67% (50 to 81) Specificity (95% CI): 92% (80 to 98) <u>USL endometriosis</u>: Sensitivity (95% CI): 50% (29 to 71) Specificity (95% CI): 94% (85 to 98) <u>Vaginal wall involvement:</u> Sensitivity (95% CI): 91% (76 to 98) Specificity (95% CI): 89% (77 to 96) <u>Bladder endometriosis*:</u> Sensitivity (95% CI): 100% (40 to 100) Specificity (95% CI): 100% (96 to 100)	review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? unclear Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? unclear risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Grasso 2010 **Clinical presentation:** pain (dysmenorrhoea, dyspareunia, chronic pelvic pain) 18/33, infertility 5/33, adnexal masses and/or tenderness at physical examination 10/33 **Age:** mean 35, range 22 to 53 years **Number enrolled:** 33 women **Number available for analysis:** MRI 33 women; 3D-TVUS 24 women **Setting:** University Hospital, Villa Valeria Hospital and Campus Bio Medico University of Rome **Place of study:** Rome, Italy **Period of study:** June 2006 to June 2008 **Guerriero 1996a** **Clinical presentation:** symptoms and clinical findings: persistent adnexal mass 118/118 (100%), infertility 45/118 (53%) **Age:** mean 33.3 ± 9.6 years, range 14 to 54 years **Number enrolled:** 118 women **Number available for analysis:** 118 women	**Scarella 2013** **Index test:** TVUS (USTV-PI, with bowel preparation) **Reference test:** laparoscopy, laparotomy (numbers for each procedure not specified) + histopathology **Ubaldi 1998** **Index test:** TVUS **Reference test:** laparoscopy 133/133 (100%) + histopathology	independently performed examinations: 1 operator performed all TVS, second operator performed RWC-TVS. Operators were informed that rectovaginal endometriosis was suspected, but they were not aware of the findings of vaginal or rectal examination, and they were not informed of the findings of previous radiological examinations and results of other index tests **Pascual 2010** TVUS: scans carried out by 3 experienced examiners, using the same scanning protocol; stored 3D volumes analysed by just 1 examiner; unclear whether blinded to clinical data **Piessens 2014** TVUS: all examinations	**Guerriero 2014** <u>Posterior DIE (tg-TVUS):</u> Sensitivity (95% CI): 71% (61 to 80) Specificity (95% CI): 88% (81 to 94) <u>Posterior DIE (3D-TVUS):</u> Sensitivity (95% CI): 87% (78 to 93) Specificity (95% CI): 94% (87 to 97) <u>Rectosigmoid endometriosis (tg-TVUS):</u> Sensitivity (95% CI): 95% (87 to 99) Specificity (95% CI): 93% (87 to 97) <u>Rectosigmoid endometriosis (3D-TVUS):</u> Sensitivity (95% CI): 91% (82 to 96) Specificity (95% CI): 97% (92 to 99) **Holland 2010** <u>Pelvic endometriosis:</u> Sensitivity (95% CI): 56% (47 to 65) Specificity (95% CI): 95% (87 to 99) <u>DIE:</u>	**Dessole 2003** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? No Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Setting: University Hospital, University of Cagliari **Place of study:** Cagliari, Italy **Period of study:** November 1994 to November 1995 **Guerriero 1996b** **Clinical presentation:** not specified **Age:** range 20 to 49 years, mean not provided **Number enrolled:** 101 women **Number available for analysis:** 101 women **Setting:** University Hospital, University of Cagliari **Place of study:** Cagliari, Italy **Period of study:** November 1993 to October 1994 **Guerriero 2007** **Clinical presentation:** pelvic pain in all 50 women: dyspareunia 19/50, dysmenorrhoea 42/50, infertility 5/50; previous medical treatment for persistent pelvic pain (estrogens, progestins and/or gonadotropin-releasing hormone agonist and non-steroidal anti-		performed by a single operator who is a gynaecologist with a subspecialty degree in ultrasound and more than 10 years' experience, but no prior experience in detecting DIE; operator was not blinded to symptoms and history of women **Piketty 2009** TVUS: DIE defined as presence of hypoechoic and irregular nodes in assessed pelvic structures; intestinal DIE (ileum - rectum) defined as previously published (referenced to Bazot et al., 2007) and described; TRUS: showed up as hypoechoic peridigestive nodules of rounded or roughly triangular shape (ileum - rectum); diagnosis of bowel infiltration in accordance with previously published (referenced to	Sensitivity (95% CI): 61% (43 to 76) Specificity (95% CI): 96% (91 to 98) Posterior DIE: Sensitivity (95% CI): 45% (27 to 64) Specificity (95% CI): 100% (98 to 100) PoD: Sensitivity (95% CI): 72% (51 to 88) Specificity (95% CI): 97% (93 to 99) **Hudelist 2011** RVS (rectovaginal septum) endometriosis: Sensitivity (95% CI): 78% (40 to 97) Specificity (95% CI): 100% (97 to 100) Rectosigmoid endometriosis: Sensitivity (95% CI): 90% (74 to 98) Specificity (95% CI): 99% (94 to 100) USL endometriosis: Sensitivity (95% CI): 63% (44 to 80) Specificity (95% CI): 97% (89 to 100) Vaginal wall involvement:	Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? unclear Did all patients receive the same reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	inflammatory drugs) for ≥ 2 years 50/50 **Age:** mean 33 ± 5 years, range 22 to 41 years **Number enrolled:** 50 women **Number available for analysis:** 50 women **Setting:** University Hospital, University of Cagliari **Place of study:** Cagliari, Italy **Period of study:** January 2005 to May 2005 **Guerriero 2008** **Clinical presentation:** pelvic pain in all 88 patients: dyspareunia 40/88, dysmenorrhoea 71/88, infertility 10/88; previous medical treatment for persistent pelvic pain (estrogens, progestins and/or GnRH agonist and non-steroidal anti-inflammatory drugs) for ≥ 2 years 88/88 **Age:** mean 33 ± 5 years, range 20 to 45 years **Number enrolled:** 88 women **Number available for analysis:** 88 women **Setting:** University Hospital, University of Cagliari		Chapron et al., 1998) and described **Reid 2013** TVUS: single examiner; level of expertise and blinding to clinical data not reported **Reid 2014** Sonovaginography: all SVG examinations performed by 2 operators (1 was an expert gynaecological sonologist with experience in diagnosis of DIE; the other was a gynaecological ultrasound fellow supervised by an experienced operator). Same person who performed SVG performed the gynaecological examination and TVS. Operators were not blinded to clinical history **Ribeiro 2008** TRUS: performed by a senior echographer, single operator; unclear	Sensitivity (95% CI): 64% (31 to 89) Specificity (95% CI): 99% (95 to 100) PoD: Sensitivity (95% CI): 76% (53 to 92) Specificity (95% CI): 100% (97 to 100) Bladder endometriosis*: Sensitivity (95% CI): 25% (1 to 81) Specificity (95% CI): 100% (97 to 100) Ovarian endometriosis: Sensitivity (95% CI): 96% (81 to 100) Specificity (95% CI): 96% (90 to 99) **Hudelist 2013** Rectosigmoid endometriosis: Sensitivity (95% CI): 85% (69 to 95) Specificity (95% CI): 96% (90 to 99) **Leon 2014** PoD endometriosis: Sensitivity (95% CI): 89% (71 to 98) Specificity (95% CI): 92% (73 to 99)	Were all patients included in the analysis? Y Could the patient flow have introduced bias? unclear risk **Eskenazi 2001** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' No Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? high concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Place of study: Cagliari, Italy **Period of study:** December 2005 to December 2007 **Guerriero 2014** **Clinical presentation:** chronic pelvic pain 101/202, dyspareunia 51/202, dysmenorrhoea 132/202; previous surgery for pelvic pain 20/202; hormonal treatment at the time of ultrasound examination 43/202 **Age:** mean 34 ± 6 years, range 18 to 52 years **Number enrolled:** 240 women **Number available for analysis:** 202 women **Setting:** University Hospital, Ospedale San Giovanni di Dio, University of Cagliari **Place of study:** Cagliari, Italy **Period of study:** January 2009 to September 2012 **Holland 2010** **Clinical presentation:** dysmenorrhoea 142/201, chronic pelvic pain 104/201, dyspareunia 78/201, infertility 38/201, dyschezia 7/201, cyclical rectal		whether examiners were blinded to clinical data DCBE: performed by a single operator under supervision of a radiologist technician; images were then reviewed by a skilled radiologist **Said 2014** TVUS: performed by an experienced sonographer; unclear whether blinded to clinical data **Savelli 2011** TVUS and DCBE: both performed by 2 groups of physicians specialising in endometriosis with training and expertise in gynaecological imaging studies, who were aware of each patient's history, symptoms and pelvic examination but were blinded to the results of other index tests **Scarella 2013**	Bladder endometriosis*: Sensitivity (95% CI): 20% (1 to 72) Specificity (95% CI): 100% (93 to 100) **Mangler 2013** Rectosigmoid endometriosis: Sensitivity (95% CI): 20% (10 to 34) Specificity (95% CI): 79% (60 to 92) **Menada 2008** RVS (rectovaginal septum) endometriosis (TVUS-BP): Sensitivity (95% CI): 93% (84 to 98) Specificity (95% CI): 90% (70 to 99) RVS (rectovaginal septum) endometriosis (RWC-TVUS): Sensitivity (95% CI): 97% (90 to 100) Specificity (95% CI): 100% (84 to 100) **Pascual 2010** RVS (rectovaginal septum) endometriosis:	test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	bleeding 2/201; single presenting symptom present in 72/201, 2 presenting symptoms in 78/201 and ≥ 3 symptoms in 51/201 **Age:** mean 34.9 ± 6.79 years (95% CI 33.98 to 35.86), range 19 to 51 years **Number enrolled:** 211 women **Number available for analysis:** 201 women **Setting:** University Hospital, King's College Hospital **Place of study:** London, UK **Period of study:** July 2006 to December 2008 **Hudelist 2011** **Clinical presentation:** dysmenorrhoea 111/129, dyspareunia 72/129, dyschezia 39/129, dysuria 6/129, chronic pelvic pain 45/129, subfertility 20/129 **Age:** mean 32.2 ± 5.4 years, range 17 to 44 years **Number enrolled:** 153 women **Number available for analysis:** 129 women **Setting:** 3 tertiary referral service Hospitals: Worthing and Southlands Hospital, Ashford and St Peters		TVUS: all examinations performed by a single experienced examiner; blinding to clinical data not reported **Ubaldi 1998** TVUS: all scans performed by 2 physicians, each with ≥ 3 years' expertise in ultrasound scanning; physicians not told about clinical histories of patients	Sensitivity (95% CI): 89% (67 to 99) Specificity (95% CI): 95% (74 to 100) **Piessens 2014** <u>Bowel endometriosis:</u> Sensitivity (95% CI): 88% (69 to 97) Specificity (95% CI): 93% (84 to 98) <u>Vaginal wall involvement endometriosis:</u> Sensitivity (95% CI): 80% (52 to 96) Specificity (95% CI): 100% (95 to 100) <u>PoD:</u> Sensitivity (95% CI): 88% (73 to 97) Specificity (95% CI): 90% (79 to 97) <u>Bladder endometriosis*:</u> Sensitivity (95% CI): 33% (13 to 59) Specificity (95% CI): 100% (95 to 100) <u>Ovarian endometriosis:</u> Sensitivity (95% CI): 100% (80 to 100) Specificity (95% CI): 93% (84 to 98)	Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Falco 2011** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? highw risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Hospital, Villach Hospital (endometriosis centre) **Place of study:** Villach, Austria; Worthing and Chertsey, UK **Period of study:** not stated **Hudelist 2013** **Clinical presentation:** dysmenorrhoea 116/117, dyspareunia 74/117, dyschezia 31/117, dysuria 9/117, chronic pelvic pain 32/117, subfertility 22/117 **Age:** mean 31.6 ± 6.5 years **Number enrolled:** 142 women **Number available for analysis:** 117 women **Setting:** Department of O&G, Stage III Center for Endometriosis & Pelvic Pain, Wilhelminen Hospital **Place of study:** Vienna, Austria **Period of study:** July 2011 to May 2012 **Leon 2014** **Clinical presentation:** dysmenorrhoea 51/51, dyspareunia 39/51, dyschezia 34/51, chronic pelvic pain 46/51, hematochezia 5/51;			**Piketty 2009** <u>Bowel endometriosis (TVUS):</u> Sensitivity (95% CI): 91% (82 to 96) Specificity (95% CI): 97% (88 to 100) <u>Bowel endometriosis (TRUS):</u> Sensitivity (95% CI): 96% (89 to 99) Specificity (95% CI): 100% (94 to 100) **Reid 2013** <u>RVS (rectovaginal septum) endometriosis:</u> Sensitivity (95% CI): 25% (3 to 65) Specificity (95% CI): 100% (96 to 100) <u>Rectosigmoid endometriosis:</u> Sensitivity (95% CI): 85% (62 to 97) Specificity (95% CI): 91% (83 to 96) <u>USL endometriosis:</u> Sensitivity (95% CI): 40% (12 to 74) Specificity (95% CI): 96% (89 to 99) <u>PoD:</u> Sensitivity (95% CI): 83% (65 to 94)	Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	suspicious bimanual vaginal examination 26/51 **Age:** mean 32.9 ± 4.7 years, range 23 to 43 years **Number enrolled:** 110 women **Number available for analysis:** 51 women **Setting:** Department of Obstetrics and Gynecology, Ultrasound and Human Reproduction Unit of the Indisa Clinic **Place of study:** Santiago, Chile **Period of study:** August 2011 to October 2012 **Mangler 2013** **Clinical presentation:** dysmenorrhoea 73%, bowel symptoms (dyschezia, cyclical constipation, diarrhoea) 68%; overall 97% presented with symptoms; previous surgery for pelvic pain 78%; hormonal treatment 69% **Age:** mean 34 years, range 19 to 51 years **Number enrolled:** 79 women **Number available for analysis:** 79 women **Setting:** University Hospital, Charité Campus Mitte			Specificity (95% CI): 97% (90 to 100) **Reid 2014** <u>RVS (rectovaginal septum) endometriosis:</u> Sensitivity (95% CI): 18% (2 to 52) Specificity (95% CI): 100% (98 to 100) <u>Posterior DIE:</u> Sensitivity (95% CI): 86% (74 to 94) Specificity (95% CI): 92% (87 to 96) <u>Rectosigmoid endometriosis:</u> Sensitivity (95% CI): 88% (75 to 96) Specificity (95% CI): 93% (75 to 100) <u>USL endometriosis:</u> Sensitivity (95% CI): 40% (12 to 74) Specificity (95% CI): 98% (94 to 99) <u>Vaginal wall involvement:</u> Sensitivity (95% CI): 18% (2 to 52) Specificity (95% CI): 99% (97 to 100) <u>PoD:</u> Sensitivity (95% CI): 83% (69 to 92)	Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Fedele 1998** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Place of study: Berlin, Germany **Period of study:** September 2007 to February 2010 **Menada 2008** **Clinical presentation:** dysmenorrhoea 84/90, dyspareunia 68/90, chronic pelvic pain 62/90, infertility 32/90, diarrhoea and/or constipation 61/90, bowel movement pain or cramping 69/90, pain on defecation 32/90, rectal bleeding 16/90, lower back pain 57/90; previous medical treatments for endometriosis 82/90 **Age:** median 32 years, range 18 to 42 years **Number enrolled:** 90 women **Number available for analysis:** 90 women **Setting:** University Hospital, San Martino Hospital, University of Genoa **Place of study:** Genoa, Italy **Period of study:** October 2006 to November 2007 **Pascual 2010** **Clinical presentation:** dyspareunia and/or dysmenorrhoea 39/39, infertility 15/39; previous			Specificity (95% CI): 98% (94 to 100) **Ribeiro 2008** <u>Rectosigmoid endometriosis:</u> Sensitivity (95% CI): 100% (87 to 100) Specificity (95% CI): 90% (55 to 100) **Said 2014** <u>Pelvic endometriosis:</u> Sensitivity (95% CI): 85% (75 to 93) Specificity (95% CI): 81% (68 to 90) **Savelli 2011** <u>Posterior DIE:</u> Sensitivity (95% CI): 85% (74 to 93) Specificity (95% CI): 100% (16 to 100) <u>Rectosigmoid endometriosis:</u> Sensitivity (95% CI): 91% (80 to 97) Specificity (95% CI): 100% (75 to 100) **Scarella 2013** <u>RVS (rectovaginal septum) endometriosis:</u>	If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	treatment for persistent pelvic pain with estrogens, progestins and/or GnRH agonist and non-steroidal anti-inflammatory drugs for ≥ 1 year 39/39 **Age:** mean 35.6 ± 5.7 years, range 25 to 44 years **Number enrolled:** 39 women **Number available for analysis:** 38 women **Setting:** University Hospital, Instituto Universitario Dexeus of Barcelona **Place of study:** Barcelona, Spain **Period of study:** January 2008 to July 2009 **Piessens 2014** **Clinical presentation:** dysmenorrhoea (63%), dyschezia (53%), dyspareunia (44%), infertility (22%), abnormal bleeding (20%), chronic pain (21%), rectal bleeding (8%); past history of endometriosis (72%) **Age:** range 18 to 48 years **Number enrolled:** 205 women **Number available for analysis:** 85 women			Sensitivity (95% CI): 96% (82 to 100) Specificity (95% CI): 100% (88 to 100) <u>DIE:</u> Sensitivity (95% CI): 94% (81 to 99) Specificity (95% CI): 100% (85 to 100) <u>USL endometriosis:</u> Sensitivity (95% CI): 86% (42 to 100) Specificity (95% CI): 100% (93 to 100) <u>Ovarian endometriosis:</u> Sensitivity (95% CI): 97% (83 to 100) Specificity (95% CI): 100% (87 to 100) **Ubaldi 1998** <u>Ovarian endometriosis:</u> Sensitivity (95% CI): 90% (55 to 100) Specificity (95% CI): 97% (92 to 99) *bladder data from the original paper	A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Ferrero 2011** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? No Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Setting: Monash Health, Clayton; Monash University **Place of study:** Clayton Victoria, Australia **Period of study:** November 2009 to September 2011 **Piketty 2009** **Clinical presentation:** dysmenorrhoea, deep dyspareunia, non-cyclical chronic pelvic pain, gastrointestinal symptoms, lower urinary tract symptoms; previous hormonal treatment for endometriosis 134/134, previous surgery for endometriosis 88/134 **Age:** mean 32.1 ± 5.0 years, range 22 to 47 years **Number enrolled:** 134 women **Number available for analysis:** 134 women **Setting:** University Hospital, Université Paris Descartes **Place of study:** Paris, France **Period of study:** January 2005 to July 2007 **Reid 2013** **Clinical presentation:** cyclical pain 70/100, pain requiring strong analgesia				knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	49/100, pain affecting life despite strong analgesia 53/100, pain preventing daily activities 55/100, dyspareunia 56/100, dyschezia 51/100, tenesmus 29/100, cyclical constipation 32/100, cyclical diarrhoea 37/100 (37%), cyclical hematuria 3/100 (3%), cyclical hematochezia 16/100 (16%), constant pain 2/100 (2%), non-cyclical pain 2/100; pain location: left iliac fossa pain 49%, lower abdominal pain 65%, right iliac fossa pain 44%, left upper quadrant pain 7%, epigastric pain 2%, right upper quadrant pain 2% and back pain 2%; median duration of pelvic pain 18 months; history of in vitro fertilisation (13%), irregular menstrual periods (19%), use of contraception (30%), history of infertility (30%) and history of endometriosis (60%) **Age:** mean 32.78 ± 6.28 years; median 33.0 years, range 19 to 48 years **Number enrolled:** 100 women? (see note below) **Number available for analysis:** 100 women **Setting:** 4 university teaching hospitals, tertiary				Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Ghezzi 2005** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	referral centres: Nepean Hospital, Royal Hospital for Women, Royal Prince Alfred Hospital, Liverpool Hospital; 5 private hospitals: Norwest Private Hospital, Hurstville Private Hospital, St. Luke's Private Hospital, Prince of Wales Private Hospital, St. George Private Hospital **Place of study:** NSW, Australia **Period of study:** January 2009 to November 2011 **Reid 2014** **Clinical presentation:** chronic pelvic pain, dysmenorrhoea, dyspareunia, dyschezia; mean duration of pain 39.7 ± 47.5 months; history of infertility 44/220; history of endometriosis 92/220; history of bowel DIE in the past 10/220 **Age:** mean 32.2 ± 7.5 years **Number enrolled:** 220 women **Number available for analysis:** 189 women **Setting:** 4 university teaching hospitals, tertiary referral centres: Nepean Hospital, Royal Hospital for Women, Royal Prince Alfred Hospital, Liverpool Hospital;				knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	5 private hospitals: Norwest Private Hospital, Hurstville Private Hospital, St. Luke's Private Hospital, Prince of Wales Private Hospital, St. George Private Hospital **Place of study:** NSW, Australia **Period of study:** January 2009 to February 2013 **Ribeiro 2008** **Clinical presentation:** symptoms - see Inclusion criteria **Age:** mean 35.8 ± 4.4 years, range 28 to 48 years **Number enrolled:** 37 women **Number available for analysis:** 37 women **Setting:** University Hospital, Santa Casa Medical School, referral centre for endometriosis **Place of study:** São Paulo, Brazil **Period of study:** January 2004 to January 2005 **Said 2014** **Clinical presentation:** dysmenorrhoea 96/142, dyspareunia 72/142, dyschezia 33/142, non-cyclical chronic pelvic pain				Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Goncalves 2010** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	28/142, infertility 37/142, dysuria 5/142 **Age:** median 29 years, range 19 to 46 years **Number enrolled:** 142 women **Number available for analysis:** 125 women **Setting:** University Hospital, El-Shatby Maternity Hospital, Alexandria University **Place of study:** Alexandria University, Egypt **Period of study:** not specified **Savelli 2011** **Clinical presentation:** infertility 30/69, dysmenorrhoea 64/69, dyspareunia 59/69, dyschezia 45/69; nulliparous 49/69, previous surgery for endometriosis 18/69, oestrogen-progestin therapy before surgery 22/69 **Age:** median 33.6 ± 5.9 years **Number enrolled:** 94 women **Number available for analysis:** 69 women **Setting:** university hospital tertiary care referral, S. Orsola-Malpighi Hospital				knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Place of study: Bologna, Italy **Period of study:** January 2004 to December 2007 **Scarella 2013** **Clinical presentation:** infertility 29/57, moderate to severe pelvic pain 50/57, dyspareunia 30/57; nulliparous 30/57 **Age:** women of reproductive age, age range or mean not specified **Number enrolled:** 100 women **Number available for analysis:** 57 women **Setting:** 2 university hospitals: Institute of Maternal and Child Research, Iniversity of Chilie; Center for Human Reproduction, Valpraiso University **Place of study:** Santiago and Valparaiso, Chilie **Period of study:** Sepember 2011 to September 2012 **Ubaldi 1998** **Clinical presentation:** infertility, chronic pelvic pain and/or adnexal masses **Age:** range 21 to 41 years				Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Grasso 2010** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number enrolled: 133 women **Number available for analysis:** 133 women **Setting:** university hospital: Centre for Reproductive Medicine of the Dutch-speaking Free University of Brussels **Place of study:** Brussels, Belgium **Period of study:** February 1994 to April 1995 **<u>Inclusion Criteria</u>** **Abrao 2007** **Study population:** patients with clinically suspected endometriosis **Selection criteria:** not specified **Bazot 2009** **Study population:** women referred with clinical evidence of pelvic endometriosis **Selection criteria:** not specified **Bergamini 2010** **Study population:** women scheduled for surgery because of signs and symptoms of severe				Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	posterior deep infiltrating endometriosis **Selection criteria:** not specified **Dessole 2003** **Study population:** women scheduled for laparotomy or laparoscopy because rectovaginal endometriosis is suspected on the basis of patient history and clinical examination **Selection criteria:** not specified **Eskenazi 2001** **Study population:** women scheduled to undergo laparoscopy or laparotomy for pelvic pain, infertility, tubal ligation or adnexal/uterine masses **Selection criteria:** not specified **Falco 2011** **Study population:** patients scheduled for laparoscopy with ≥ 1 symptom suggestive for the presence of endometriosis **Selection criteria:** not specified **Fedele 1998**				does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Guerriero 1996a** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Study population: patients scheduled for laparoscopy or laparotomy for pelvic endometriosis, suspected on basis of history and objective findings (not specified) **Selection criteria:** not specified **Ferrero 2011** **Study population:** patients referred to the endometriosis centre **Selection criteria:** suspicion of deep pelvic endometriosis (on the basis of gynaecological symptoms and vaginal examination); presence of gastrointestinal symptoms that might be caused by bowel endometriosis; reproductive age; desire to undergo complete surgical excision of the endometriosis. **Ghezzi 2005** **Study population:** premenopausal women with adnexal mass or with clinical signs suggestive of pelvic endometriosis who were scheduled for laparoscopic surgery **Selection criteria:** not specified				Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Goncalves 2010 **Study population:** patients submitted to laparoscopy on suspicion of endometriosis **Selection criteria:** scheduled to undergo surgery for therapeutic management of endometriosis. **Grasso 2010** **Study population:** patients with clinical suspicion of pelvic endometriosis **Selection criteria:** not specified **Guerriero 1996a** **Study population:** women scheduled for laparoscopy or laparotomy for a persistent ovarian mass **Selection criteria:** premenopausal, non-pregnant women **Guerriero 1996b** **Study population:** women who were submitted to laparoscopy or laparotomy because of the presence of a persistent adnexal mass **Selection criteria:** premenopausal, non-pregnant women				does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Guerriero 1996b** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Guerriero 2007 **Study population:** women scheduled for laparoscopic surgery for rectovaginal endometriosis, suspected on the basis of patient history of pelvic pain and/or clinical examination **Selection criteria:** not specified **Guerriero 2008** **Study population:** women scheduled for laparoscopic surgery for clinically suspected endometriosis on the basis of patient history of pelvic pain and/or clinical examination **Selection criteria:** not specified **Guerriero 2014** **Study population:** all premenopausal women with clinical suspicion of deep endometriosis who were scheduled for surgery in our department **Selection criteria:** reproductive age, clinically suspected endometriosis; exclusion criteria: abdominal mass larger than 10 cm with distortion of pelvic anatomy, emergency laparoscopy due				Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	to acute pain, 2D-US or 3D-US not performed, insufficient description at surgery, pregnancy at time of diagnosis, surgery longer than 30 days after ultrasound **Holland 2010** **Study population:** women with clinically suspected or proven pelvic endometriosis **Selection criteria:** premenopausal women with clinical suspicion of endometriosis awaiting diagnostic laparoscopy; women diagnosed with pelvic endometriosis at diagnostic laparoscopy awaiting operative treatment; age ≥ 16 years; ability to provide informed consent. **Hudelist 2011** **Study population:** 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				does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Guerriero 2007** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? uncelar risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	referred (coming to the pain clinic without seeing any gynaecologist before this time for their current problems) **Selection criteria:** premenopausal women **Hudelist 2013** **Study population:** women attending pelvic pain clinic with suspected endometriosis and scheduled for laparoscopy on the basis of clinical examination and TVS findings **Selection criteria:** not specified **Leon 2014** **Study population:** 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) **Selection criteria:** clinical suspicion of DIE, patient's acceptance to undergo				A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	transvaginal sonography. Exclusion criteria: concomitant cancer, pregnancy, or pelvic inflammatory process; surgery performed at a centre other than the recruitment centre; choice of medical treatment instead of surgery; patient withdrawal before surgery **Mangler 2013** **Study population:** patients with suspected/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 **Selection criteria:** not specified **Menada 2008** **Study population:** women with suspected rectovaginal endometriosis on the basis of pain symptoms and/or gynaecological examination **Selection criteria:** not specified				Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Guerriero 2008** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Pascual 2010 **Study population:** patients with clinically suspected endometriosis based on patient history of pelvic pain and/or clinical examination **Selection criteria:** not specified **Piessens 2014** **Study population:** patients with clinically suspected endometriosis referred to TVUS **Selection criteria:** not specified **Piketty 2009** **Study population:** patients suffering from pelvic pain (alone or associated with infertility) who underwent complete surgical exeresis of deeply infiltrating endometriosis (DIE), which was suspected in all cases preoperatively (questioning, clinical examination, imaging) **Selection criteria:** not specified **Reid 2013** **Study population:** women with a history of chronic pelvic pain and/or				do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	endometriosis and scheduled for operative laparoscopy **Selection criteria:** pelvic pain, defined as chronic if it persisted for longer than 3 months and could be constant or intermittent, cyclical or non-cyclical in nature; 4 types of pelvic pain included: cyclical pain during menstruation (dysmenorrhoea), deep dyspareunia, dyschezia and non-cyclical pelvic pain; only women of reproductive age. **Reid 2014** **Study population:** women who presented to pelvic pain clinic with symptoms suggestive of endometriosis **Selection criteria:** reproductive age, history of chronic pelvic pain ± history of endometriosis, laparoscopy within 6 months of gel SVG examination. **Ribeiro 2008** **Study population:** patients with clinically suspected deeply infiltrating endometriosis (DIE) referred to gynaecological endoscopy and endometriosis clinic				B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Guerriero 2014** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability:

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Selection criteria: dysmenorrhoea or dyspareunia associated with ≥ 1 of the following signs: pouch of Douglas (POD) tenderness or nodules, pain caused by cervical mobilisation, pain during POD mobilisation; intestinal symptoms alone not considered inclusion criteria. **Said 2014** **Study population:** women with any symptoms suggestive of endometriosis who were booked for laparoscopy **Selection criteria:** reproductive age; pain in the lower abdomen or pelvis for ≥ 6 months; infertility; regular menstrual cycle; no medications for infertility or pelvic pain treatment in the preceding 3 months; availability of complete past medical, social, obstetrical and gynaecological history; normal size ovary on TVS. **Savelli 2011** **Study population:** patients with results of pelvic examination or symptoms				Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	suggestive of DIE of the posterior compartment **Selection criteria:** symptoms or examination findings indicative of DIE of the posterior compartment **Scarella 2013** **Study population:** women with chronic pelvic pain and/or suspected endometriosis **Selection criteria:** not specified **Ubaldi 1998** **Study population:** patients who had been referred for diagnostic or operative laparoscopy for infertility, chronic pelvic pain and/or adnexal masses **Selection criteria:** non-pregnant premenopausal women **Exclusion Criteria** **Abrao 2007** Not reported **Bazot 2009** Not reported **Bergamini 2010**				interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? high risk **Holland 2010** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Not reported **Dessole 2003** Not reported **Eskenazi 2001** acute conditions such as ectopic pregnancy, evaluation of endometrial or ovarian cancer, treatment of already diagnosed endometriosi **Falco 2011** Not reported **Fedele 1998** previous surgery for rectovaginal endometriosis **Ferrero 2011** previous bilateral ovariectomy; previous barium radiological examination or other examination for diagnosis of bowel endometriosis; previous bowel surgery (except appendectomy); previous episodes suggestive of intolerance to iodinated contrast medium; renal or hepatic failure; psychiatric disorders **Ghezzi 2005**				B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Yes

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	previous surgical intervention on adnexa or uterus; history of breast, gastrointestinal tract or genitourinary tract malignancy; history of infertility without symptoms or signs of endometriosis; clinical or ultrasound suspicion of malignancy **Goncalves 2010** any prior bowel surgery **Grasso 2010** Not reported **Guerriero 1996a** Not reported **Guerriero 1996b** Not reported **Guerriero 2007** Not reported **Guerriero 2008** Not reported **Guerriero 2014** Not reported **Holland 2010**				Could the reference standard, its conduct, or its interpretation have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Hudelist 2011** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	women who could not undergo TVUS scan; women who became pregnant whilst awaiting surgery **Hudelist 2011** Not reported **Hudelist 2013** Not reported **Leon 2014** concomitant cancer, pregnancy, or pelvic inflammatory process; surgery performed at a centre other than the recruitment centre; choice of medical treatment instead of surgery; patient withdrawal before surgery **Mangler 2013** Not reported **Menada 2008**a patients who were virgins or who had any type of genital malformation that made physical examination or TVS impossible; previous surgical excision of bowel endometriosis **Pascual 2010**				Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Not reported **Piessens 2014** Not reported **Piketty 2009** Not reported **Reid 2013** Not reported **Reid 2014** malignancy, menopause, pregnancy **Ribeiro 2008** previous surgical therapy for intestinal endometriosis and previous use of medical therapy for endometriosis **Said 2014** virginity, pregnancy, ovarian cyst of any type on TVS, genital malformation that made examination or TVS impossible, history of gynaecological cancer or previous abdominal or pelvic surgery, premature ovarian failure, large uterine masses **Savelli 2011** Not reported				Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Hudelist 2013** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided?According to the CSR 'Was a two-gate design avoided?' Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Scarella 2013 postmenopausal patients, patients with previous surgery of colon/sigmoid, patients with known causes of pelvic pain **Ubaldi 1998** Not reported				Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s)

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Leon 2014** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? high risk **Mangler 2013** A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Menada 2008 A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Were there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the patient flow have introduced bias? Low risk **Pascual 2010** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Piessens 2014** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? high risk **Piketty 2009** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? No Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unlcear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Reid 2013** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Reid 2014** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Ribeiro 2008** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? No Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Said 2014** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? No Could the selection of patients have introduced bias? high risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Savelli 2011** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? high risk **Scarella 2013** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability:

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					interpretation have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? high risk **Ubaldi 1998** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? high risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk

1

G.7 Review question: Diagnosis – Biomarkers: CA-125

3 **What is the accuracy of erum CA-125 in diagnosing endometriosis?**

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation Nisenblat, Vicki, Bossuyt, M. M. Patrick, Shaikh, Rabia, Farquhar, Cindy, Jordan, Vanessa, Scheffers,	**Condition** Study participants included reproductive-aged women with suspected endometriosis based on clinical symptoms, pelvic examination or both, who	**Tests** CA-125 > 35 IU/ml only **Barbati 1994** **Index test:** CA-125	**Methods** **Barbati 1994** serum levels of CA-125 were measured by immunoradiometric 'one step' sandwich	Results **Barbati 1994** Sensitivity (95% CI): 44% (22 to 69) Specificity (95% CI): 89% (71 to 98)	**Limitations** AMSTAR Checklist 1. Was an 'a priori' design provided? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
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 **Ref Id** 496572 **Country/ies where the study was carried out** New Zealand Study type **Cochrane Review** **Aim of the study** To evaluate blood biomarkers as replacement tests for diagnostic surgery and as triage tests to inform decisions on surgery for endometriosis. **Study dates** 2016	undertook the index test as well as the reference standard. **Sample size** N=141 studies but only 24 studies relevant to the present review were included **Characteristics** **Barbati 1994** **Clinical presentation:** Inertility or pelvic pain **Age:** range 23-41 years (endometriosis group), 16-55 years (controls) **Number of participants enrolled:** 45 women **Number of participants available for analysis:** 45 women (all in mid-follicular cycle phase, day 8-12) **Setting:** Institute of O&G, University of Rome 'La Sapienza' **Place of study:** Rome, Italy **Period of study:** not stated **Bilibio 2014** **Clinical presentation:** endometriosis group - infertility, pelvic pain or both; other causes of infertility were excluded by	**Reference test:** laparoscopy/laparotomy N = 45 (100%) **Bilibio 2014** **Index test:** CA-125 **Reference test:** laparoscopy n = 97 (100%) + histopathology **Chen 1998** **Index test:** CA-125 **Reference test:** laparoscopy N = 157 (100%) + histology **Colacurci 1996** **Index test:** CA-125 **Reference test:** laparoscopy N = 40 (100%) **Fedele 1989** **Index test:** CA-125 **Reference test:** laparoscopy N = 264 (100%) + histology **Ferreira 1994** **Index test:** CA-125 **Reference test:** laparoscopy/laparotomy N = 54 (100%) + histology **Franchi 1993** **Index test:** CA-125 **Reference test:** laparoscopy/laparotomy N = 120 (100%)	assay (IRMA CA-125 II K, Sorin Biomedica, Italy); minimal detectable concentration 1.4 U/ml; sample processing and experiments are described in details Bilibio 2014 CA-125 was analysed with Roche Diagnostics **Chen 1998** serum CA-125 was determined by immunoradiometric assay ELISA-CA 125 II kit (GIF-SUR-YVETTE CEDEX, France); no other details provided **Colacurci 1996** serum CA-125 levels were measured by immunoradiometric 'two-step method' (IRMA-mat, Byk-Stangtee Diagnostic GmbH&Co Kgy, Dietzenbach); sample processing and experiments are described in details	**Bilibio 2014** Sensitivity (95% CI): 27% (17 to 40) Specificity (95% CI): 97% (85 to 100) **Chen 1998** Sensitivity (95% CI): 61% (52 to 69) Specificity (95% CI): 88% (68 to 97) **Colacurci 1996** Sensitivity (95% CI): 44% (22 to 69) Specificity (95% CI): 91% (71 to 99) **Fedele 1989** Sensitivity (95% CI): 15% (8 to 23) Specificity (95% CI): 100% (93 to 100) **Ferreira 1994** Sensitivity (95% CI): 4% (0 to 22) Specificity (95% CI): 89% (65 to 99) **Franchi 1993** Sensitivity (95% CI): 51% (34 to 68) Specificity (95% CI): 87% (78 to 93) **Gagne 2003** Sensitivity (95% CI): 20% (15 to 27) Specificity (95% CI): 92% (87 to 95)	2. Was there duplicate study selection and data extraction? Y 3. Was a comprehensive literature search performed? Y 4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? No 5. Was a list of studies (included and excluded) provided? Y 6. Were the characteristics of the included studies provided? Y 7. Was the scientific quality of the included studies assessed and documented? Y 8. Was the scientific quality of the included studies used appropriately in formulating conclusions? Y 9. Were the methods used to combine the findings of studies appropriate? Y 10. Was the likelihood of publication bias assessed? No 11. Was the conflict of interest included? Y Where there is a high/unclear risk regarding applicability it is due to a two-gate design: according to Nisenblat et al. 2016

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Source of funding Internal sources Cochrane Gynaecology and Fertility Group, University of Auckland, New Zealand. Technical support The Robinson Institute, University of Adelaide, Australia. Access to academic resources External sources No sources of support supplied	hysterosalpingography, semen analysis, and measurements of serum FSH and TSH levels on the 3rd day of the menstrual cycle **Age:** mean age 33.34 ± 4.66 and 33.67 ± 7.16 years (endometriosis group); 33.03 ± 4.42 years (control group) **Number of participants enrolled:** 97 women **Number of participants available for analysis:** 97 women (all in luteal phase of menstrual cycle) **Setting:** Department of O&G, Universidade Federal do Rio Grande do Sul, Hospital de Clínicas de Porto Alegre **Place of study:** Porto Alegre, Brazil **Period of study:** not specified **Chen 1998** **Clinical presentation:** not specified **Age:** mean age 30.8 ± 7.3 years, range 15-45 **Number of participants enrolled:** 157 women **Number of participants available for analysis:** 155	**Gagne 2003** **Index test:** CA-125 **Reference test:** laparoscopy/laparotomy N = 368 (100%) **Guerriero 1996** **Index test:** CA-125 **Reference test:** laparoscopy/laparotomy + histology **Hallamaa 2012** **Index test:** CA-125 **Reference test:** laparoscopy N = 175 (100%) + histology **Harada 2002** **Index test:** CA-125 **Reference test:** laparoscopy/laparotomy N = 123 (100%) **Hornstein 1995** **Index test:** CA-125 **Reference test:** laparoscopy N = 123 (100%) **Koninckx 1996** **Index test:** CA-125 **Reference test:** laparoscopy N = 55 (100%) **Kurdoglu 2009** **Index test:** CA-125 **Reference test:** laparoscopy/laparoto	**Fedele 1989** serum CA-125 was measured by immunoradiometric assay (Sorin Biomedica, Saluggia VC, Italy) **Ferreira 1994** serum CA-125 was measured by ELISA (Cobas Core CA-125 II, EIA Roche 1992); assay sensitivity < 1 U/ml; procedure and sample handling described **Franchi 1993** serum CA-125 levels assessed by radioimmunoassay; sample processing and laboratory technique not described **Gagne 2003** serum CA-125 level was determined by using a one step-sandwich radioimmunoassay (Fujirebio America Inc.) with assay sensitivity 0.4 U/ml; sample handling and laboratory procedure	**Guerriero 1996** Sensitivity (95% CI): 59% (39 to 76) Specificity (95% CI): 79% (68 to 88) **Hallamaa 2012** Sensitivity (95% CI): 38% (30 to 47) Specificity (95% CI): 100% (93 to 100) **Harada 2002** Sensitivity (95% CI): 49% (38 to 59) Specificity (95% CI): 100% (85 to 100) **Hornstein 1995** Sensitivity (95% CI): 23% (14 to 34) Specificity (95% CI): 94% (83 to 99) **Koninckx 1996** Sensitivity (95% CI): 50% (29 to 71) Specificity (95% CI): 87% (70 to 96) **Kurdoglu 2009** Sensitivity (95% CI): 57% (47 to 67) Specificity (95% CI): 92% (75 to 99) **Lanzone 1991** Sensitivity (95% CI): 53% (42 to 64) Specificity (95% CI): 87% (72 to 96)	these are studies with two sets of inclusion criteria with respect to **Clinical presentation:** and one set of inclusion criteria with respect to reference standard (the participants with or without a clinical suspicion of endometriosis scheduled for abdominal surgery). <u>QUADAS 2</u> **Barbati 1994** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	women (all in luteal phase of menstrual cycle) **Setting:** tertiary teaching hospital Keelung Chang Gung Memorial Hospital **Place of study:** Taiwan **Period of study:** January 1993 - January 1995 **Colacurci 1996** **Clinical presentation:** infertility **Age:** mean age 31.2 ± 4.5 years (endometriosis group), 32.6 ± 6.1 years and 27.0 ± 5.8 years (controls) **Number of participants enrolled:** 45 women **Number of participants available for analysis:** 40 women, all in mid-follicular cycle phase (day 7-10) **Setting:** Institute of O&G, School of Medicine, 2nd University of Naples **Place of study:** Naples, Italy **Period of study:** not stated **Fedele 1989** **Clinical presentation:** not specified **Age:** mean 30.9 years (endometriosis), 31.2 years (controls)	my N = 127 (100%) + histopathology **Lanzone 1991** **Index test:** CA-125 **Reference test:** laparoscopy N = 270 (100%) **Maiorana 2007** **Index test:** CA-125 **Reference test:** laparoscopy N = 86 (100%) **Martinez 2007** **Index test:** CA-125 **Reference test:** laparoscopy N = 119 (100%) **Mohamed 2013** **Index test:** CA-125 **Reference test:** laparoscopy + histology N = 60 (100%) **Molo 1994** **Index test:** CA-125 **Reference test:** laparoscopy N = 35 (100%) + histology **Muscatello 1992** **Index test:** CA-125 **Reference test:** laparoscopy N = 119 (100%) **Patton 1986** **Index test:** CA-125	described in details. The bootstrap method validation was performed by drawing 200 replicate samples with replacement from the original data set **Guerriero 1996** serum Ca-125 levels assessed by immunoradiometric assay (CIS Bio International, Gif sur Yvette, France), limit of detection 0.5 U/ml; sample processing and laboratory technique not described **Hallamaa 2012** CA-125 concentrations were analysed by ELISA analysis (Fujirebio Diagnostics inc, Malvern, PA, USA) according to the manufacturer's instructions Herada 2002 serum CA-125 levels were measured by enzyme	**Maiorana 2007** Sensitivity (95% CI): 67% (54 to 78) Specificity (95% CI): 94% (71 to 100) **Martinez 2007** Sensitivity (95% CI): 47% (30 to 65) Specificity (95% CI): 97% (90 to 100) **Mohamed 2013** Sensitivity (95% CI): 70% (51 to 85) Specificity (95% CI): 83% (65 to 94) **Molo 1994** Sensitivity (95% CI): 0% (0 to 18) Specificity (95% CI): 94% (70 to 100) **Muscatello 1992** Sensitivity (95% CI): 53% (42 to 64) Specificity (95% CI): 87% (72 to 96) **Patton 1986** Sensitivity (95% CI): 14% (5 to 29) Specificity (95% CI): 93% (85 to 98) **Somigliana 2004** Sensitivity (95% CI): 27% (15 to 42) Specificity (95% CI): 97% (85 to 100)	Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number of participants enrolled: 264 women **Number of participants available for analysis:** 154 women (menstrual cycle phase not specified) **Setting:** Tteaching hospital, Luigi Mangiagalli, University of Milan **Place of study:** Milan, Italy **Period of study:** October 1985 - July 1987 **Ferreira 1994** **Clinical presentation:** infertility, not specified otherwise **Age:** median 30 years, range 20-50 years **Number of participants enrolled:** 54 women **Number of participants available for analysis:** 41 women (menstrual cycle phase not specified) **Setting:** University hospital, Federal University of Minas Gerais **Place of study:** Belo Horizonte, Brazil **Period of study:** January 1992 - June 1993 **Franchi 1993** **Clinical presentation:** pelvic mass, not specified	**Reference test:** laparoscopy + histology N = 113 (100%) **Somigliana 2004** **Index test:** CA-125 **Reference test:** laparoscopy N = 80 (100%) **Vigil 1999** **Index test:** CA-125 **Reference test:** laparoscopy N = 49 (100%) + histology **Yang 1994** **Index test:** CA-125 **Reference test:** laparoscopy n = 42 (100%) **Zeng 2005** **Index test:** CA-125 **Reference test:** laparoscopy/laparotomy N = 58 (100%)	immunoassay (TFB Co,Tokyo, Japan) and were expressed in arbitrary units based on a primary reference standard **Hornstein 1995** serum CA-125 concentrations were determined by immunoradiometric assay (Centocor, Malvern, PA, USA): older assay and the new, a second-generation assay, which utilises M-II murine monoclonal OC125 antibody **Koninckx 1996** A-125 assay by second generation IRMA kit (CA-125 II, Centocor, Malvern, Pa); all the samples assayed in duplicate using kits from the same production batch **Kurdoglu 2009** Details of the index test procedure not reported **Lanzone 1991**	**Vigil 1999** Sensitivity (95% CI): 44% (30 to 60) Specificity (95% CI): 67% (9 to 99) **Yang 1994** Sensitivity (95% CI): 36% (19 to 56) Specificity (95% CI): 86% (57 to 98) **Zeng 2005** Sensitivity (95% CI): 44% (28 to 62) Specificity (95% CI): 82% (60 to 95)	Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Bilibio 2014** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided? Y Did the study avoid inappriate exclusions? Y Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Age: median age 34 years, range 20-51 years (endometriosis); median age 32 years, range 27-42 years (controls) **Number of participants enrolled:** 120 women **Number of participants available for analysis:** 46 women (cycle phase not specified) **Setting:** Department of O&G, University of Pavia, 2nd School of Medicine **Place of study:** Varese, Italy **Period of study:** June 1991 - December 1992 **Gagne 2003** **Clinical presentation:** infertility (7% controls, 16% cases); pain (19% controls, 33% cases); pelvic mass (8% controls, 13% cases); fibroids (9% controls, 15% cases); menorrhagia (2% controls, 4% cases); tubal ligation (60% controls, 25% cases); hysterectomy (19% controls, 32% cases); diagnostic laparoscopy (20% controls, 43% cases); history of endometriosis (3% controls, 16% cases) **Age:** random sampling from a population with		serum CA-125 levels measured with radioimmunoassay (CIS Diagnostici); all samples from the same patient were assayed at the same time **Maiorana 2007** serum CA-125 levels were measured by enzyme immunoassay and were expressed in arbitrary units based on a primary reference standard; no other information provided **Martinez 2007** serum CA-125 levels were measured by enzyme immunoassay and were expressed in arbitrary units based on a primary reference standard; no other information provided. Serum CA-125 level performed using a commercially		do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	mean age of 37.3 ± 6.4 years **Number of participants enrolled:** 368 women **Number of participants available for analysis:** 368 women (in luteal phase of menstrual cycle) **Setting:** biotech firm - MetrioGene BioSciences (a subsidiary of PROCREA BioSciences) **Place of study:** Montreal, Canada **Period of study:** July 1997 - May 2001 **Guerriero 1996** **Clinical presentation:** pelvic mass - 100%, symptoms not specified **Age:** range 20-49 years **Number of participants enrolled:** 101 women **Number of participants available for analysis:** 101 women (only moderate-severe endometriosis included; all in follicular cycle phase) **Setting:** Department of O&G, University of Cagliari **Place of study:** Cagliari, Italy		available chemiluminescent microparticle immunoassay (ARCHITECT CA-125 II Abbott Diagnositics, Spain) with assay sensitivity of < 1.0 IU/ml **Mohamed 2013** CA-125 was measured by ELISA kit for Can-Ag CA-125 (Fujirebio Diagnostics, Inc, Goteborg, Sweden) according to manufacturer instructions (expected value 5.06–47.9 U/ml) **Molo 1994** plasma concentrations of CA-125 were measured by radioimmunoassay (Contocor Inc, Malvern, PA) **Muscatello 1992** serum concentration of CA-125 measured by using a commercially available		interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Chen 1998** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?'Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Period of study: November 1993 - October 1994 **Hallamaa 2012** **Clinical presentation:** endometriosis - not specified; controls - women requesting tubal ligation; hormonal medication was used by 78 (43.3%) women **Age:** mean age 34 years, range 18-48 years **Number of participants enrolled:** 180 women **Number of participants available for analysis:** 175 (7 in menstrual, 32 in proliferative and 60 in secretory cycle phase; 61 had inactive/atrophic endometrium) **Setting:** 2 central hospitals and 2 university central hospitals **Place of study:** Turku, Finland **Period of study:** October 2005 - October 2007 **Harada 2002** **Clinical presentation:** not specified **Age:** mean age 35.4 ± 6.7 years, range 21-52 years		radioimmunoassay (CIS Diagnostici); all assays were performed in duplicate; concentration assessed with a standard curve; sample handling described **Patton 1986** serum CA-125 levels were measured using radioimmunoassay (RIA); sample handling and laboratory techniques not described, but referenced to a primary source (referenced to the original source) **Somigliana 2004** serum level of CA-125 assessed by using a commercially available chemiluminescent immunometric assay (Roche Diagnostics GmbH, Germany) with assay sensitivity 0.6 IU/ml; serum IL-6 levels assessed		B. Concerns regarding applicability: Athere concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number of participants enrolled: 123 women **Number of participants available for analysis:** 123 women (menstrual cycle phase not specified) **Setting:** Department of Reproductive Medicine, Tokyo Medical and Dental University Hospital **Place of study:** Tokyo, Japan **Period of study:** not stated **Hornstein 1995** **Clinical presentation:** not specified **Age:** not specified; all patients had menstrual cycles; implies reproductive age **Number of participants enrolled:** 123 women **Number of participants available for analysis:** 123 women (in follicular phase of menstrual cycle) **Setting:** 2 teaching hospitals: Fertility Unit of Brigham and Women's Hospital and the Reproductive Endocrine/Infertility Service of the Cooper Hospital University Medical Center		by using 2 methods: a commercially available ELISA kit (R&D Systems, Inc, USA) with assay sensitivity 0.7 pg/ml and a sequential immunometric assay (Diagnostic Prod Corp, Medical Systems, Italy); sample handling described **Vigil 1999** CA-125 levels analysed by the IRMA-COUNT OM-MA method; sample handling and laboratory technique not described **Yang 1994** CA-125 was measured by emission immunoassay kit (Syntron Biotech Co, USA) according to manufacturers instructions with a lower limit of detection of 5000 U/l; sample handling and laboratory		knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Colacurci 1996** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?'Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Place of study: Boston, MA, USA and Camden, NJ, USA **Period of study:** not stated **Koninckx 1996** **Clinical presentation:** infertility (n = 33), pain (n = 13), infertility + pain (n = 6), hydrosalpinx (n = 1), ovarian cyst (n= 2) **Age:** range 20-45 years (personal communication with the author) **Number of participants enrolled:** 61 women **Number of participants available for analysis:** 55 women (only DIE, endometrioma and severe pelvic adhesions included; all in menstrual, follicular and early luteal phase of menstrual cycle) **Setting:** division of endoscopic surgery, University Hospital Gasthiusberg, University of Leuven **Place of study:** Leuven, Belgium **Period of study:** not stated **Kurdoglu 2009** **Clinical presentation:** indications for surgery:		technique described **Zeng 2005** serum CA-125 was determined by chemiluminescence assay; sample handling and laboratory technique not described		Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s)

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	suspected pelvic and ovarian endometriosis, infertility, adnexal cystic mass, chronic pelvic pain, desire for sterilisation **Age:** mean age 31.12 ± 5.97 years (endometriosis group), 33.46 ± 9.48 years (controls) **Number of participants enrolled:** 179 participants **Number of participants available for analysis:** 127 participants (cycle phase not specified) **Setting:** Department of Obstetrics and Gynecology, Gazi University School of Medicine **Place of study:** Ankara, Turkey **Period of study:** January 2002 - March 2005 **Lanzone 1991** **Clinical presentation:** pelvic pain, infertility or both **Age:** mean age 30 ± 6.5 years, range 19-44 years (endometriosis group), 30 ± 6.9 years, range 19-41 years (controls) **Number of participants enrolled:** 270 participants **Number of participants available for analysis:** 119				Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Fedele 1989** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	participants (all in luteal cycle phase) **Setting:** Department of O&G, Universita Catolica del Sacro Cuore **Place of study:** Rome, Italy **Period of study:** January 1987 - December 1988 **Maiorana 2007** **Clinical presentation:** In endometriosis group: dysmenorrhoea - 52%, dyspareunia - 26%, asymptomatic - 22%; controls - ovarian cysts **Age:** mean age 33.6 ± 7.3 years, range 21-54 years **Number of participants enrolled:** 86 women **Number of participants available for analysis:** 86 women (in follicular phase of menstrual cycle) **Setting:** obstetrics and gynaecology units, Civic Hospital **Place of study:** Paleromo, Italy **Period of study:** not stated **Martinez 2007** **Clinical presentation:** indications for laparoscopy were pelvic pain (n = 5),				Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?'Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	infertility (n = 11), tubal sterilisation (n = 37), myomas (n = 16), suspicion of endometrioma (n = 33) and other benign ovarian pathologies (n = 26) **Age:** reproductive age **Number of participants enrolled:** 128 women **Number of participants available for analysis:** 119 women (all in follicular cycle phase) **Setting:** Department of O&G, Hospital Universitario Dr Peset **Place of study:** Valencia, Spain **Period of study:** February 2003 - February 2005 **Mohamed 2013** **Clinical presentation:** endometriosis group: chronic pelvic pain - 30 women, dysmenorrhoea - 26 women, history of PID - 7 women; controls: chronic pelvic pain - 2 women, dysmenorrhoea - 9 women, history of PID - 5 women **Age:** range 18-40 years **Number of participants enrolled:** 60 women **Number of participants available for analysis:** 60				Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Ferreira 1994** A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	women (all in in follicular phase of menstrual cycle) **Setting:** Cytogenetic and Endoscopy Unit, Department O&G, Zagazig University Hospital **Place of study:** Zagazig, Egypt **Period of study:** April 2008 - August 2010 **Molo 1994** **Clinical presentation:** infertility **Age:** reproductive age **Number of participants enrolled:** 35 women **Number of participants available for analysis:** 35 women (all in late proliferative phase - mid-cycle phase) **Setting:** Department of O&G, Rush Medical College and Rush-Presbyterian-St Luke's Medical Centre **Place of study:** Chicago, IL **Period of study:** not specified **Muscatello 1992** **Clinical presentation:** infertility, pelvic pain or both				Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Age: mean age 30 ± 6 years, range 19-41 years (endometriosis) and 29 ± 5 years, range 19-44 years (controls) **Number of participants enrolled:** 119 women **Number of participants available for analysis:** 119 women (all in luteal cycle phase) **Setting:** Department of O&G, Universiti Cattolica, S. Cuore **Place of study:** Rome, Italy **Period of study:** January 1089 - February 1990 **Patton 1986** **Clinical presentation:** indications for surgery: infertility - 44%, pain - 10%, elective sterilisation - 43%, premature ovarian failure - 2.6% **Age:** mean 30.5 years, range 16-48 years **Number of participants enrolled:** 113 women **Number of participants available for analysis:** 113 women (menstrual cycle phase not specified)				review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Setting: Department of O&G, Mayo Clinic, tertiary care centre **Place of study:** Rochester, Minnesota **Period of study:** January 1985 - June 1985 **Somigliana 2004** **Clinical presentation:** endometriosis group: not specified, other concomitant pathologies (fibroids, benign ovarian masses) - 14/45; control group: the main diagnoses were PID - 6/35, ovarian cysts - 19/35, myomas - 2/35, normal pelvis in patients with infertility/ pelvic pain - 5/35 **Age:** mean age 32.0 ± 4.2 years (endometriosis group), 32.6 ± 6.4 years (controls) **Number of participants enrolled:** 80 women **Number of participants available for analysis:** 80 women (11 in menstrual, 12 in peri-ovulatory, 23 in luteal cycle phase; for 27 participants cycle phase was not determined) **Setting:** an academic department specialising in gynaecologic laparoscopy -				**Franchi 1993** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Department of O&G, Clinica L.Mangiagalli, University of Milano **Place of study:** Milan, Italy **Period of study:** October 2002 - January 2003 **Vigil 1999** **Clinical presentation:** chronic pelvic pain, dysmenorrhoea, infertility **Age:** mean age 28.16, range 16-41 years **Number of participants enrolled:** 49 women **Number of participants available for analysis:** 49 women (different phases of menstrual cycle, not specified) **Setting:** Research Center of Reproductive Health at the Pontificia Catholic University Chile **Place of study:** Santiago, Chile **Period of study:** not provided **Yang 1994** **Clinical presentation:** infertility - 40, suspected endometriosis - 2 **Age:** mean age 31.36 years, range 24-39 years				Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number of participants enrolled: 42 participants **Number of participants available for analysis:** 42 participants (all in luteal cycle phase) **Setting:** Chang Zheng Hospital, Second Military Medical College **Place of study:** Shanghai, China **Period of study:** July 1992 - December 1992 **Zeng 2005** **Clinical presentation:** infertility or pelvic pain **Age:** mean age 33 ± 4 years, range 26-40 years (endometriosis), 32 ± 4 years, range 25-39 years (controls) **Number of participants enrolled:** 58 women **Number of participants available for analysis:** 58 women (31 women in follicular and 27 women in luteal cycle phase) **Setting:** Department of O&G, Third Xiangya Hospital, Central South University **Place of study:** Changsha, China				Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Gagne 2003** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' No Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? high concern Index Test

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Period of study: March 2003 - February 2004 **Inclusion Criteria** **Barbati 1994** women undergoing laparotomy or diagnostic laparoscopy for infertility or pelvic pain with no hormonal medications at least 3 months before surgery, mid-follicular cycle phase **Bilibio 2014** inclusion criteria for endometriosis group: superficial peritoneal implants confirmed by biopsy, regular menstrual cycles, negative transvaginal ultrasonography for endometrioma and deep endometriosis **Chen 1998** patients undergoing laparoscopy for dysmenorrhoea **Colacurci 1996** women undergoing laparoscopy for infertility in mid-follicular cycle phase **Fedele 1989**				A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? High concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	women undergoing laparoscopy for infertility, pelvic pain or both **Ferreira 1994** women scheduled for laparoscopy or laparotomy for investigation of infertility **Franchi 1993** patients of reproductive age undergoing laparotomy or laparoscopy for pelvic mass **Gagne 2003** patients of pre-menopausal age who had never been pregnant, luteal phase of the menstrual cycle (based on the last period and further confirmed by histology), regular cycles (21-35 days), not acute salpingitis, no hormonal treatment or intrauterine device in previous 3 months. **Hallamaa 2012** patients undergoing laparoscopy for suspected endometriosis or tubal ligation **Harada 2002** atients who underwent laparotomy or laparoscopy with the preoperative diagnosis of infertility, myoma uteri, adenomyosis or endometriosis (cases)				Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	and patients who underwent laparoscopy for infertility investigation (controls) **Hornstein 1995** patients with the preoperative diagnosis of endometriosis, pelvic pain, or infertility recruited from 2 fertility units **Koninckx 1996** women scheduled for laparoscopy for suspected endometriosis **Kurdoglu 2009** women undergoing laparoscopy or laparotomy or various indications **Lanzone 1991** women undergoing laparoscopy for infertility or pelvic pain during luteal phase of the cycle **Maiorana 2007** women who underwent laparoscopy for infertility, ovarian cyst or suspected endometriosis (endometriosis group) and women operated for ovarian cysts and confirmed not to have endometriosis (controls) **Martinez 2007** productive age and regular menstrual cycles; exclusion criteria: administration of				Could the patient flow have introduced bias? Low risk **Guerriero 1996** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	any medication over the previous 2 years, acute inflammatory diseases or neoplasms, 2 or more concomitant findings at laparoscopy **Mohamed 2013** women referred for laparoscopy for unexplained primary infertility, chronic pelvic pain or both with regular menses, follicular cycle phase; only patients with advanced disease selected **Molo 1994** consecutive patients undergoing laparoscopy for infertility investigation **Muscatello 1992** women who underwent laparoscopy for infertility, pelvic pain or both at the authors' institution **Patton 1986** women who underwent laparoscopy with no systemic diseases **Somigliana 2004** women who underwent gynaecologic laparoscopy for benign gynaecological pathologies; reproductive age, gynaecological indications for laparoscopic surgery **Vigil 1999**				Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	women who underwent laparoscopy for dysmenorrhoea and pelvic pain not responding to medical management, with or without infertility **Yang 1994** women who underwent laparoscopy for infertility or suspected endometriosis **Zeng 2005** reproductive age regular menstrual cycle; exclusion criteria: hormonal treatment for 3/12 months prior reproductive age, preoperative diagnosis of uterine fibroids, adenomyosis. **<u>Exclusion Criteria</u>** **Barbati 1994** Not reported **Bilibio 2014** endocrine disorders, drugs that could affect the parameters of the tests employed, irregular menstrual cycles, infertility or pain were not caused by endometriosis, any hormonal medications in 3/12 months before surgery **Chen 1998** Not reported **Colacurci 1996**				Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Not reported **Fedele 1989** Not reported **Ferreira 1994** endocrine abnormalities, systemic disease, abnormal laboratory investigations, uterine fibroids, PID, pelvic pathology other than endometriosis identified at surgery **Franchi 1993** Not reported **Gagne 2003** Not reported **Hallamaa 2012** suspicion of malignancy, pregnancy or infection **Harada 2002** patients with malignant tumours or inflammatory disease **Hornstein 1995** Not reported **Koninckx 1996** hormonal treatment or medical treatment for endometriosis in the 3 months preceding laparoscopy, refusal a clinical examination during menstruation (only DIE considered) **Kurdoglu 2009**				**Hallamaa 2012** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' No Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? high concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	suggested or ascertained diagnosis of myoma uteri, adenomyosis, pelvic inflammatory disease or malignancy, salpingitis, other benign ovarian tumour and refusal to participate in the study **Lanzone 1991** peritoneal fluid positive for mycoplasma and chlamydia **Maiorana 2007** patients with malignant tumours or inflammatory disease **Martinez 2007** administration of any medication over the previous 2 years, acute inflammatory diseases or neoplasms, 2 or more concomitant findings at laparoscopy **Mohamed 2013** hormonal treatment for 3 months prior to surgery, history of ovarian cancer, ovarian failure, pelvic inflammatory disease or other gynaecological pathologies, previous pelvic surgery, obesity, smokers **Molo 1994** Not reported **Muscatello 1992**				the reference standard? Unclear If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Not reported **Patton 1986** Not reported **Somigliana 2004** suspected or ascertained diagnosis of malignancy, pregnancy, menopausal age, refusal to participate in the study **Vigil 1999** Not reported **Yang 1994** Not reported **Zeng 2005** hormonal treatment for 3/12 months prior reproductive age, preoperative diagnosis of uterine fibroids, adenomyosis.				Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Harada 2002** A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Hornstein 1995** A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					interpretation have introduced bias? Unclearrisk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Koninckx 1996** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? No

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Kurdoglu 2009** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					CSR 'Was a two-gate design avoided?' NO Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? high concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Lanzone 1991** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?'Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Maiorana 2007** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' No Did the study avoid inappropriate exclusions? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? high concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Martinez 2007** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' No Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Unclearrisk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? high concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Mohamed 2013** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' No Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability:

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Are there concerns that the included patients and setting do not match the review question? lhigh concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Molo 1994** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' unclear Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					do not match the review question? unclear concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Muscatello 1992** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Patton 1986** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' No Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? high concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Unclear Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s)

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Somigliana 2004** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?'Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					V**igil 1999** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclearrisk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Unclear Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the patient flow have introduced bias? Unclear risk **Yang 1994** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Zeng 2005** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: low concern Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk

1

G.8 Review question: Diagnosis – Biomarkers: HE-4

3 **What is the accuracy of of HE-4 in diagnosing endometriosis?**

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation 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 **Ref Id** 417763	**Condition** Women diagnosed with pelvic mass and scheduled for surgery **Sample size** N=68 **Characteristics** Not reported **Inclusion Criteria** • Women diagnosed with pelvic mass	**Tests** HE-4 Surgery and histology	**Methods** Serum HE4 was obtained from women prior to surgery. Serum HE-4 levels were measured using the EIA assay, and the upper limit for HE-4 was 114 pM. A cut-off point corresponding to the highest accuracy was determined by the authors. Pathology reports were also reviewed at the time for histopathological classification of benign neoplasms. Patients were stratified by benign disease classification. Percentages of elevated biomarker levels were determined. The P values for comparison of the proportion of patients with elevated HE-4 and Ca125 in various benign	**Results** Endometriosis/endometrioma; 17 women in the endometriosis or endometrioma subgroup were found not to have elevated HE-4 levels. Sensitivity (95% CI): 0% Specificity (95% CI): 98% (90 - 100)* *calculated using a binomial calculator for the confidence intervals (http://statpages.info/confint.html)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? unclear Was a case-control design avoided? No Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the included patients and setting do not match the review question? No Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Country/ies where the study was carried out China **Study type** Prospective study **Aim of the study** To measure human epididymis protein 4 (HE-4) and Ca125 levels in Chinese women with benign gynaecological disorders **Study dates** February 2010 to July 2012 **Source of funding** Not reported	undergoing surgery **Exclusion Criteria** Not reported		histopathological classifications were determined.		Were the index test results interpreted without knowledge of the results of the reference standard? No If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was there an appropriate interval between index test and reference standard? unclear Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? unclear risk

1

G.9 Review question: Diagnosis – Biomarkers in endometrial tissues (the nerve fibre marker Protein Gene Product 9.5 (PGP 9.5)

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?

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation 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	**Condition** Study participants included reproductive-aged women (puberty to menopause) with suspected endometriosis based on clinical symptoms, pelvic examination or both, who undertook the index test as well as the reference standard.	**Tests** **Al-Jefout 2007** **Index test:** endometrial nerve fibres: PGP 9.5 **Reference test:** laparoscopy + histology **Al-Jefout 2009** **Index test:** endometrial nerve fibres: PGP 9.5 **Reference test:** laparoscopy + histology **Bokor 2009**	**Methods** **Al-Jefout 2007** Description of positive case definition by index test as reported: presence of nerve fibres in the functional layer of endometrium, measured by IHC staining for PGP 9.5 (immunostaining was carried out on a Dako Autostainer Model S3400 (Dako Cytomation, Inc, CA); images analysed by using an Olympus BX51 digital camera (Olympus, Japan)); laboratory technique described; 3 pathologists, 2	**Results** **Al-Jefout 2007** Sensitivity (95% CI): 100% (83 to 100) Specificity (95% CI): 100% (80 to 100) **Al-Jefout 2009** Sensitivity (95% CI): 98% (92 to 100) Specificity (95% CI): 83% (66 to 93) **Bokor 2009** Sensitivity (95% CI): 95% (75 to 100)	**Limitations** AMSTAR Checklist 1. Was an 'a priori' design provided? Y 2. Was there duplicate study selection and data extraction? Y 3. Was a comprehensive literature search performed? Y 4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? No 5. Was a list of studies (included and excluded) provided? Y 6. Were the characteristics of the included studies provided? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Systematic Reviews, 2016 **Ref Id** 496552 **Country/ies where the study was carried out** New Zealand **Study type** Cochrane Review **Aim of the study** To investigate the influence of heterogeneity on the diagnostic accuracy of endometrial biomarkers for endometriosis **Study dates** 2016 **Source of funding** Internal sources Cochrane Menstrual Disorders and	**Sample size** N=54 studies only 8 studies relevant to the present review were included **Characteristics** **Al-Jefout 2007** **Clinical presentation:** chronic pelvic pain, infertility or both **Age:** reproductive age, not specified **Number enrolled:** 37 women **Number available for analysis**: 37 women (menstrual cycle phase not specified) **Setting:** Royal Prince Alfred Hospital, a tertiary referral centre **Place of study:** Sydney, Australia **Period of study:** 1 January 2006 to 1 December 2006	**Index test:** endometrial neural marker PGP 9.5 **Reference test:** laparoscopy + histology **Elgafor el Sharkwy 2013** **Index test:** endometrial nerve fibres - PGP 9.5 **Reference test:** laparoscopy **Leslie 2013** **Index test:** endometrial functional layer nerve fibres - PGP 9.5 **Reference test:** laparoscopy + histology **Makari 2012** **Index test:** endometrial nerve fibres - PGP 9.5 **Reference test:** laparoscopy + histology **Meibody 2011** **Index test:** endometrial small nerve fibres in eutopic	of whom had good experience in nerve fibre counting; 'blinded counting' **Al-Jefout 2009** Description of positive case definition by index test as reported: presence of endometrial nerve fibres in functional layer by IHC staining for PGP 9.5 (Immunostaining on a Dako Autostainer Model S3400 (Dako, Australia); image analysis by using an Olympus microscope BX51 and digital camera DP70 (Olympus, Japan)); laboratory technique described; 2 people with experience in nerve fibre counting, blinded to the patients' data and each others' results **Bokor 2009** Description of positive case definition by index test as reported: nerve fibre density was defined as total number of nerve fibres divided by the total surface area of the examined endometrium; nerve fibres were evaluated by IHC for each marker and counted in HPF areas for the slide section (antibody detection with REAL Detection System, Alkaline	Specificity (95% CI): 75% (51 to 91) **Elgafor el Sharkwy 2013** Sensitivity (95% CI): 92% (79 to 98) Specificity (95% CI): 80% (64 to 91) **Leslie 2013** Sensitivity (95% CI): 19% (9 to 33) Specificity (95% CI): 71% (48 to 89) **Makari 2012** Sensitivity (95% CI): 100% (69 to 100) Specificity (95% CI): 50% (19 to 81) **Meibody 2011** Sensitivity (95% CI): 100% (74 to 100) Specificity (95% CI): 80% (52 to 96) **Yaday 2013** Sensitivity (95% CI): 80% (61 to 92) Specificity (95% CI): 100% (88 to 100)	7. Was the scientific quality of the included studies assessed and documented? Y 8. Was the scientific quality of the included studies used appropriately in formulating conclusions? Y 9. Were the methods used to combine the findings of studies appropriate? Y 10. Was the likelihood of publication bias assessed? No 11. Was the conflict of interest included? Y Where there is a high/unclear risk regarding applicability it is due to a two-gate design: according to Gupta et al. 2016 these are studies with two sets of inclusion criteria with respect to **Clinical presentation:** and one set of inclusion criteria with respect to **Reference test:** the participants with or without a clinical suspicion of endometriosis scheduled for abdominal surgery QUADAS 2 **Al-Jefout 2007** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Subfertility Group, University of Auckland, New Zealand. Technical support The Robinson Institute, University of Adelaide, Australia. Access to academic resources External sources No sources of support supplied	**Al-Jefout 2009** **Clinical presentation:** pelvic pain symptoms alone (n = 52), infertility alone (n = 24), pelvic pain + infertility (n = 20), no pain and no infertility (n = 3) **Age:** mean age 33.9 years (range 20-50 years) **Number enrolled:** 103 women **Number available for analysis**: 99 women (menstrual cycle phase n = 15; proliferative n = 39; mid-cycle n = 14; secretory n = 31) **Setting:** Royal Prince Alfred Hospital, a tertiary referral centre **Place of study:** Sydney, Australia **Period of study:** 12 December	endometrium - PGP 9.5 **Reference test:** Laparoscopy/laparotomy + histology **Yaday 2013** **Index test:** endometrial nerve fibres **Reference test:** laparoscopy + histology	Phosphatase/RED, Rabbit/Mouse (Dako); analysis by image analysis software KS400 3.0 (Zeiss, Germany) linked to a Zeiss microscope); the whole surface of each section was evaluated on high-power images; procedure described; thresholds not pre-specified; reported cut-off values: PGP 9.5 − 0.49, VIP − 0.08, CGRP − 0.23, SP − 0.2, NPY − 0.13, NF − 0.19; 1 examiner who was blinded to the diagnosis **Elgafor el Sharkwy 2013** Description of positive case definition by index test as reported: presence of nerve fibres in the functional layer of endometrium, assessed by IHC staining for PGP 9.5 (an average of 4–5 sections per specimen were examined by using an Olympus microscope); 2 pathologists, both of whom have good experience in nerve fibre identification **Leslie 2013** Description of positive case definition by index test as reported: presence of functional layer nerve fibres as detected by PGP 9.5 IHC staining (lower uterine,		Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	2007 to 10 December 2008 **Bokor 2009** **Clinical presentation:** infertility, 100%; dysmenorrhoea, 25% **Age:** mean age 33 ± 10 years, endometriosis; 32 ± 5 years, controls **Number enrolled:** 40 women (retrospective selection) **Number available for analysis**: 40 women (all in secretory phase of menstrual cycle) **Setting:** University Hospital Gasthuisberg **Place of study:** Leuven, Belgium **Period of study:** not provided		cervical and basal layer staining was not considered; magnification using a Leica DM2500 light microscope); laboratory technique described; single pathologist unaware of the results for the reference standard; positive and equivocal biopsies were blindly reviewed by the 2nd pathologist, disagreement resolved by consensus **Makari 2012** Description of positive case definition by index test as reported: presence of nerve fibres as detected by IHC staining for PGP 9.5 (evaluatIoin under × 400 magnification, microscope Olympus BX51; the number of immunoreactive nerve fibres was also calculated for each cross-sectional area to assess nerve fibre density) **Meibody 2011** Description of positive case definition by index test as reported: Presence of nerve fibres detected by IHC staining for PGP 9.5 seen in 10 HPF (IHC by using Dako Denmark A/S Produktionsej42 DK-2600, Denmark and Olympus microscope; assessment of 3-4 sections per slide;		Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Al-Jefout 2009** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided?According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Elgafor el Sharkwy 2013 **Clinical presentation:** (n/N): infertility - 91/114; dysmenorrhoea - 64/114; dyspareunia - 17/114; dyschezia - 6/114; other pelvic pain - 35/114 **Age:** mean age 29 ± 0.6 years, controls; 31 ± 1.1 years, endometriosis **Number enrolled:** 114 women **Number available for analysis**: 78 women (all in follicular cycle phase; only control and endometriosis stage I-II were analysed) **Setting:** University hospital - Zagazig University Hospital		density of NF was also calculated by intensity of staining); laboratory technique described; pathologist was blinded to reference standard result **Yaday 2013** Description of positive case definition by index test as reported: positive IHC staining for PGP 9.5 identified as single cell positive or linear nerve fibres; technique described; senior pathologist blinded to patients' data		Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Place of study: Zagazig, Egypt **Period of study:** December 2010 to April 2012 **Leslie 2013** Clincial presentation: pain - 45/68, infertility - 14/68; adnexal mass/ menorrhagia - 7/68; hormonal therapy - 11/68; information was not available in 1 control and 11 cases **Age:** mean age 35 years (range 21–53) **Number enrolled:** 68 women **Number available for analysis**: 68 women (25 in proliferative, 19 in secretory cycle phase; 24 - unclear/hormonal treatment) **Setting:** university hospital				Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Bokor 2009** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? No Could the conduct or interpretation of the index test have introduced bias? High risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	- King Edward Memorial Hospital and private hospital - Hollywood Hospital **Place of study:** Perth, Australia **Period of study:** 2006-2011 **Makari 2012** **Clinical presentation:** dysmenorrhoea - 10/20, chronic pelvic pain - 11/20, infertility, dyspareunia, dysuria, dyschezia **Age:** mean age 36.1 ± 6.10, endometriosis; 30 13 ± 6.38 years, controls **Number enrolled:** 20 women **Number available for analysis**: 20 women (15 in proliferative and 5 in secretory cycle phase)				B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Setting: university hospital - Hospital of Lithuanian University of Health Sciences Kaunas Clinics **Place of study:** Kaunas, Lithuiania **Period of study:** 2009-2011 **Meibody 2011** **Clinical presentation:** chronic pelvic pain - 23/27, dyspareunia - 5/27, dysmenorrhoea - 7/27, infertility - 5/27 **Age:** mean age 39.5 ± 5.9 years, endometriosis; 41.6 ± 5.7 years, controls **Number enrolled:** 27 women **Number available for analysis**: 27 women (all in				**Elgafor el Sharkwy 2013** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the included patients and setting do not match the review question? high concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s)

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	proliferative cycle phase) **Setting:** university hospital - Minimally Invasive Surgery Research Center, Rassoul Akram Hospital, Iran University of Medical Sciences **Place of study:** Tehran, Iran **Period of study:** 2007-2009 **Yaday 2013** **Clinical presentation:** infertility - 32/60, CPP - 19/60, infertility + pain symptoms (dysmenorrhoea, dyspareunia, dyschezia) - 9/60; regular menstrual cycle - 57/60 **Age:** range 15-45 years **Number enrolled:** 60 women **Number available for analysis**: 60				Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Leslie 2013** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	women (cycle phase not specified) **Setting:** university hospital - O&G Department, University College of Medical Sciences and Guru Teg Bahadur Hospital **Place of study:** Delhi, India **Period of study:** November 2009 to April 2012 **<u>Inclusion Criteria</u>** **Al-Jefout 2007** reproductive-aged women undergoing laparoscopy for suspected endometriosis or infertility **Al-Jefout 2009** reproductive-aged women undergoing laparoscopy for infertility, pelvic pain or both **Bokor 2009**				Could the selection of patients have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	reproductive-aged women undergoing laparoscopy for infertility, pelvic pain or both with no medical treatment for 3/12 months preceding surgery, secretory phase of menstrual cycle **Elgafor el Sharkwy 2013** women undergoing laparoscopy for infertility, pelvic pain or both, reproductive age, follicular phase of the cycle and regular menstrual cycle; **Leslie 2013** patients undergoing laparoscopy for suspected endometriosis **Makari 2012** patients that presented for laparoscopy for infertility, pelvic pain or both; reproductive				Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Makari 2012** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	age (18-45 years); exclusion criteria: hormonal treatment 3/12 months before surgery, pregnancy or oncology cases **Meibody 2011** women undergoing laparoscopy/laparotomy for infertility or pelvic pain; reproductive age, regular menstrual cycle **Yaday 2013** patients who underwent laparoscopy for infertility/pelvic pain/suspected endometriosis **Exclusion Criteria** **Al-Jefout 2007** current hormonal treatment for endometriosis, pregnancy and unwillingness to participate **Al-Jefout 2009**				Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	hormonal treatment for 3/12 months prior to surgery, pregnancy, unwillingness to participate **Bokor 2009** not reported **Elgafor el Sharkwy 2013** any current infection, any medication within 1 month prior to laparoscopy, previous surgery for endometriosis and smoking or drinking alcohol **Leslie 2013** histological diagnosis not available (ablated lesions). Hormonal pretreatment was not an exclusion **Makari 2012** not reported **Meibody 2011** unwillingness to participate and use of hormonal medications for				Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Meibody 2011** A. Risk of Bias Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? unclear concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	the past 3/12 months **Yaday 2013** hormonal therapy in the preceding 3/12 months, acute PID, suspected pregnancy, suspected or diagnosed genital malignancy, undiagnosed vaginal bleeding, documented genital tuberculosis, contraindication for laparoscopy or unwillingness to undergo surgery				B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Yaday 2013 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the included patients and setting do not match the review question? unclear concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s)

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk

1

G.10 Review question: Diagnosis – MRI

3 **What is the accuracy of MRI in diagnosing endometriosis?**

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation	**Condition**	**Tests**	**Methods**	**Results**	**Limitations**

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Nisenblat, Vicki, Farquhar, Cindy, Akoum, Ali, Fraser, Ian, Bossuyt, M. M. Patrick, Hull, Louise M., Non-invasive tests for the diagnosis of endometriosis, Cochrane Database of Systematic Reviews, 2012 **Ref Id** 359883 **Country/ies where the study was carried out** New Zealand **Study type** Cochrane Review **Aim of the study** To provide estimates of the diagnostic accuracy of imaging modalities for the diagnosis of	Study participants included women of reproductive age (puberty to menopause) with suspected endometriosis based on clinical symptoms and/or pelvic examination, who undertook both the index test and the reference standard. **Sample size** N=49 studies involving 4807 women (for both ultrasound and MRI) **Characteristics** **Abrao 2007** **Clinical presentation:** dysmenorrhoea 53/104, deep dyspareunia 66/104, acyclical pelvic pain 17/104, infertility 55/104, cyclical bowel symptoms (pain/bleeding)	**Abrao 2007** **Index test:** MRI (T1/T2-w) **Reference test:** laparoscopy 104/104 (100%) + histopathology **Ascher 1995** **Index test:** MRI 3 types (T1/T2-w (CSE); T1/T2-w + fat-suppressed (CSE/TIFS); T1/T2-w + fat-suppressed + Gd (CSE/TIFS/Gd-TIFS)) **Reference test:** laparoscopy 24/31 (77.4%), laparotomy 7/31 (22.6%) **Bazot 2009** **Index test:** MRI (T1/T2-w + fat-suppressed/Gd) **Reference test:** laparoscopy 79/92 (85.9%), laparotomy 13/92 (14.1%) + histopathology **Bazot 2013** **Index test:** MRI 2 types: 2-dimensional fast	**Abrao 2007** MRI: carried out independently by a single examiner who was blinded to participants' clinical data and to results of other imaging; level of expertise not reported **Ascher 1995** MRI: prospectively evaluated by 2 radiologists experienced in pelvic MRI; readers aware of clinical suspicion of endometriosis **Bazot 2009** MRI: each examination interpreted according to a standardised protocol, retrospectively by 1 radiologist with 2 years' experience in gynaecological imaging. Readers informed of women's clinical history and symptoms but blinded to results of physical and previous imaging examinations **Bazot 2013** MRI: images independently analysed by 2 radiologists with different degrees of experience in female MRI (1 reader with > 20 years' experience; second reader a junior radiologist). Both	**Abrao 2007** RVS (rectovaginal septum) endometriosis: Sensitivity (95% CI): 0.76 (0.60 to 0.88) Specificity (95% CI): 0.68 (0.55 to 0.79) Anterior DIE: Sensitivity (95% CI): 0.83 (0.71 to 0.92) Specificity (95% CI): 00.98 (0.89 to 1.00) Rectovaginal: Sensitivity (95% CI): 76% (60 to 88) Specificity (95% CI): 68% (55 to 79) Rectosigmoid: Sensitivity (95% CI): 83% (71 to 92) Specificity (95% CI): 98% (89 to 100) **Ascher 1995** Pelvic endometriosis (T1-/T2-w): Sensitivity (95% CI): 76% (53 to 92) Specificity (95% CI): 60% (26 to 88) Pelvic endometriosis (T1-/T2-w + fat-supressed):	AMSTAR Checklist 1. Was an 'a priori' design provided? Y 2. Was t. here duplicate study selection and data extraction? Y 3. Was a comprehensive literature search performed? Y 4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? No 5. Was a list of studies (included and excluded) provided? Y 6. Were the characteristics of the included studies provided? Y 7. Was the scientific quality of the included studies assessed and documented? Y 8. Was the scientific quality of the included studies used appropriately in formulating conclusions? Y 9. Were the methods used to combine the findings of studies appropriate? Y 10. Was the likelihood of publication bias assessed? No 11. Was the conflict of interest included? Y QUADAS 2 **Abrao 2007** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
pelvic endometriosis, ovarian endometriosis and deeply infiltrating endometriosis (DIE) versus surgical diagnosis as a reference standard. To describe performance of imaging tests for mapping of deep endometriotic lesions in the pelvis at specific anatomical sites. **Study dates** 2016 **Source of funding** Internal sources Cochrane Menstrual Disorders and Subfertility Group, University of Auckland, New Zealand.	59/104, cyclical urinary symptoms 14/104 **Age:** mean 33.8 ± 6.1 years, range 18 to 45 years **Number enrolled:** 104 women **Number available for analysis:** 104 women **Setting:** tertiary university hospital, referral centre for endometriosis, São Paulo University **Place of study:** São Paolo, Brazil **Period of study:** August 2004 to October 2006 **Ascher 1995** **Clinical presentation:** not specified **Age:** mean 34.1 years, range 21 to 46 years **Number enrolled:** 38 women	spin echo T2-w (2D FSE T2-w MRI); 3-dimensional fast spin echo T2-w MRI (3D FSE T2-w MRI) **Reference test:** laparoscopy (n = 20), laparotomy (n = 3) + histopathology. **Biscaldi 2014** **Index test:** MDCT-e; MRI jelly method (MRI-e) **Reference test:** laparoscopy 260/260 (100%) + histopathology **Chamie 2009** **Index test:** MRI (T1/T2-w + fat-suppressed/Gd) **Reference test:** laparoscopy 92/92 (100%) + histopathology **Grasso 2010** **Index test:** MRI (T1/T2-w + fat-suppressed + Gd)	readers blinded to clinical and ultrasonographic findings **Biscaldi 2014** MRI: 2 radiologists blindly reviewed images at a PACS workstation; they were not aware of clinical findings and patient history, knowing only that the presence of bowel endometriosis was clinically suspected; level of expertise not reported **Chamie 2009** MRI: analysed prospectively by 2 radiologists (same examiners) who were blinded to each patient's history, physical findings and ultrasound results; level of expertise not reported **Grasso 2010** MRI: analysed prospectively by 1 radiologist who was blinded to clinical and sonographic findings; level of expertise not reported. **Ha 1994** MRI: reviewed independently by 2 radiologists; level of expertise not reported. Observer knew only that patients had suspected endometriosis **Hottat 2009** MRI: 2 investigators with 8 years' and 1 year experience	Sensitivity (95% CI): 86% (64 to 97) Specificity (95% CI): 50% (19 to 81) <u>Pelvic endometriosis (T1-/T2-w + fat-supressed/Gd):</u> Sensitivity (95% CI): 81% (58 to 95) Specificity (95% CI): 50% (19 to 81) **Bazot 2009** <u>DIE:</u> Sensitivity (95% CI): 97% (91 to 99) Specificity (95% CI): 0% (0 to 84) <u>Rectovaginal:</u> Sensitivity (95% CI): 55% (23 to 83) Specificity (95% CI): 99% (93 to 100) <u>Rectosigmoid:</u> Sensitivity (95% CI): 87% (77 to 94) Specificity (95% CI): 97% (91 to 100) <u>USL:</u> Sensitivity (95% CI): 84% (75 to 91) Specificity (95% CI): 90% (55 to 100)	Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Technical support The Robinson Institute, University of Adelaide, Other. Access to academic resources External sources No sources of support supplied	**Number available for analysis:** 31 women **Setting:** not specified **Place of study:** USA **Period of study:** 11-month period, dates not specified **Bazot 2009** **Clinical presentation:** dysmenorrhoea 79/92, dyspareunia 63/92, dyschezia 32/92, dysuria 3/92, infertility 21/92; history of surgery for endometriosis 31/92 **Age:** median age 31.8 years, range 20 to 50 years **Number enrolled:** 92 women **Number available for analysis:** 92 women	**Reference test:** laparoscopy 33/33 (100%) + histopathology **Ha 1994** **Index test:** MRI 2 types (T1/T2-w MRI; fat-suppressed T1-w MRI) **Reference test:** laparoscopy 31/31 (100%) **Hottat 2009** **Index test:** MRI (3.0T Magnetom system (3.0T MRI)) **Reference test:** laparoscopy 34/41; laparotomy 7/41 + histopathology (100%) **Manganaro 2012a** **Index test:** MRI (3.0T Magnetom system (3.0T MRI)) **Reference test:** laparoscopy 46/46 (100%) **Managaro 2012b** **Index test:** MRI (3.0T Magnetom	in MRI; blinded to clinical findings; independently and prospectively analysed all images Manganaro 2012a MRI: 2 radiologists with, respectively, 10 years' and 5 years' experience in female pelvis imaging; blinding to clinical data not reported **Managaro 2012b** MRI: 2 radiologists with 12 years' and 7 years' experience in female pelvis imaging; blinded to clinical data **Manganaro 2013** MRI: radiologist who analysed images had > 13 years' experience in imaging of the female pelvis (single operator) and was blinded to results of previous imaging or clinical examination **Okada 1995** MRI: numbers or level of expertise of surgeons or pathologists not reported; unclear whether blinded to results of the index test **Stratton 2003** MRI: 2 experienced, board-certified radiologists analysed preoperative magnetic resonance images	Vaginal wall involvement: Sensitivity (95% CI): 80% (61 to 92) Specificity (95% CI): 85% (74 to 93) Ovarian: Sensitivity (95% CI): 92% (78 to 98) Specificity (95% CI): 88% (76 to 95) **Bazot 2013** Posterior DIE (2D FSE T2-w): Sensitivity (95% CI): 89% (65 to 99) Specificity (95% CI): 20% (1 to 72) Posterior DIE (3D): Sensitivity (95% CI): 100% (81 to 100) Specificity (95% CI): 20% (1 to 72) Rectosigmoid (2D FSE T2-w): Sensitivity (95% CI): 85% (55 to 98) Specificity (95% CI): 100% (69 to 100) Rectosigmoid (3D): Sensitivity (95% CI): 85% (55 to 98)	Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Setting: tertiary care Tenon Hospital, referral centre for endometriosis and Surgical Centre Trocadero **Place of study:** Paris, France **Period of study:** April 2000 to May 2005 **Bazot 2013** **Clinical presentation:** dysmenorrhoea, deep dyspareunia, dyschezia, dysuria or infertility **Age:** median age 34 years, range 24 to 46 years **Number enrolled:** 110 women **Number available for analysis:** 23 women **Setting:** tertiary care hospital, Tenon Hospital,	system (3.0T MRI)) **Reference test:** laparoscopy 19/19 (100%) **Manganaro 2013** **Index test:** MRI (3.0T MRI) Reference standard: laparoscopy 42/42 (100%) + histopathology **Okada 1995** **Index test:** MRI (T1-w fat-saturated MRI) Reference standard: laparoscopy 47/74 (63.5%), laparotomy 27/74 (36.5%) + histopathology **Stratton 2003** **Index test:** MRI (T1/T2-w + fat-suppressed + Gd) **Reference test:** laparoscopy 48/48 (100%) + histopathology **Sugimura 1993** **Index test:** MRI (T1/T2-w)	and recorded a consensus reading of the extent and location of possible endometriosis. Radiologists were aware of the clinical possibility of deep endometriosis in all participants but did not know the results of surgery, pelvic ultrasound, history, physical exam findings or histopathology **Sugimura 1993** MRI: prospectively read by 2 study authors who were aware that patients had a clinical history of suspected endometriosis; level of expertise not reported **Takeuchi 2005** MRI: read preoperatively by 1 radiologist who was blinded to clinical findings; level of expertise not reported **Thomeer 2014** MRI: 2 experienced radiologists (blinded), with 13 years' and 12 years' experience in abdominal MRI, analysed independently and blindly data on a PACS workstation. They had no information regarding clinical data; disagreements about image interpretation were sorted by consensus	Specificity (95% CI): 90% (55 to 100) USL (2D FSE T2-w): Sensitivity (95% CI): 88% (64 to 99) Specificity (95% CI): 33% (4 to 78) USL (3D): Sensitivity (95% CI): 88% (64 to 99) Specificity (95% CI): 33% (4 to 78) Vaginal wall involvement (2D FSE T2-w): Sensitivity (95% CI): 60% (15 to 95) Specificity (95% CI): 94% (73 to 100) Vaginal wall involvement (3D): Sensitivity (95% CI): 80% (28 to 99) Specificity (95% CI): 100% (81 to 100) PoD (2D FSE T2-w): Sensitivity (95% CI): 71% (42 to 92) Specificity (95% CI): 100% (66 to 100) PoD (3D): Sensitivity (95% CI): 71% (42 to 92)	Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Ascher 1995** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability:

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	referral centre for endometriosis **Place of study:** Paris, France **Period of study:** February 2010 to May 2010 **Biscaldi 2014** **Clinical presentation:** dysmenorrhoea 185/260, dyspareunia 157/260, chronic pelvic pain 142/260, infertility 54/260, diarrhoea 57/260, constipation 85/260, bloating 122/260, dyschezia 130/260; previous surgery for endometriosis 113/260, previous medical treatment: oral contraceptive pill 79/260, contraceptive vaginal ring 14/260 **Age:** mean 32.6 ± 4.3 years	**Reference test:** laparoscopy 13/35 (37%), laparotomy 22/35 (63%) + histopathology **Takeuchi 2005** **Index test:** MRI (T1/T2-w + fat-suppressed, jelly method) **Reference test:** laparoscopy 31/31 (100%) + histopathology Thomeer 2014 **Index test:** MRI 3.0T Reference standard: laparoscopy 40/40 (100%)		Specificity (95% CI): 100% (66 to 100) **Biscaldi 2014** <u>Rectosigmoid:</u> Sensitivity (95% CI): 99% (96 to 100) Specificity (95% CI): 96% (90 to 99) **Chamie 2009** <u>Rectovaginal:</u> Sensitivity (95% CI): 89% (79 to 96) Specificity (95% CI): 92% (75 to 99) <u>Rectosigmoid:</u> Sensitivity (95% CI): 86% (73 to 94) Specificity (95% CI): 93% (81 to 99) <u>Vaginal wall involvement:</u> Sensitivity (95% CI): 73% (39 to 94) Specificity (95% CI): 100% (96 to 100) <u>Ureteral:</u> Sensitivity (95% CI): 50% (16 to 84) Specificity (95% CI): 100% (96 to 100) <u>Bladder:</u>	Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Unlcear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number enrolled: 260 women **Number available for analysis:** 260 women **Setting:** tertiary care university hospital, San Martino Hospital, referral centre for endometriosis, Galliera Hospital **Place of study:** Genoa, Italy **Period of study:** not specified **Chamie 2009** **Clinical presentation:** dysmenorrhoea 89/92, dyspareunia 54/92, acyclical pain 72/92, dysuria 8/92, dyschezia 44/92, infertility 40/92; painful palpable nodules on examination 58/92			Sensitivity (95% CI): 23% (5 to 54) Specificity (95% CI): 100% (95 to 100) **Grasso 2010** Pelvic endometriosis: Sensitivity (95% CI): 57% (39 to 73) Specificity (95% CI): 98% (90 to 100) DIE: Sensitivity (95% CI): 96% (80 to 100) Specificity (95% CI): 86% (42 to 100) **Ha 1994** Pelvic endometriosis (T1-/T2-w): Sensitivity (95% CI): 52% (33 to 71) Specificity (95% CI): 100% (16 to 100) Pelvic endometriosis (fat-supressed): Sensitivity (95% CI): 76% (56 to 90) Specificity (95% CI): 100% (16 to 100) **Hottat 2009** DIE:	Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Bazot 2009** A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Age: mean 33 years, range 20 to 52 years **Number enrolled:** 92 women **Number available for analysis:** 92 women **Setting:** tertiary university hospital, referral centre for endometriosis, São Paulo University **Place of study:** São Paolo, Brazil **Period of study:** November 2005 to July 2007 **Grasso 2010** **Clinical presentation:** pain (dysmenorrhoea, dyspareunia, chronic pelvic pain) 18/33, infertility 5/33, adnexal masses and/or tenderness at physical			Sensitivity (95% CI): 96% (81 to 100) Specificity (95% CI): 100% (77 to 100) <u>Anterior DIE:</u> Sensitivity (95% CI): 75% (35 to 97) Specificity (95% CI): 100% (89 to 100) <u>Rectosigmoid:</u> Sensitivity (95% CI): 100% (75 to 100) Specificity (95% CI): 96% (82 to 100) <u>USL:</u> Sensitivity (95% CI): 82% (60 to 95) Specificity (95% CI): 89% (67 to 99) <u>Vaginal wall involvement:</u> Sensitivity (95% CI): 82% (48 to 98) Specificity (95% CI): 97% (83 to 100) <u>PoD:</u> Sensitivity (95% CI): 95% (76 to 100) Specificity (95% CI): 100% (83 to 100) <u>Ovarian:</u> Sensitivity (95% CI): 95% (76 to 100)	Patient Sampling Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	examination 10/33 **Age:** mean 35, range 22 to 53 years **Number enrolled:** 33 women **Number available for analysis:** MRI 33 women; 3D-TVUS 24 women **Setting:** University Hospital, Villa Valeria Hospital and Campus Bio Medico University of Rome **Place of study:** Rome, Italy **Period of study:** June 2006 to June 2008 **Ha 1994** **Clinical presentation:** not specified **Age:** mean 35 years, range 20 to 52 years **Number enrolled:** 31 women			Specificity (95% CI): 95% (75 to 100) **Manganaro 2012a** <u>Pelvic endometriosis:</u> Sensitivity (95% CI): 97% (84 to 100) Specificity (95% CI): 100% (77 to 100) <u>DIE:</u> Sensitivity (95% CI): 96% (78 to 100) Specificity (95% CI): 100% (85 to 100) <u>USL:</u> Sensitivity (95% CI): 95% (74 to 100) Specificity (95% CI): 91% (72 to 99) <u>Ovarian:</u> Sensitivity (95% CI): 100% (82 to 100) Specificity (95% CI): 96% (81 to 100) **Managaro 2012b** <u>PoD:</u> Sensitivity (95% CI): 93% (68 to 100) Specificity (95% CI): 75% (19 to 99) **Manganaro 2013** <u>USL</u>:	Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number available for analysis: 31 women **Setting:** University Hospital, Catholic University Medical College **Place of study:** Seoul, Korea **Period of study:** 12-month period, dates not specified **Hottat 2009** **Clinical presentation:** dysmenorrhoea 19/41, chronic pelvic pain 29/41, dyspareunia 5/41, suspicious clinical examination 15/41, past hx of endometriosis 7/41 **Age:** mean 33 years, range 20 to 46 years **Number enrolled:** 106 women **Number available for**			Sensitivity (95% CI): 95% (74 to 100) Specificity (95% CI): 91% (72 to 99) **Okada 1995** Pelvic endometriosis: Sensitivity (95% CI): 88% (77 to 95) Specificity (95% CI): 67% (30 to 93) **Stratton 2003** Pelvic endometriosis: Sensitivity (95% CI): 67% (50 to 80) Specificity (95% CI): 75% (19 to 99) **Sugimura 1993** Pelvic endometriosis: Sensitivity (95% CI): 73% (52 to 88) Specificity (95% CI): 67% (30 to 93) **Takeuchi 2005** Posterior DIE: Sensitivity (95% CI): 94% (71 to 100) Specificity (95% CI): 100% (77 to 100) PoD:	Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Bazot 2013** A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear risk Could the selection of patients have introduced bias? unclear risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	analysis: 41 women **Setting:** endometriosis referral centre, Erasme Hospital, Universite´ Libre de Bruxelles **Place of study:** Brussels, Belgium **Period of study:** March 2007 to August 2008 **Manganaro 2012a** **Clinical presentation:** chronic pelvic pain, infertility; transvaginal ultrasound suggestive of endometriosis 23/46; treatment with combined oral contraceptive pill 17/46 **Age:** mean 30.4 years, range 20 to 43 years **Number enrolled:** 46 women **Number available for**			Sensitivity (95% CI): 91% (71 to 99) Specificity (95% CI): 78% (40 to 97) **Thomeer 2014** <u>Pelvic endometriosis:</u> Sensitivity (95% CI): 81% (65 to 92) Specificity (95% CI): 100% (29 to 100) <u>PoD:</u> Sensitivity (95% CI): 100% (69 to 100) Specificity (95% CI): 100% (88 to 100)	B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s)

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	analysis: 46 women **Setting:** University Hospital: Umberto I Hospital, Sapienza University of Rome **Place of study:** Rome, Italy **Period of study:** February 2010 to September 2010 **Managaro 2012b** **Clinical presentation:** transvaginal ultrasound examination positive for endometriosis, chronic pelvic pain, symptomatic patients with negative ultrasound examination **Age:** mean 26 years, range 19 to 35 years **Number enrolled:** 19 women				Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number available for analysis: 19 women **Setting:** University Hospital: Umberto I Hospital, Sapienza University of Rome **Place of study:** Rome, Italy **Period of study:** October 2010 to April 2011 **Manganaro 2013** **Clinical presentation:** severe pain symptoms such as dyspareunia, dysmenorrhoea and acyclical pain (visual analogue scale (VAS) > 7/10) **Age:** mean 28 years, range 19 to 45 years **Number enrolled:** 42 women **Number available for**				Could the patient flow have introduced bias? High risk **Biscaldi 2014** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	analysis: 42 women **Setting:** University Hospital, Umberto I Hospital, "Sapienza" University of Rome **Place of study:** Rome, Italy **Period of study:** July 2010 to July 2012 **Okada 1995** **Clinical presentation:** infertility, lower abdominal pain, menstrual pain, dyspareunia; suspected endometriosis on pelvic examination or transvaginal ultrasonography **Age:** mean 37.4 years, range 26 to 49 years **Number enrolled:** 74 women **Number available for**				Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? High risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	analysis: 74 women **Setting:** University Hospital, Shimane Medical University **Place of study:** Izumo, Japan **Period of study:** August 1991 to December 1993 **Stratton 2003** **Clinical presentation:** pelvic pain (menstrual, coital and non-menstrual pelvic pain) confirmed by standardised questionnaire using a visual analogue scale; none treated for endometriosis in the past 6 months nor had taken hormonal medication in the past 3 months; prior surgical diagnosis of endometriosis 38/58				B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Chamie 2009** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Age: range 20 to 44 years **Number enrolled:** 58 women **Number available for analysis:** 46 women **Setting:** university hospitals, Warren G. Magnusen Clinical Center, National Institutes of Health, Georgetown University Medical Center **Place of study:** Bethesda, MD, Washington, DC, USA **Period of study:** January 1999 to November 2000 **Sugimura 1993** **Clinical presentation:** not specified **Age:** mean 36 years, range 24 to 48 years				Did the study avoid inappropriate exclusions? No Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number enrolled: 35 women **Number available for analysis:** 35 women **Setting:** university hospital, Shimane Medical University **Place of study:** Izumo, Japan **Period of study:** March 1991 to August 1992 **Takeuchi 2005** **Clinical presentation:** dysmenorrhoea 31/31, dyspareunia 10/31, chronic pelvic pain 7/31; sonography suggestive for endometrioma 25/31; none had a history of previous pelvic surgery, and none had received hormonal therapy within 6 months				Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Unclear Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Unclear risk **Grasso 2010** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	preceding the study **Age:** mean 32.1 ± 4.2 years **Number enrolled:** 31 women **Number available for analysis:** 31 women **Setting:** university hospital, Juntendo University School of Medicine **Place of study:** Tokyo, Japan **Period of study:** January 2001 to July 2002 **Thomeer 2014** **Clinical presentation:** pain, subfertility and other symptoms suggestive of endometriosis (not specified) **Age:** median 25 years, range 18 to 39 years				Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability: Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Number enrolled: 40 women **Number available for analysis:** 40 women **Setting:** university hospital, Erasmus Medical Centre, Rotterdam University **Place of study:** Rotterdam, The Netherlands **Period of study:** November 2010 to December 2012 **<u>Inclusion Criteria</u>** **Abrao 2007** Study population: patients with clinically suspected endometriosis Selection criteria: not specified **Ascher 1995** Study population: women with clinically				A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	suspected endometriosis who were scheduled for surgery Selection criteria: not specified **Bazot 2009** Study population: women referred with clinical evidence of pelvic endometriosis Selection criteria: not specified **Bazot 2013** Study population: patients referred for pelvic MRI because of clinical suspicion of endometriosis Selection criteria: not specified **Biscaldi 2014** Study population: patients referred to (our) endometriosis centre Inclusion criteria: reproductive age,				Could the patient flow have introduced bias? Low risk **Ha 1994** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability: Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	suspicion of deep pelvic endometriosis on the basis of symptoms and vaginal examination, gastrointestinal symptoms that might be caused by rectosigmoid endometriosis. **Chamie 2009** Study population: women who had a history and findings of a physical exam consistent with endometriosis Inclusion criteria: symptoms consistent with endometriosis, such as pelvic pain, dysmenorrhoea, deep dyspareunia, acyclical pelvic pain, dyschezia and infertility; pelvic examination revealing thickening of posterior cul-de-				Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	sac and/or nodules; transvaginal ultrasound results showing ovarian cysts with thickened low-amplitude echoes; no previous pelvic surgery for endometriosis **Grasso 2010** Study population: patients with clinical suspicion of pelvic endometriosis Selection criteria: not specified **Ha 1994** Study population: patients with suspected endometriosis Selection criteria: not specified **Hottat 2009** Study population: patients referred for pelvic MR imaging because of clinical				B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Hottat 2009** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	suspicion of endometriosis Inclusion criteria: not reported **Manganaro 2012a** Study population: women with clinical ± sonographic suspicion of endometriosis Inclusion criteria: transvaginal ultrasound examination positive for endometriosis; patients with chronic pelvic pain; symptomatic patients with negative ultrasound; infertile patients **Managaro 2012b** Study population: women with clinical ± sonographic suspicion of endometriosis Inclusion criteria: transvaginal				Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Low risk B. Concerns regarding applicability Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	ultrasound examination positive for endometriosis; patients with chronic pelvic pain; symptomatic patients with negative ultrasound; infertile patients **Manganaro 2013** Study population: patients with suspected USL DIE based on clinical symptoms, abnormal gynaecological examination or transvaginal ultrasound findings Selection criteria: not specified **Okada 1995** Study population: women visiting outpatient department with suspected endometriosis based on Clinical presentation:				Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	(symptoms and pelvic examination), transvaginal ultrasonography and/or blood test for Ca-125 Selection criteria: not specified **Stratton 2003** Study population: women 18 to 45 years of age with pelvic pain, who were otherwise in good health, were evaluated to exclude other causes of pain (from a cohort of women recruited for a randomised, double-blind, placebo-controlled study of surgical excision followed by innovative medical treatment for endometriosis) Selection criteria: not specified **Sugimura 1993**				Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Manganaro 2012a** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Patient characteristics and setting

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Study population: women with clinically suspected endometriosis Selection criteria: not specified **Takeuchi 2005** Study population: women scheduled to undergo laparoscopy for suspected rectovaginal endometriosis based on clinical symptoms, rectal/pelvic examination findings and preoperative sonographic examination results Selection criteria: not specified **Thomeer 2014** Study population: patients with clinical suspicion of endometriosis scheduled to undergo laparoscopy				Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Selection criteria: not specified **Exclusion Criteria** **Abrao 2007** exclusion criteria: virgin or individual with any type of genital malformation that made physical examination or transvaginal ultrasonography impossible; unable to tolerate MRI **Ascher 1995** Not reported **Bazot 2009** Not reported **Bazot 2013** Not reported **Biscaldi 2014** Exclusion criteria: previous bilateral ovariectomy, previous radiological exams of the bowel requiring contrast media, previous bowel surgery (except appendectomy),				Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear risk Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Managaro 2012b**

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	history of intolerance to iodinated contrast media, renal or hepatic failure, contraindications to MR examination, psychiatric disorders **Chamie 2009** Not reported **Grasso 2010** Not reported **Ha 1994** Not reported **Hottat 2009** exclusion criteria: common contraindications to MRI (pacemaker, metallic foreign bodies, claustrophobia), age < 18 years, postmenopausal status **Manganaro 2012a** Not reported **Managaro 2012b** Not reported **Manganaro 2013** Not reported				A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
	Okada 1995 Not reported **Stratton 2003** Not reported **Sugimura 1993** Not reported **Takeuchi 2005** Not reported **Thomeer 2014** exclusion criteria: use of contraceptives or hormonal suppressive medication, contraindication to MRI (pacemaker, different metallic bodies, claustrophobia), age younger than 18, postmenopausal status				If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Unclear Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Manganaro 2013** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the patient flow have introduced bias? Low risk **Okada 1995** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Stratton 2003** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? Unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? High risk **Sugimura 1993** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability Patient characteristics and setting

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Takeuchi 2005** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? No Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? High risk B. Concerns regarding applicability Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					If a threshold was used, was it pre-specified? Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Unclear Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the patient flow have introduced bias? Unclear risk **Thomeer 2014** A. Risk of Bias Patient Sampling Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? According to the CSR 'Was a two-gate design avoided?' Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? Low risk B. Concerns regarding applicability Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					from the review question? Low concern Reference Standard A. Risk of Bias Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk
Full citation Arrive, L., Hricak, H., Martin, M. C., Pelvic endometriosis: MR imaging,	**Condition** Clinically suspected endometriosis **Sample size**	**Tests** MR Laparoscopy, laparotomy	**Methods** • Laparoscopy, and laparotomy procedure reports, photographs obtained during procedures and histological slides, when available, were	**Results** <u>Pelvic endometriosis:</u> Sensitivity (95% CI): 64% (43 to 82) Specificity (95% CI): 60% (15 to 95)	**Limitations** <u>QUADAS 2</u> Patient Selection A. Risk of Bias Patient Sampling

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Radiology, 171, 687-92, 1989 **Ref Id** 401020 **Country/ies where the study was carried out** USA **Study type** Prospective cohort study **Aim of the study** To analyse the value of MRI in detecting, characterising, and staging endometriosis, including evaluation of endometriosis, endometrial adhesions, and endometrial implants. **Study dates** 1989	N=30 (Consecutive patients) **Characteristics** Not reported **Inclusion Criteria** Clinically suspected endometriosis **Exclusion Criteria** Not reported		reviewed by one of the authors • Degree of severity of endometriosis was classified according to the AFS system • MRI: Spin-echo images were obtained, T1 and T2 predominant images were obtained in all patients • MRI images were analysed and recorded independently, the observers knew only the clinical history of suspected endometriosis • Lesion location, size and shape were recorded. Thickness, signal intensity of the lesion, distinctness of the interface of the lesion with adjacent organs, appearance of the lesion, position of the uterus, and presence of free fluid in the cul-de-sac • Endometrioma was diagnosed when heterogeneous ovarian lesion with multilocularity and/or loss of clear interface with adjacent organs was demonstrated • Haemorrhagic cyst was diagnosed when a		Was a consecutive or random sample of patients enrolled? Yes Was a case-control design avoided? Yes Did the study avoid inappropriate exclusions? Yes Could the selection of patients have introduced bias? Low risk B. Concerns regarding applicability: Patient characteristics and setting Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Source of funding French Foreign Office			unilocular, heterogeneous ovarian lesion demonstrated a clear interface with adjacent organs. • MRI imaging and surgical findings were compared (sensitivity, specificity, accuracy were calculated)		Is the reference standards likely to correctly classify the target condition? Yes Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? yes Did all patients receive the same reference standard? Unclear Were all patients included in the analysis? Yes Could the patient flow have introduced bias? Unclear risk

G.11 Review question: Diagnosis – Surgical diagnosis with or without histological confirmation

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

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Full citation 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 **Ref Id** 401663 **Country/ies where the study was carried out** Germany **Study type** Case-series **Aim of the study** To analyse the accuracy of	**Condition** clinical suspicion of endometriosis **Sample size** n=164 **Characteristics** 59.8% stage I endometriosis 8.5% stage II 17% stage III 14.6%stageIV **Inclusion Criteria** • laparoscopic data on 164 endometriosis patients recorded in the German Complication Register were analysed **Exclusion Criteria** Not reported	**Tests** laparoscopy histological diagnosis	**Methods** The German Complications Register is a computerised database established by the Institute of Natural Intelligence in Bremen which compiles data from 41 German endoscopic surgery centers. In this study only the data from one centre in Kiel was evaluated. Laparoscopy was performed with the patient under general anaesthesia. Magnification was used to get better view of the abdominal wall and the organs of the minor pelvis. Under observation, any lesion was taken as suspicious for endometriosis. To verify diagnosis biopsies were taken by grasping the red black or white lesion and punching it out with punch biopsy forceps. In case of ovarian endometriomas the cysts were enucleated in the typical manner in attempt to extract the endometriotic lesion.	**Results** Endometriosis (number of patients): Positive test: 138/164 (84%) Endometriosis (number of biopsy specimens): Positive test: 142/264 (54%)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s)

Study details	Participants	Tests	Methods	Outcomes and results	Comments
laparsocopic visualisation in diagnosing the various endometriotic sites as confirmed histologically **Study dates** January 1998 to September 2000 **Source of funding** Not reported					Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? low risk **Other information** None
Full citation de Almeida Filho, D. P., de Oliveira, L. J., do Amaral, V. F., Accuracy of laparoscopy for	**Condition** women undergoing laparoscopy for pelvic pain and/or infertility **Sample size**	**Tests** laparoscopy histopathology	**Methods** During the laparoscopy they performed biopsies on anatomical abnormalities that presented the macroscopic appearance	**Results** Sensitivity (95% CI): 98% (95 to 99) Specificity (95% CI): 79% (76 to 82) Endometriosis (number of patients):	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
assessing patients with endometriosis, Sao Paulo Medical Journal, 126, 305-308, 2008 **Ref Id** 416856 **Country/ies where the study was carried out** Brazil **Study type** Some other intervention type **Aim of the study** Cross-sectional study to test the efficacy of laparoscopy alone for diagnosing endometriosis and to evaluate the lateratility of endometriosis among the study population	n=976 **Characteristics** mean age 30.85 (SD 5.54) acute or chronic pelvic pin 98.84% dysmenorrhea 37.39% primary infertility 20% secondary infertility 6.66% **Inclusion Criteria** • subject needed to be in the menacme and presenting pelvic pain, dyspareunia, dysmenorrhea or infertility and the results from complementary tests such as CA125 determination and ultrasound needed to reveal pelvis masses or blood in the pelvis. **Exclusion Criteria** • patients who had not reached menarche yet • menopausal patients • cases of laparosccopic reinterventions		of endometriosis (ie typical lesions such as "powder burn", of reddish colour, light colour or even on fibrotic lesions. The lesions suggestive of endometriosis were biopsied and histopathologically examined in the pathological anatomy department. The endometriosis was staged in accordance with the 1985 American Fertility Society classification, and the staging was compared with the result from the histopathological analysis on the biopsies	Positive test: 337/468 (72%) Negative test: 500/508 (98%)	Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Study dates 1994 to 2004 **Source of funding** None declared					Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? low risk **Other information** None
Full citation Chatman, D. L., Zbella, E. A., Biopsy in laparoscopically diagnosed endometriosis, Journal of Reproductive Medicine, 32, 855-7, 1987 **Ref Id** 380977	**Condition** patients with the primary complaint of pelvic pain **Sample size** n=273 **Characteristics** pain duration 2months-several years 84% aged between 20-40	**Tests** laparoscopy histology	**Methods** Laparoscopy performed under general anaesthesia with the use of a double puncture technique. The severity of the endometriosis was classified according to the criteria of Acosta et al 1973 (Obstet Gynaecol 42:19) Peritoneal and ovarian biopsies were	**Results** <u>Endometriosis (number of patients):</u> Positive test: 74/115 (64%) Only 115 with laparoscopically visualised endometriosis had biopsies 158 were not biopsied because it was thought that	**Limitations** <u>QUADAS 2</u> A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability:

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Country/ies where the study was carried out USA **Study type** Case-series **Aim of the study** To correlate the findings of endometriosis observed at laparoscopy with the histologic diagnosis of specimen obtained at biopsy **Study dates** Not reported more specifically than "over a 4 year period" **Source of funding** Not reported	**Inclusion Criteria** laparoscopy only after a constellation of suggestive symptoms (dysmenorreha, dyspareunia) and/or physical signs (nodularity of the uterosacral ligaments, retroversion of the uterus, enlargement of ovaries)indicated possible presence of the disease **Exclusion Criteria** not reported		performedto obtain histologic confirmation of endometriosis Peritoneal biopsies were performed using Eder 388 biopsy forceps or Olympus 0517 biopsy forceps. Ovarian biopsies performed with Eder 688 ovarian biopsy forceps Pathologic specimens consiting of 5- to 10-mm tissue samples were processed and stained with hematoxylin and eosin. Histologic confirmation of endometriosis was established with light microscopy only in the presence of endometrial glands with or without stroma	biopsy would be superfluous or because endometriotic implants were in areas deemed unsafe for biopsies.	Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? high risk **Other information** None
Full citation 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 **Ref Id** 401276 **Country/ies where the**	**Condition** Women undergoing laparoscopy for pelvic pain **Sample size** N=63, however in n=48 excision of endometriotic lesions was undertaken. In other 15 cases the lesions were either very small or too superficial **Characteristics** Age ranged from 23 to 54 y (50% were older than 35 y) **Inclusion Criteria**	**Tests** Laparoscopy Histology	**Methods** The same operative technique was used in all patients, high-pressure entry technique 25 mmHg using 2-3 ports in addition to the 10 mm umbilical port; 5 mm ports were inserted under direct vision in the right and left iliac fossae lateral to the deep inferior epigastric vessels and one suprapubically.	**Results** <u>Endometriosis (biopsy specimens):</u> Positive histology: 104/132(78.8%) Negative histology: 11/132 (16.7%), 4.5% were non-diagnostic <u>Endometriosis (number of patients):</u> Positive histology: 36/48 (75%) Negative histology: 9/48 (18.7%), 6.3% were non-diagnostic	**Limitations** <u>QUADAS 2</u> A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
study was carried out UK **Study type** Retrospective cohort study **Aim of the study** To determine the correlation between laparoscopic diagnosis of endometriosis and histological confirmation. **Study dates** Not stated **Source of funding** Not stated	• Women undergoing laparoscopy for pelvic pain. **Exclusion Criteria** Not stated				Were the index test results interpreted without knowledge of the results of the reference standard? Unclear If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Other information** None
Full citation 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 **Ref Id** 401118 **Country/ies where the**	**Condition** Consecutive women with pain or infertility **Sample size** N=118 69 women were laparoscopically diagnosed with endometriosis (137 samples taken). **Characteristics** Mean age 29.5 y; mean weight 63.3 kg. **Inclusion Criteria** • Women with pain or infertility **Exclusion Criteria** Not stated	**Tests** Laparoscopy Histology	**Methods** A retrospective analysis of all surgical reports between 1994 and 1999 with the clinical diagnosis of minimal and mild endometriosis. Indications for surgery were pain or infertility. Surgery was performed by 10 surgeons.	**Results** Endometriosis (number of patients): Positive test: 49/69 (42%) Endometriosis (number of biopsy specimens): Positive test: 77/137 (56%)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? NA

Study details	Participants	Tests	Methods	Outcomes and results	Comments
study was carried out Germany **Study type** Retrospective cohort study **Aim of the study** Study has attempted to determine to what extent relevant terms such as pigmented and nonpigmented endometriosis are taken into account during routine surgery, outside research conditions. **Study dates** 1994 to 1999 **Source of funding** Not stated					Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Could the patient flow have introduced bias? low risk **Other information** None
Full citation Emmert, C., Romann, D., Riedel, H. H., Endometriosis diagnosed by laparoscopy in adolescent girls, Archives of Gynecology & Obstetrics, 261, 89-93, 1998 **Ref Id** 401280 **Country/ies where the study was carried out** Germany **Study type** Some other intervention type **Aim of the study** To review the incidence, type and clinical	**Condition** Adolescent girls undergoing laparoscopy/pelviscopy. Indications for laparoscopy included chronic or acute pelvic pain and right-sided lower abdominal pain. For this question only girls with laparoscopic ally diagnosed endometriosis were included (n=37). **Sample size** N = 105 (number of lesions not given) 37 were diagnosed with laparoscopic diagnosed endometriosis and 14 of these received both laparoscopy and histological examination. **Characteristics** Mean age of all 105 girls undergoing surgery: 17.3 years Age range of 37 girls with laparoscopic diagnosed endometriosis: 11-19 yrs	**Tests** Laparoscopy/pelviscopy Histological examination	**Methods** Laparoscopy: 105 adolescent girls with pain underwent laparoscopy/pelviscopy. Each case of endometriosis was staged according to the endoscopic endometriosis classification by Semm (EEC). 37 were diagnosed with endometriosis Histological examination: Of the 37 girls diagnosed with endometriosis after laparoscopy, 14 girls (37.8%) had histological examination of biopsies. No criteria for the histological examination a re provided in the paper.	Results Endometriosis (biopsy specimens): Not given Endometriosis (number of patients): Positive histology: 6/14 (42.8%)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have introduced bias? Y - it is unclear whether the patients were consecutive or chosen based on other factors. No information was provided for why the patients who had samples sent for histological examination (14/37) were chosen and they may have shared risk factors which could cause bias. B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A

Study details	Participants	Tests	Methods	Outcomes and results	Comments
stage of endometriotic lesions of adolescent girls with chronic pelvic pain **Study dates** January 1996 to June 1997 **Source of funding** Not stated	**Inclusion Criteria** • Adolescent girls with indications for laparoscopy included chronic or acute pelvic pain and right-sided lower abdominal pain. **Exclusion Criteria** None stated.				Could the conduct or interpretation of the index test have introduced bias? high risk - Laparoscopy was considered as the gold standard for detection of endometriosis B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern. Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear. Details about the criteria for diagnosis on histological examination are not provided. Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear. Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk. Not enough information is provided in the paper. B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Unclear Were all patients included in the analysis? Unclear - no indication of whether patients were consecutive. Could the patient flow have introduced bias? Low risk **Other information** None
Full citation 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 **Ref Id** 402082	**Condition** Women who presented with chronic pelvic pain or known endometriosis (diagnosed histologically or by visualization) refractory to medical treatment at the Department of Gynecologic Surgery at Mayo Clinic Scottsdale. **Sample size** N=44 **Characteristics** Age at operation: 14-48 years, mean 33 years (SD 9) Parity: 0 - 57% 1 - 11%	**Tests** Laparoscopy- visual appearance Histology	**Methods** Laparoscopy: all areas of typical and atypical endometriosis were documented on a pelvic diagram (lesion type, location), completely excised, fixed in formalin, assessed pathologically Endometriosis definition: presence of glands and stroma Mayo pathologists blinded to the type of lesion (if any) Lesion definitions: puckered pigmented, scarred, red, vesicular, peritoneal pockets, adhesions and yellow lesions	**Results** Endometriosis: Sensitivity (95% CI): 97% (90 to 100) Specificity (95% CI): 77% (72 to 82) Endometriosis (number of biopsy specimens): Positive test: 67/138 (49%) Negative test: 240/242 (99%)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? low risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Country/ies where the study was carried out USA **Study type** Prospective cohort study **Aim of the study** To correlate the diagnosis of endometriosis on the basis of visualisation at laparoscopy with the pathologic diagnosis. **Study dates** July 1997-March 1999. **Source of funding** None described.	2 - 30% 4 - 2% Prevalence of previous treatments: laparoscopy and ablation on excision, once n=7, twice n=6, three time n=1, hysterectomy n=7, leuprolide n=6 All women presented with a primary complaint of pelvic pain, dysmenorrhea, or dyspareunia **Inclusion Criteria** As per condition listed above **Exclusion Criteria** • Recently completed therapy with gonadotropin releasing hormone agonists (within 6 months of laparoscopic evaluation)		Normal pelvic peritoneum also sampled- multiple site specific biopsies (R and L USL, post. and ant. of the cul-de-sac, ovarian fossae, peritoneum overlying right psoas muscle If abnormal peritoneum no additional samples taken No abnormal peritoneum: 9 biopsy specimens (~0.5cm)taken at the specified sites Disease stage: American Fertility Society Classification (AFS), visual and histological scores (substracting the score of lesions that were visually consistent with endometriosis but not confirmed on pathology) Ovarian endometriomas excised and histology examination Pathology examination: 1 of 6 pathologists and re-reviewed by 1 pathologist Specimen fixed in formalin, embedded in paraffin and 3-4µm sections obtained every 50-60µm Sections stained in hematoxylin and eosin		If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y Could the reference standard, its conduct, or its interpretation have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
			4-6 sections per specimen - evaluated by light microscopy		Were all patients included in the analysis? Y Could the patient flow have introduced bias? low risk **Other information** AFS scores were also reported.
Full citation 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 **Ref Id** 401717 **Country/ies where the study was carried out** Belgium **Study type** Some other intervention type	**Condition** Women undergoing laparoscopy for infertility. **Sample size** N=118 women in total study Reported here are results from the 86 women had laparoscopy diagnosed endometriosis (138 biopsies). **Characteristics** Age range and other baseline characteristics are not given. **Inclusion Criteria** • Patients who were undergoing laparoscopy for infertility **Exclusion Criteria** None stated.	**Tests** Laparoscopic surgery Histological examination	**Methods** Laparoscopy: peritoneal biopsies were taken from areas of the pelvic peritoneum bearing foci of endometriosis (brownish, bluish, or purplish hemorrhagic areas often associated with stellate scarring) and/or from areas of visually normal peritoneum (uterosacral ligaments). Biopsies were taken with a biopsy punch forceps and were 3 to 5mm large. The laparoscope was placed 4 to 5 cm from the peritoneum to evaluate its surface. Thereafter, the laparoscope was placed close to the peritoneum to achieve some magnification. The peritoneum was considered as normal peritoneum if no lesion described before was seen.	**Results** Endometriosis (biopsy specimens): With macroscopically visible endometriotic lesion: Positive histology: 80/86 (93.0%) With macroscopically normal peritoneum: Positive histolology: 7/52 (13.5%) Endometriosis (number of patients): Positive histology: 80/86 (93.0%)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Unclear – no exclusion reasons given Could the selection of patients have introduced bias? Unclear – no information how patients were selected B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Aim of the study To evaluate histologically, biopsies of peritoneal endometriosis and of visually normal peritoneum taken from patients undergoing a laparoscopy for infertility. **Study dates** Not stated. **Source of funding** Not stated.			Histological examination: All biopsy specimens were fixed in formaldehyde and embedded in paraffin. Three micrometer serial sections were stained with Gomori's Trichrome and examined, on a blind basis, with a Leitz Orthoplan microscope (Leitz, Wetzlar, West Germany). In all cases, the mitotic index was calculated as previously described by counting mitotic figures (prometaphase, metaphase, anaphase, and telophase) for 2,000 epithelial cells per biopsy. The epithelial height was measured with the help of an ocular micrometer. Fifty cells were selected in which the plane of section clearly passed through the cell nucleus parallel to the longitudinal axis of the cell. Blind interpretation of histological results was done systematically. Results (epithelial height) were' expressed as the mean ± SD. The x2 test and the median test were		Could the conduct or interpretation of the index test have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Y – papers states the assessors of the histological examination was ‘blinded’. Could the reference standard, its conduct, or its interpretation have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
			used for statistical analysis. The microscopic criteria for endometriosis were the presence of both glandular epithelium and stroma		Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Other information** None
Full citation 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 **Ref Id** 417376 **Country/ies where the study was carried out** UK **Study type** Prospective cohort study	**Condition** Women with chronic pelvic pain. **Sample size** N=62 but biopsies from 3 patients were unsuitable for histological evaluation and were excluded from the study **Characteristics** No data on sample characteristics **Inclusion Criteria** • Women with chronic pelvic pain **Exclusion Criteria** Not stated	**Tests** Laparoscopy Histology	**Methods** Preoperative bowel preparation was given to all patients in anticipation of surgical intervention. All procedures were done under the direct supervision of the same senior laparoscopic surgeon.	**Results** Endometriosis (biopsy specimens): positive test 85/150 (56.7%) Endometriosis (patients): positive test 43/59 (72.9%)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? Y Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Aim of the study To histologically evaluate peritoneal lesions laparoscopically suspicious for endometriosis, which had been excised from different pelvic anatomical sites in patients with the presenting complaint of chronic pelvic pain, irrespective of previous pelvic surgery or the earlier diagnosis of endometriosis. **Study dates** October 1997 to October 1998 **Source of funding** Not stated					Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? No Could the patient flow have introduced bias? high risk **Other information**

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					None
Full citation 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 **Ref Id** 402778 **Country/ies where the study was carried out** USA **Study type** Prospective cohort study **Aim of the study** To better understand the	**Condition** Women with chronic pelvic pain thought to be due to endometriosis. **Sample size** N=77 **Characteristics** Not given **Inclusion Criteria** Women with chronic pelvic pain undergoing surgery as part of a clinical trial of a potential new treatment for endometriosis. All women had had pelvic pain for at least 6 months and were otherwise healthy, with regular menstrual cycles. **Exclusion Criteria** Not stated	**Tests** Laparoscopy Histology	**Methods** All women entered into the study underwent laparoscopy at the same University hospital. At laparoscopy, the goal was to remove all visible implants that might be endometriosis. all lesions suspicious for endometriosis were excised by using a contact neodymium:yttrium-aluminum-garnet laser after careful, systematic inspection of the peritoneal surfaces throughout the pelvis and the abdomen.	**Results** Endometriosis (number of patients): Positive test: 57/65 (88%) Endometriosis (number of biopsy specimens): Positive test: 189/314 (60%) No negative test results reported No sensitivity or specificity reported	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
clinical characteristics of histologically proven endometriosis lesions. To develop criteria that would predict histologic confirmation of endometriosis and to determine the accuracy of visualization of lesions for making a diagnosis. **Study dates** Not stated **Source of funding** Supported by the intramural program of the National Institute of Child Health and Human Development					Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? low risk **Other information** None
Full citation Jansen, R. P., Russell, P., Nonpigmented endometriosis:	**Condition** Women who underwent laparoscopy for infertility (n=70) or other indications (n=7) including pelvic pain	**Tests** Laparoscopy Histology	**Methods** The patients were a subset of those seen between June 1982 and September 1984 in an	**Results** Endometriosis (number of biopsy specimens):	**Limitations** QUADAS 2 A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
clinical, laparoscopic, and pathologic definition, American Journal of Obstetrics & Gynecology, 155, 1154-9, 1986 **Ref Id** 401456 **Country/ies where the study was carried out** Australia **Study type** Prospective cohort study **Aim of the study** To describe the morphologic characteristics and clinical importance of peritoneal lesions that have the histologic features of endometriosis but are devoid of	and assessment for sterilization reversal **Sample size** N=77 **Characteristics** No description of the study population **Inclusion Criteria** • women undergoing laparoscopy for infertility or other indications including pelvic pain and assessment for sterilization reversal **Exclusion Criteria** Not stated		endocrine-infertility practice. A full medical history was obtained for all patients, including responses to questions for dysmenorrhea, deep dyspareunia, and premenstrual spotting.	Positive test: 73/137 (53%) No negative test results reported No sensitivity or specificity reported	Was a consecutive or random sample of patients enrolled? unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
the pigmented stigmas typical of this disease. **Study dates** June 1982 and September 1984 **Source of funding** Not stated					Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? low risk **Other information** None
Full citation Vercellini, P., Vendola, N., Bocciolone, L., Rognoni, M. T., Carinelli, S. G., Candiani, G. B., Reliability of the visual diagnosis of ovarian	**Condition** Women who underwent a laparotomy for an "ovarian cyst" **Sample size** N=245 **Characteristics**	**Tests** Laparotomy (visual) Histology of ovarian cyst	**Methods** Endometrioma visual definition: ovarian cyst no >12cm in diametre adhesions to the pelvic side wall and/or the posterior broad ligament 'powder burns' and minute red or blue spots with	**Results** <u>Endometrioma (number of ovarian cysts):</u> Positive test: 213/218 (98%) Negative test: 106/113 (94%) Sensitivity (95% CI): 97% (94 to 99)	**Limitations** <u>QUADAS 2</u> A. Risk of Bias Was a consecutive or random sample of patients enrolled? unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Y

<table>
<tr><th>Study details</th><th>Participants</th><th>Tests</th><th>Methods</th><th>Outcomes and results</th><th>Comments</th></tr>
<tr>
<td>endometriosis, Fertility & Sterility, 56, 1198-200, 1991
Ref Id
402067

Country/ies where the study was carried out
Italy

Study type
Case-series

Aim of the study
To compare the surgical and histological diagnoses in women of reproductive age who underwent laparotomy for ovarian cysts in the last 5 years with the aim of evaluating the reliability of the visual diagnosis of endometrioma.</td>
<td>Median age 29 years.
<table>
<tr><th>Characteristic</th><th>Endometrioma group n=138</th><th>Non endometrioma group n=77</th><th>Mixed group n=30</th></tr>
<tr><td>Median age, yrs (range)</td><td>30 (23-40)</td><td>29 (20-40)</td><td>28 (21-38)</td></tr>
<tr><td>Median parity (range)</td><td>0.4 (0-4)</td><td>0.5 (0-3)</td><td>0.3 (0-3)</td></tr>
<tr><td>Surgical intervention
Cyst enucleation
Unilateral
Bilateral
Unilateral SO</td><td>
93
57
36
18
7
20</td><td>
48
44
4
16
1
12</td><td>
26
-
26
-
-
4</td></tr>
</table>
</td>
<td></td>
<td>adjacent puckering on the surface
tarry, thick, chocolate coloured fluid content
Histology
Cysts enucleated or removed with the ovary
fixed in formalin immediately and embedded in paraffin
≥10 serial sections for each specimen, hematoxylin and eosin stained
Light microscope: 10X and 40X magnifications
Ovarian endometrioma definition: ≥2 of the following characteristics: endometrial eptithelium, endometrial glands or gland like structures, endometrial stroma, hemosiderin laden macrophages</td>
<td>Specificity (95% CI): 95% (90 to 99)</td>
<td>Could the selection of patients have introduced bias? unclear risk
B. Concerns regarding applicability:
Are there concerns that the included patients and setting do not match the review question? low concern
Index Test
A. Risk of Bias
Were the index test results interpreted without knowledge of the results of the reference standard? unclear
If a threshold was used, was it pre-specified? Y
Could the conduct or interpretation of the index test have introduced bias? unclear risk
B. Concerns regarding applicability
Are there concerns that the index test, its conduct, or interpretation differ from the review question? low concern
Reference Standard
A. Risk of Bias
Target condition and reference standard(s)
Is the reference standards likely to correctly classify the target condition? Y
Were the reference standard results interpreted without knowledge of the results of the index tests? unclear
Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk
B. Concerns regarding applicability</td>
</tr>
</table>

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Study dates January 1986-December 1990 **Source of funding** None described.	TAH and unilateral SO TAH and bilateral SO **Inclusion Criteria** • 20-40 years old • Absence of clinicial and/or ultrasound suspicions of malignancy • First laparotomy except for appendectomy • Non administration of steroid or estrogen suppressing drugs in the preceding 6 months • availability of adequate tissue for histologic study for each of the ovarian cysts diagnosed at laparotomy **Exclusion Criteria** None described				Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? low risk **Other information** None
Full citation Fernando, S., Soh, P. Q., Cooper, M.,	**Condition** Women with suspected endometriosis because of pain or infertility	**Tests** Laparoscopy Histology	**Methods** This study is a part of an longitudinal cohort study which was aiming to	**Results** Endometriosis (biopsy specimens):	**Limitations** QUADAS 2 A. Risk of Bias

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Evans, S., Reid, G., Tsaltas, J., Rombauts, L., Reliability of visual diagnosis of endometriosis, Journal of Minimally Invasive Gynecology, 20, 783-9, 2013 **Ref Id** 401307 **Country/ies where the study was carried out** Australia **Study type** Prospective cohort study **Aim of the study** The authors investigated whether the accuracy of visual diagnosis is affected by disease stage, accounting for other covariates.	**Sample size** N=431 **Characteristics** Patient mean (SD) age was 31.8 (7.2) and BMI was 23.6 (4.5). The median number of previous laparoscopic and/or laparotomic procedures was 1 (range, 0-8), and median parity was 0 (range, 0-7). **Inclusion Criteria** • Women with suspected diagnosis of endometriosis because of pain or infertility before laparoscopy. **Exclusion Criteria** • Patients were excluded before laparoscopy if they had a suspected gynecologic malignancy, known current or chronic relapsing pelvic inflamatory disease, or current pregnancy or if they were unable to provide informed consent.		assess pain and fertility outcomes after laparoscopic surgery performed to treat endometriosis. 533 patients were identified as potentially eligible for enrollment on the basis of a presumed diagnosis of endometriosis because of pain or infertility before laparoscopy. Of these, 62 either did not have any visual features of endometriosis or, if biopsies were taken, none contained histologically proven endometriosis. In another 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. Thus, 102 patients were excluded from this analysis, leaving 431 women, from whom a total of 1439 biopsy specimens were obtained.	Positive test: 1082/1439 (75.2%)	Was a consecutive or random sample of patients enrolled? unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? unclear Could the selection of patients have introduced bias? unclear risk B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? unclear If a threshold was used, was it pre-specified? NA Could the conduct or interpretation of the index test have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Study dates September 2003 to July 2007 **Source of funding** Supported by an unconditional grant from the Australian Gynaecological Endoscopy & Surgery Society awarded to the AWARE group.			Preoperatively, all patients completed a questionnaire to collect demographic, biometric and clinical data including age, BMI, and gynecologic and medical history.		Were the reference standard results interpreted without knowledge of the results of the index tests? unclear Could the reference standard, its conduct, or its interpretation have introduced bias? unclear risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Y Could the patient flow have introduced bias? low risk **Other information** None
Full citation Stripling, M. C., Martin, D. C., Chatman, D. L., Zwaag, R. V., Poston, W. M., Subtle appearance of pelvic endometriosis,	**Condition** Postoperative diagnosis of endometriosis. The paper does not state the reasons for the women undergoing laparoscopy/laparotomy. **Sample size** N = 109 (164 lesions)	**Tests** Laparoscopy Laparotomy +/- laparoscopy Histological examination	**Methods** Lesion excision: Patients undergoing laparotomy and/or laparoscopy had suspected endometriosis lesions removed using either the C02 laser, scissors, or biopsy forceps.	**Results** Endometriosis (biopsy specimens): Positive histology: 148/164 (90.2%) Endometriosis (number of patients): Positive histology: 106/109 (97.2%)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Unclear Could the selection of patients have

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Fertility & Sterility, 49, 427-31, 1988 **Ref Id** 417800 **Country/ies where the study was carried out** USA **Study type** Retrospective cohort study **Aim of the study** To investigate whether lesions excised by laparotomy or laparoscopic surgery were endometriosis (diagnosed histologically) and to determine the rates. **Study dates** January 1986 to October 1986	**Characteristics** The paper does not provide baseline characteristics (e.g. age, reason for laparoscopy/laparotomy or any other risk factors) **Inclusion Criteria** • Consecutive patients with a postoperative diagnosis of endometriosis **Exclusion Criteria** None stated.		Histologic examination. Excised lesions were sent to the pathology department and standard hematoxylin and eosin stains were performed on all specimens. Endometriosis was diagnosed when both glands and stroma were found. Trichrome stains were performed on four fibromuscular scar lesions for the analysis of the fibrous and muscular components.		introduced bias? Y B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Unclear risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? No Could the reference standard, its conduct, or its interpretation have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Source of funding Not stated.					Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Y Were all patients included in the analysis? Unclear Could the patient flow have introduced bias? Low risk **Other information** None
Full citation 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	**Condition** Consecutive patients who were undergoing laparoscopy for infertility (group 1, n = 52), chronic pelvic pain (group 2, n = 18) or tubal sterilization (group 3, n = 30), **Sample size** N = 100 women (119 biopsies, of which 19 were of lesions laparoscopically diagnosed as endometriosis) Group 1 - infertility:n = 52 (26 had laparoscopically diagnosed endometriosis) Group 2 - chronic pelvic pain: n = 18 (8 had laparoscopically diagnosed endometriosis)	**Tests** Laparoscopy Histological examination	**Methods** Laparoscopy: systematic laparoscopic evaluation of all pelvic peritoneal surfaces was carried out. The laparoscope was placed 4-5 cm from the peritoneum to evaluate its surface; thereafter, the laparoscope was placed close to the peritoneum to achieve some magnification. Peritoneum eligible for study had to have a perfectly smooth surface with no fibrosis or abnormal vascular patterns, and transparency with no associated colour or suggestion of sub-peritoneal cystic structures. Systematic	**Results** Although it indicates that 47 women had laparoscopically diagnosed endometriosis the paper states "Biopsy of the endoscopically suspected endometriosis in 19 patients revealed the presence of endometrial glands and stroma in 17 cases (89.5%), while the two other biopsies showed fibrosis with haemosiderin-laden macrophages and endometrium-like stroma alone respectively."	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Unclear - although the collection of 'endometriotic' biopsies from people with laparoscopically diagnosed endometriosis did not occur in all cases (19/47 = 40.4%). No details about why some patients had biopsies taken and others didn't is not reported in the paper. Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Unclear - as per question 1; above it is not clear the criteria for selecting the 19/47 patients with laparoscopically diagnosed endometriosis were identified.

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Ref Id 417928 **Country/ies where the study was carried out** Spain **Study type** Prospective cohort study **Aim of the study** The specific aims of this study were (1) to investigate prospectively the prevalence of endometriosis at laparoscopy in the three groups of patients (infertile patients, patients with chronic pelvic pain and asymptomatic fertile women) and (2) to evaluate histologically biopsies of	Group 3 - tubal sterilization: n = 30 (13 had laparoscopically diagnosed endometriosis) **Characteristics** Age: Infertility: 32.1 ± 3.9 years; Chronic pelvic pain: 32.6 ± 4.9 years; tubal sterilization: 33.8 ± 4.8 years Mean parity: Chronic pelvic pain: 1.5 (range 0-6); tubal sterilization: 2.4 (range 1-13) No patients had been pregnant within the past year. Hormonal treatment for endometriosis No patients had been treated with hormonal treatment for endometriosis. **Inclusion Criteria** • Consecutive patients who were undergoing laparoscopy for infertility, chronic pelvic pain or tubal sterilization. **Exclusion Criteria** None stated.		biopsy of visually normal peritoneum overlying the uterosacral ligaments, biopsies of suspicious lesions were taken when the visual diagnosis of endometriosis was in doubt (19 cases). Biopsies were taken with a 5-mm Wolf punch biopsy forceps. Histological examination: All biopsy specimens were evaluated by the same expert gynaecological pathologist who was unaware of diagnostic groups. Several step sections (one every 100-150 µm) were made of each specimen. Standard haematoxylin and eosin stains were performed on all specimens. Endometriosis was diagnosed by the presence of both endometrial glands and stroma. Intra-mesothelial endometriosis (surface endometrial epithelium without stroma and glands) was not considered in the present study.	Positive histology: 17/19 (89.5%); Negative histology: 2/19 (10.5%) **Infertility** Endometriosis from 'NORMAL uterosacral ligaments' (number of patients): Positive histology: 3/26 (11.5%); Negative histology: 23/26 (88.5%) **Chronic Pelvic Pain** Endometriosis from 'NORMAL uterosacral ligaments' (number of patients): Positive histology: 1/8 (12.5%); Negative histology: 7/8 (87.5%) **Tubal sterilisation** Endometriosis from 'NORMAL uterosacral ligaments' (number of patients):	Could the selection of patients have introduced bias? Y B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear - as only 19 biopsies of endometriotic lesions were collected it is unclear whether the assessors completing outcome assessment

Study details	Participants	Tests	Methods	Outcomes and results	Comments
visually normal peritoneum taken from all these women, and (3) to investigate the relation between oral contraception and the risk of pelvic endometriosis in those three well-defined groups of patients.. **Study dates** Not stated. **Source of funding** Not stated.				Positive histology: 1/13 (7.7%); Negative histology: 12/13 (92.3%)	knew that these were people with laparoscopically diagnosed endometriosis. Could the reference standard, its conduct, or its interpretation have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? No - only 19/47 patients had the reference standard applied. Were all patients included in the analysis? No Could the patient flow have introduced bias? high risk **Other information** None
Full citation Cornillie, F. J., Oosterlynck, D., Lauweryns, J. M., Koninckx, P. R., Deeply infiltrating pelvic endometriosis:	**Condition** Consecutive women undergoing laparoscopies for infertility, pain or both. **Sample size** N= 179 laparoscopies. Infertility n = 105 ; pain n =	**Tests** Laparscopy Histological examinatio n	**Methods** Laparoscopy: Pelvic implants were excised with a CO2 laser and the depth of infiltration of endometriosis was accurately assessed during and after excision	**Results** Endometriosis (number of patients with lesions with depth greater than 3mm): Positive histology: 84/110 (76.4%)	**Limitations** QUADAS 2 A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y Was a case-control design avoided? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
histology and clinical significance, Fertility & Sterility, 53, 978-83, 1990 **Ref Id** 403149 **Country/ies where the study was carried out** Belgium **Study type** Prospective cohort study **Aim of the study** To investigate systemically the histological characteristics and the activity of deeply infiltrating pelvic endometriosis. **Study dates** October 1988 to July 1989	60; infertility AND pain n = 14. Total laparoscopically diagnosed with endometriosis: 142/179 (80.4%): Infertility n=81; pain n=49; infertility AND pain n= 12 Biopsy samples taken from N=110 women with lesions penetrating deeper than 3mm **Characteristics** Age or other risk factors were not stated in the paper. **Inclusion Criteria** • Patients in whom laparoscopy was performed for infertility, pelvic pain or both. Biopsies were taken from all lesions penetrating deeper than 3mm. **Exclusion Criteria** • Women with ovarian endometriosis only and women using medical suppressive therapy for endometriosis were excluded.		by comparing the depth of excision and the height of the biopsy with the graded tip of a second puncture instrument. Histological examination: Biopsies were fixed in phosphate-buffered formalin, dehydrated through alcohols, and embedded in paraffin. The deep implants were divided into two tissue blocks, from which at least 2 sections were made perpendicularly to the peritoneal surface, and were stained with hematoxylin and eosin. All biopsies were studied by one of the authors and endometriosis was diagnosed only when ectopic glands together with stroma were found		Did the study avoid inappropriate exclusions? Y - although those with endometrial lesions of 3mm or less were not included in the results. Could the selection of patients have introduced bias? No B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? Low concern - although may not be representative of all patients (i.e those without deep endometrial lesions) Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Y

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Source of funding Not stated.					Were the reference standard results interpreted without knowledge of the results of the index tests? No - it appears samples were only taken from people with laparoscopically diagnosed endometriosis. Could the reference standard, its conduct, or its interpretation have introduced bias? low risk B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? No - although 144 people had laparoscopically diagnosed endometriosis, only those with lesion depth greater than 3mm had histological examination. Were all patients included in the analysis? Y (all patients with lesion depth greater than 3mm) Could the patient flow have introduced bias? Low risk **Other information** Results given are only for deep lesions of greater than 3mm.
Full citation	**Condition**	**Tests**	**Methods**	**Results**	**Limitations**

Study details	Participants	Tests	Methods	Outcomes and results	Comments
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 **Ref Id** 403331 **Country/ies where the study was carried out** USA **Study type** Retrospective cohort study **Aim of the study** To assess a correlation between endosalpingiosis and pelvic pain. **Study dates** August 1992 – October 1993.	Patients undergoing laparoscopy for chronic pelvic pain. **Sample size** N: 51 surgeries completed (due to the nature of the study this is likely to be 51 separate patients). 37 of 51 cases showed some evidence of laparscopically diagnosed endometriosis. **Characteristics** Not clearly stated. The paper reports: "The patients with endosalpingiosis were similar in age to those with biopsy-proven endometriosis and those without evidence of endometriosis, averaging 35.0, 34.3, and 32.9, years, respectively." **Inclusion Criteria** • Patients with chronic pelvic pain. **Exclusion Criteria** None stated.	Laparoscopy Histological examination	Laproscopy: Details about technique are not provided in the paper. The paper only says that surgical approach to endometriosis involved excision of nearly all visible endometriosis, to enable the authors to evaluate the rate and location of endosalpingiosis found in association with chronic pelvic pain. Histological examination: Details of method and criteria are not provided. The paper only says that all specimens were fixed in paraffin, underwent hematoxylin and eosin staining.	<u>Endometriosis (biopsy specimens):</u> Positive histology: 21/37 (56.8%) <u>Endometriosis (number of patients):</u> Positive histology: 21/37 (56.8%)	<u>QUADAS 2</u> A. Risk of Bias Was a consecutive or random sample of patients enrolled? Y – consecutive samples although patients were included based on an a retrospective review Was a case-control design avoided? Y Did the study avoid inappropriate exclusions? Unclear – no exclusion reasons provided Could the selection of patients have introduced bias? Unclear – results from one surgeon only B. Concerns regarding applicability: Are there concerns that the included patients and setting do not match the review question? Low concern Index Test A. Risk of Bias Were the index test results interpreted without knowledge of the results of the reference standard? Y If a threshold was used, was it pre-specified? N/A Could the conduct or interpretation of the index test have introduced bias? Unclear – no details of the intervention test were provided. B. Concerns regarding applicability Are there concerns that the index test, its conduct, or interpretation differ from the review question? Low concern

Study details	Participants	Tests	Methods	Outcomes and results	Comments
Source of funding Not stated.					Reference Standard A. Risk of Bias Target condition and reference standard(s) Is the reference standards likely to correctly classify the target condition? Unclear – lack of information provided in the paper. Were the reference standard results interpreted without knowledge of the results of the index tests? Unclear – no information provided Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear- lack of information given. B. Concerns regarding applicability Are there concerns that the target condition as defined by the reference standard does not match the question? Low concern Flow and Timing A. Risk of Bias Was there an appropriate interval between index test and reference standard? Y Did all patients receive the same reference standard? Unclear – no information given Were all patients included in the analysis? Y Could the patient flow have introduced bias? Low risk **Other information**

Study details	Participants	Tests	Methods	Outcomes and results	Comments
					Note: the paper was really looking for the rate of endosalpingiosis.

G.12 Review question: Staging Systems

What is the effectiveness of using endometriosis-staging systems to guide treatment of endometriosis?

No clinical evidence was identified for this review.

G.13 Review question: Pharmacological management – Analgesics

What is the effectiveness of analgesics for reducing pain in women with endometriosis, including recurrent and asymptomatic endometriosis?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Kauppila, A., Ronnberg, L., Naproxen sodium in dysmenorrhea secondary to endometriosis, Obstetrics & Gynecology, 65, 379-83, 1985 **Ref Id** 346834 **Country/ies where the study was carried out** Finland **Study type** RCT **Aim of the study**	**Sample size** N = 24 women **Characteristics** N = randomized: 24 N= analysed: 20 **Inclusion criteria** Women with endometriosis classified by the American Fertility Society (mild endometriosis n=7; moderate endometriosis n=8; severe endometriosis n=6). Women were diagnosed by pelvic examination, history of menstrual distress and by direct visualisation of pelvic regions at laporoscopy or laparatomy	**Interventions** **Group 1** (Naproxen Sodium - NSAID - was given for 2 menstrual cycles, then crossover to placebo for 2 menstrual cycles) **Group 2** (Placebo was given for 2 menstrual cycles, then crossover to Naproxen Sodium - NSAID	**Details** **Overall Pain relief:** all self-reported using a questionnaire completed by the patient immediately after each menstrual cycle	**Results** <u>Overall pain relief</u> Naproxen sodium: 10/11 (90.9%) Placebo: 5/8 (62.5%) RR 1.45 (0.82 to 2.57)* <u>Unintended effects of treatment</u> Naproxen sodium: 4/11 (36.4%) Placebo: 7/9 (77.8%) RR 0.47 (0.2 to 1.1)* <u>Supplementary analgesia needed</u> Naproxen sodium: 1/11 (9.1%) Placebo: 2/8 (25%)	**Limitations** Adequate sequence generation: unclear Allocation concealment: unclear Blinding: moderate risk of bias Incomplete outcome data: low risk of bias Free of selective reporting: unclear risk of bias Free of other bias: high risk of bias **Other information** None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study dates **Source of funding**	**Exclusion criteria** not clear	- for 2 menstrual cycles		RR 0.36 (0.04 to 3.35)* * Calculated by NGA technical team from first period results	

G.14 Review question: Pharmacological management – Neuromodulators

2 **What is the effectiveness of neuromodulators for treating endometriosis, including recurrent and asymptomatic endometriosis?**

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation 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 **Ref Id** 405528	**Sample size** Assigned to bupivacaine, n=32; n=2 lost to follow-up; analysed, n=30 Assigned to placebo, n=30; analysed, n=30 **Characteristics** Bupivacaine, n=30 / Placebo, n=30 / P-value Age: 32.8 ±5.0 / 33.0 ±2.6 / 0.63 Parity: 2.7 ±1.2 / 3.0 ±1.1 / 0.39 Body mass index: 27.2 ±2.1 / 29 ±1.0 / 0.65 Laparoscopic stage: / / Stage 1: 14 / 16 /	**Interventions** Buivacaine: 10ml diluted bupivacaine (0.25%; Marcaine, Astra Zenica, Istanbul, Turkey) plus 100ml Ringer solution, infused through a catheter over 15 to 20 minutes Placebo: 10ml placebo infusion (sterile water) plus 100ml Ringer solution The allocated study solution was provided to the surgeon	**Details** Participants were randomly assigned 1:1 to bupivacaine or placebo according to computer-generated randomisation sequence using numbered, sealed envelopes. All participants and investigators were masked to group allocations, including during data analysis. One treatment was given before ovulation on day 7 to 12 of their cycle. Under paracervical block and using Ringer	**Results** **Bupivacaine (n=30)** VAS (1 to 10), Mean (95% confidence interval), p-value is comparison with baseline Baseline: 7.7 (7.9 to 8.2) 1 month: 6.1 (5.5 to 6.3), P<0.05 2 months: 5.6 (5.8 to 6.0), P<0.01 3 months: 5.4 (4.9 to 5.0), P<0.001 Verbal rating scale (1 to 100), p-value is comparison with baseline Baseline: 90.2 (90.5 to 91.9) 1 month: 35.4 (29.3 to 41.6), P<0.05 2 months: 34.2 (28.6 to 39.8), P<0.01 3 months: 38.6 (32.4 to 44.8), P<0.001 **Placebo (n=30)** VAS (1 to 10), Mean (95% confidence interval), p-value is comparison with baseline Baseline: 7.9 (8.2 to 6.8) 1 month: 7.4 (7.5 to 6.7), P<0.05	Limitations Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments

Country/ies where the study was carried out

Mansoura, Egypt

Study type

Randomised, placebo-contolled, double-blind study

Aim of the study

To assess the effectiveness of hysteroscopic-guided pertubal diluted bupivacaine infusion for endometriosis-associated chronic pelvic pain

Study dates

1 June 2010 and 30 July 2013

Source of funding

Not reported

Stage 2	10	8	
Stage 3	4	4	
Stage 4	2	2	

Patients stopped all analgesics before beginning the study

Inclusion criteria

- CPP for at least six months, pain score on the visual analogue scale (VAS) of more than 5 (0 to 10 scale), laparoscopically confirmed stage I to IV pelvic endometriosis and patent fallopian tubes

Exclusion criteria

- Younger than 18 years of age, any hormonal therapy in previous 3 months, a desire to conceive within 1 year, occulded fallopian tubes with or without pelvic adhesions, non-gynecological causes of CPP (intestinal, urinary or musculoskeletal), and known hypersensitivity or contraindications to bupivacaine or any amide local anesthetic agent.

intraoperatively by senior nursing staff. Solutions were indistibguishable and were preloaded into identical unlabelled Ringer solution bottles.

solution as a uterine distending medium, an office hysteroscope was passed and one tubal orifice was identified. Under hysteroscopic guidance, a 3-Fr ureteric catheter was introduced, cannulated through the tubal ostium, and passed proximally for 2 to 3cm. After successful cannulation, the participants received study treatment or placebo intraoperatively. No adjunctive measures or analgesics were given after treatment. Follow-up visits were made at 1, 2 and 3 months. All participants completed a daily diary about pain during the month preceding the procedure and follow-up visits. They

2 months: 7.5 (7.9 to 6.8), P<0.01

3 months: 7.7 (7.5 to 6.6), P<0.001

<u>Verbal rating scale (1 to 100), p-value is comparison with baseline</u>

Baseline: 91.8 (91.3 to 92.3)

1 month: 91.2 (90.5 to 91.9), P<0.05

2 months: 89.9 (92.1 to 93.1), P<0.01

3 months: 90.2 (92.0 to 88.9), P<0.001

Patient satisfaction at 3 months:

Degree of satisfaction	Bupivacaine (n=30)	Placebo (n=30)	P-value (x2 test)
Satisfied	22	2	0.18
Uncertain	4	2	0.32
Dissatisfied	4	26	0.36

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td>provided a subjective assessment of the severity of pelvic pain on a VAS (0 - no pain to 10 - severe pain). mean VAS scores for the month were calculated for each patient. At monthly follow-up appointment, participants provided a monthly pain score on a verbal rating scale (VRSmonthly) (0- no pain to 100 - maximum pain).</td><td></td><td></td></tr>
<tr><td>Full citation
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</td><td>Sample size
Lignocaine, n=24; Placebo, n=18 (ITT)

Characteristics
Placebo
Age, mean (SD)=33.4 (4.4)
Weight (kg), mean (SD)= 67.6 (12.2)
Height (cm), mean (SD)=167.4 (8.6)
Duration of endometriosis (years), mean (SD)=4.25 (4.51)
Number of smokers=0
VAS at inclusion, mean (SD)=78.22 (18.62)</td><td>Interventions
Study treatment: pertubation with lignocaine 1 mg/ml in Ringer solution
Placebo: pertubation with Ringer solution
Three treatments given preovulatory on cycle day 6</td><td>Details
At the first visit baseline measurements were collected. At the second visit, patients were randomised sequentially in blocks of treatment (three placebo and four study treatment). The treatment was given over three sequential</td><td>Results
<u>EPH-30 questionnaire baseline:</u>
<table>
<tr><td>EHP-30 dimension</td><td>n</td><td>Lidocaine, Mean (SD)</td><td>n</td><td>Placebo, Mean (SD)</td></tr>
<tr><td>Pain</td><td>23</td><td>51.7 (20.0)</td><td>17</td><td>50.8 (19.9)</td></tr>
<tr><td>Control and powerlessness</td><td>23</td><td>59.6 (23.5)</td><td>18</td><td>67.1 (17.9)</td></tr>
<tr><td>Emotional well-being</td><td>20</td><td>54.2 (15.8)</td><td>18</td><td>53.7 (18.1)</td></tr>
</table></td><td>Limitations
<u>Withdrawals</u>
Lignocaine: after 6 months (n=4); 2 pregnant, 1 did not fill in EHP-30 at baseline and 1 did not fill in EHP-30 at six months. After 12 months (n=8); 2</td></tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td>Scandinavica, 92, 1375-82, 2013
Ref Id
338611

Country/ies where the study was carried out
Sweden

Study type
Randomised double-blind controlled-trial

Aim of the study
To evaluated the effect of pertubation with Ringer-Lignocaine on dysmenorrhea in women with endometriosis

Study dates
22 March 2007 to 3 June 2009

Source of funding
An unconditional research grant from the Stockholm</td>
<td>Diastolic BP at inclusion, mean (SD)=74 (7.9)
Systolic BP at inclusion, mean (SD)=118 (13.0)
Caucasians=14
Oriental=3
Other=1
Patients using SSRI=4
Patients using analgesics=18
Patients using paracetamol=12
Patients using NSAIDs=13
Patients using codeine=6
Patients using tramadol=1
Patients using dextropropoxyphene=1
Patients using other opiods=2
Patients using oral contraceptive=3
Patients using intrauterine device=0
Patients using corpus luteum cyst=3
Patients using endometrioma=0

Lignocaine
Age, mean (SD)=33.08 (5.5)
Weight (kg), mean (SD)=69.5 (11.1)
Height (cm), mean (SD)=164.0 (4.6)
Duration of endometriosis (years), mean (SD)=5.62 (4.28)
Number of smokers=4
VAS at inclusion, mean (SD)=73.58 (19.0)
Diastolic BP at inclusion, mean (SD)=77 (9.8)
Systolic BP at inclusion, mean (SD)=121 (12.2)</td>
<td>to 12 in three sequential menstrual cycles.
4:3 treatment/placebo randomisation rate
Note: all patients used analgesics when needed</td>
<td>menstrual cycles and was considered successful if three treatnmebts were given during a maximum of five consecutive menstrual cycles. The pertubations were carried out on menstrual cycle Day 6 to 12. A thin plastic catheter (PBN-Medicals, Stenlose, Denmark) was inserted in the cervical canal and the small, intraluminal rubber balloon on the catheter was inflated with saline to prevent retrograde leakage. Blood pressure and heart rate were measured and recorded before and five minutes after the treatment. A 10ml quantity of solution was infused through the uterine cavity</td>
<td>
<table>
<tr><td>Social support</td><td>22</td><td>52.3 (22.6)</td><td>18</td><td>47.9 (20.8)</td></tr>
<tr><td>Self-image</td><td>22</td><td>34.1 (17.6)</td><td>18</td><td>25.5 (18.4)</td></tr>
<tr><td>Sexual intercourse</td><td>21</td><td>41.8 (27.3)</td><td>17</td><td>41.1 (24.1)</td></tr>
</table>
<u>Change after six months:</u>
<table>
<tr><td>EHP-30 dimension</td><td>n</td><td>Lidocaine, Median (IQR)</td><td>n</td><td>Placebo, Median (IQR)</td><td>p-value</td></tr>
<tr><td>Pain</td><td>20</td><td>-13.6 (-27.3 to -2.3)</td><td>15</td><td>-11.4 (-22.7 to -2.3)</td><td>0.99</td></tr>
<tr><td>Control and powerlessness</td><td>20</td><td>-8.3 (-33.3 to -2.1)</td><td>16</td><td>-6.3 (-35.4 to -2.1)</td><td>0.84</td></tr>
<tr><td>Emotional well-being</td><td>18</td><td>-4.2 (-37.5 to -4.17)</td><td>16</td><td>-12.5 (-20.8 to -6.25)</td><td>0.99</td></tr>
<tr><td>Social support</td><td>19</td><td>-18.8 (-31.25 to 0)</td><td>16</td><td>-6.3 (-12.5 to -6.25)</td><td>0.034</td></tr>
</table>
</td>
<td>pregnant, 2 endometriotic cysts and 1 escalting pain with need for other therapies (she did not fill in EHP-30 at baseline). 3 did not fill in the EHP-30 questionnaire at 12 months.
Placebo: after 6 months (n=2); 1 pregnant and 1 did not fill in EHP-30 at six months. After 12 months (n=8); 3 pregnant, 3 escalating pain with need for other therapies and 2 did not fill in EHP-30 questionnair</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td>County Council, Sweden</td>
<td>Caucasians=22
Oriental=0
Other=2
Patients using SSRI=3
Patients using analgesics=24
Patients using paracetamol=14
Patients using NSAIDs=22
Patients using codeine=5
Patients using tramadol=2
Patients using dextropropoxyphene=4
Patients using other opiods=3
Patients using oral contraceptive=2
Patients using intrauterine device=1
Patients using corpus luteum cyst=1
Patients using endometrioma=2

Inclusion criteria
• Presence of peritoneal or ovarian endometriosis as verified by laparoscopy and dysmenorrhea with a pain score of >50 mm on the visual analogue scale (VAS).
• Age >20 years; normal fallopian tubes; regular menstual cycles 21 to 35 days; treatment with oral contraceptive ongoing >1 month and continued during trial; previous hormonal treatment discontinued >1 month (OC, gestations) and >6 months (GnRH agonist); no wish for pregnancy during study; normal pap smear; negative chlamydia test; negative pregnancy test</td>
<td></td>
<td>and pertubated into the peritoneal cavity.
Quality of life was evaluated with the EHP-30 questionnaire, filled out at baseline, with follow-up after the 7th and 13th menstrual periods, i.e. 6 and 12 months after treatment. All dimensions and items on the questionnaire were collected. On the modular questionnaire, only the score concerning sexual intercourse (5 items) were included, since this is a frequent problem for women with endometriosis. If one or more items were missing from any dimension on the core and modular questionnaire, a scale score could not be calculated for that individual.</td>
<td>
<table>
<tr><td>Self-image</td><td>19</td><td>-8.3 (-16.7 to 0)</td><td>16</td><td>0.0 (-16.67 to -8.33)</td><td>0.24</td></tr>
<tr><td>Sexual interco urse</td><td>15</td><td>-10.0 (-25.0 to -10.0)</td><td>14</td><td>5.0 (-10 to -5)</td><td>0.24</td></tr>
</table>
<u>Change after 12 months:</u>
<table>
<tr><td>EHP-30 dimen sion</td><td>n</td><td>Lidocain e, Median (IQR)</td><td>n</td><td>Placeb o, Media n (IQR)</td><td>p-valu e</td></tr>
<tr><td>Pain</td><td>14</td><td>-8.0 (-29.5 to -2.3)</td><td>9</td><td>-11.4 (-20.5 to -4.5)</td><td>0.69</td></tr>
<tr><td>Contro l and powerl essnes s</td><td>13</td><td>-12.5 (-37.5 to -8.3)</td><td>10</td><td>-20.8 (-41.7 to -0)</td><td>0.74</td></tr>
<tr><td>Emotio nal well-being</td><td>12</td><td>-20.8 (-37.5 to -0)</td><td>10</td><td>-12.5 (-25.0 to -4.17)</td><td>0.63</td></tr>
<tr><td>Social suppor t</td><td>15</td><td>-12.5 (-37.5 to -0)</td><td>10</td><td>-6.3 (-31.25 to -12.5)</td><td>0.50</td></tr>
</table>
</td>
<td>e at 12 months.

Other information
This publicat ion is from the same study as Wickstom 2013, Pertu bation with lignocaine as a new treatment of dysmenorrh ea due to endometrios is: a randomised controlled trial, Human Reproductio n, Vol.27, No.3, 695-701</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td></td>
<td>Exclusion criteria
• Reduced patency in the Fallopian tubes and intention to achieve pregnancy during the forthcoming year.
• Continuous treatment with medication that may increase risk of infection; clinical signs of pelvic inflammmoatory disease; hyperreactivity to local anesthesia; fibroids >2 cm; ongoing treatment with GnRH agonist; ongoing continuous treatment with high-dose gestagens; pregnancy; peritubal adhesions; occluded fallopian tubes; inability to understanding information or comply with study procedures; Participation in a clinical study within one year before the present study; any disease or laboratory finding considered of importance by the investigator</td>
<td></td>
<td>If an item was misssin in any dimension at baseline then this specific score was withdrawn from further analysis.</td>
<td>
<table>
<tr><td>Self-image</td><td>15</td><td>-8.3 (-16.7 to 0)</td><td>10</td><td>0.0 (-16.7 to 0)</td><td>0.57</td></tr>
<tr><td>Sexual intercourse</td><td>12</td><td>-7.5 (-15.0 to -5)</td><td>8</td><td>-7.5 (-20.0 to -7.50)</td><td>0.97</td></tr>
</table>
</td>
<td></td>
</tr>
<tr>
<td>Full citation
Wickstrom, K., Bruse, C., Sjosten, A., Spira, J., Edelstam, G., Pertubation with lignocaine as a new treatment of dysmenorrhea due to endometriosis: A randomized</td>
<td>Sample size
Lignocaine, n=24; Placebo, n=18 (ITT)

Characteristics
Placebo
Age, mean (SD)=33.4 (4.4)
Weight (kg), mean (SD)= 67.6 (12.2)
Height (cm), mean (SD)=167.4 (8.6)
Duration of endometriosis (years), mean (SD)=4.25 (4.51)</td>
<td>Interventions
Study treatment: pertubation with lignocaine 1 mg/ml in Ringer solution
Placebo: pertubation with Ringer solution</td>
<td>Details
At the first visit baseline measurements were collected. At the second visit, patients were randomised sequentially in blocks of treatment (three placebo and four</td>
<td>Results
<u>Number of successful treatments in the PP population after three pertubations</u>
<u>Definition of success is improved >=50% on VAS scale from baseline)</u>
Lignocaine, n=9 (After 1st treatment, n=3; after second treatment, n=5; Success, first menstrual period after third treatment, n=9; 3rd menstrual period after third treatment, n=4; 6th menstrual period after third treatment, n=2; 9th menstrual period after third treatment, n=4)</td>
<td>Limitations
Five patients became pregnant and were withdrawn from further evaluation (lignocaine, n=2; placebo, n=3)</td>
</tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
controlled trial, Obstetrical & Gynecological Survey, 68, 286-7, 2013 **Ref Id** 405550 **Country/ies where the study was carried out** Sweden **Study type** Randomised double-blind controlled-trial **Aim of the study** To evaluated the effect of pertubation with Ringer-Lignocaine on dysmenorrhea in women with endometriosis **Study dates** 22 March 2007 to 3 June 2009 **Source of funding** An unconditional res	Number of smokers=0 VAS at inclusion, mean (SD)=78.22 (18.62) Diastolic BP at inclusion, mean (SD)=74 (7.9) Systolic BP at inclusion, mean (SD)=118 (13.0) Caucasians=14 Oriental=3 Other=1 Patients using SSRI=4 Patients using analgesics=18 Patients using paracetamol=12 Patients using NSAIDs=13 Patients using codeine=6 Patients using tramadol=1 Patients using dextropropoxyphene=1 Patients using other opiods=2 Patients using oral contraceptive=3 Patients using intrauterine device=0 Patients using corpus luteum cyst=3 Patients using endometrioma=0 **Lignocaine** Age, mean (SD)=33.08 (5.5) Weight (kg), mean (SD)=69.5 (11.1) Height (cm), mean (SD)=164.0 (4.6) Duration of endometriosis (years), mean (SD)=5.62 (4.28) Number of smokers=4 VAS at inclusion, mean (SD)=73.58 (19.0)	Three treatments given preovulatory on cycle day 6 to 12 in three sequential menstrual cycles. 4:3 treatment/placebo randomisation rate	study treatment). The treatment was given over three sequential menstual cycles and was considered successful if three treatnmebts were given during a maximum of five consecutive menstrual cycles. The pertubations were carried out on menstrual cycle Day 6 to 12. A thin plastic catheter (PBN-Medicals, Stenlose, Denmark) was inserted in the cervical canal and the small, intraluminal rubber balloon on the catheter was inflated with saline to prevent retrograde leakage. Blood pressure and heart rate were measured and recorded before and five minutes after the treatment. A 10ml	**Placebo,** n=1 (After 1st treatment, n=0; After second treatment, n=0; success, first menstrual period after third treatment, n=1; 3rd menstrual period after third treatment, n=1; 6th menstrual period after third treatment, n=0; 9th menstrual period after third treatment, n=0) <u>Definition of success is <20 mm on the VAS-scale</u> **Lignocaine** = after the third treatment, n=6 **Placebo** = after the third treatment, n=0	**Withdrawals** **Lignocaine:** n=2 had endometriosis >25 mm diagnosed 1 and 4 months after the third treatment; n=1 discontinued 5 days after third treatment because of such painful endometriosis that continuous OC had to be initiated **Placebo:** n=3 due to escalation pain and the need for other therapies such as high doses of gestagens or GnRH agonists **Other information**

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
earch grant from the Stockholm County Council, Sweden	Diastolic BP at inclusion, mean (SD)=77 (9.8) Systolic BP at inclusion, mean (SD)=121 (12.2) Caucasians=22 Oriental=0 Other=2 Patients using SSRI=3 Patients using analgesics=24 Patients using paracetamol=14 Patients using NSAIDs=22 Patients using codeine=5 Patients using tramadol=2 Patients using dextropropoxyphene=4 Patients using other opiods=3 Patients using oral contraceptive=2 Patients using intrauterine device=1 Patients using corpus luteum cyst=1 Patients using endometrioma=2 **Inclusion criteria** • Presence of peritoneal or ovarian endometriosis as verified by laparoscopy and dysmenorrhea with a pain score of >50 mm on the visual analogue scale (VAS). • Age >20 years; normal fallopian tubes; regular menstual cycles 21 to 35 days; treatment with oral contraceptive ongoing >1 month and continued during trial; previous hormonal treatment discontinued >1 month (OC, gestations) and >6 months (GnRH		quantity of solution was infused through the uterine cavity and pertubated into the peritoneal cavity. Dysmenorrhea was evaluated with a VAS scale and a pain questionnaire (revised version derived from Biberoglu and Behrman, 1981), initially filled out at the menstruation before the first treatment. thereafter the VASE scale and questionnaire were completed during the second, third and fourth period, i.e. after every treatment. the final follow-up took place after the 7th, 10th and 13th menstrual treatment, i.e. 6, 9 and 12 months after initial treatment. The maximum pain		This publication is from the same study as Wickstom 2013, Quality of life in patients with endometriosis and the effect of pertubation with lidocaine - a randomised controlled trial, Acta Obstetricia et Gynecologica Scandinavica, 92, 1375-1382.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	agonist); no wish for pregnancy during study; normal pap smear; negative chlamydia test; negative pregnancy test **Exclusion criteria** • Reduced patency in the Fallopian tubes and intention to achieve pregnancy during the forthcoming year. • Continuous treatment with medication that may increase risk of infection; clinical signs of pelvic inflammmoatory disease; hyperreactivity to local anesthesia; fibroids >2 cm; ongoing treatment with GnRH agonist; ongoing continuous treatment with high-dose gestagens; pregnancy; peritubal adhesions; occluded fallopian tubes; inability to understanding information or comply with study procedures; Participation in a clinical study within one year before the present study; any disease or laboratory finding considered of importance by the investigator		during every menstrual period was recorded and a decrease on the VAS scale of >=50% from baseline was defined as a success.		

G.15 Review question: Pharmacological management – Hormonal medical treatments

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?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation	**Sample size**	**Interventions**	**Details**	Results	Limitations

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Brown,J., Pan,A., Hart,R.J., Gonadotrophin-releasing hormone analogues for pain associated with endometriosis, Cochrane Database of Systematic Reviews, 12, CD008475-, 2010 **Ref Id** 112047 **Country/ies where the study was carried out** New Zealand, Australia **Study type:** Cochrane systematic review **Aim of the study:** To determine the effectiveness and safety of GnRHas in the treatment of the painful symptoms associated with endometriosis. **Study dates:**	N=41 RCTs examining GnRHas as treatment for pain associated with endometriosis versus no treatment, placebo, danazol, intra-uterine progestagens, or other GnRHas. **Characteristics** Randomised trials reporting the following comparisons were included: • GnRHas versus no treatment for relieving painful symptoms associated with endometriosis and its related adverse effects • GnRHas versus placebo for relieving painful symptoms associated with endometriosis and its related adverse effects • GnRHas versus analgesics for relieving painful symptoms associated with endometriosis and its related adverse effects • GnRHas versus danazol for relieving painful symptoms associated with endometriosis and its related adverse effects • GnRHas versus intra-uterine progestagen for relieving painful symptoms associated with endometriosis and its related adverse effects • Different doses of GnRHas for relieving painful symptoms associated with endometriosis and its related adverse effects	**Agarwal 1997**: Nafarelin 200mcg BD IN + placebo every 4 weeks IM for 6 months (n=105) vs LA Depot 3.75mg every 4 weeks IM + placebo BD IN for 6 months (n=103) **Bergqvist 1998**: Triptorelin 3.75mg IM depot every 4 weeks for 24 weeks (n=24) vs placebo IM every 4 weeks for 24 weeks (n=25) **Burry 1992**: Nafarelin 400mcg daily IN for 6 months (n=111) vs Danazol 600mg daily PO for 6 months (n=58) **Cheng 2005**: Nafarelin acetate 200mcg	**Agarwal 1997**: Multicentre, randomised, double-blind, double-placebo study **Bergqvist 1998**: Prospective, randomised, placebo-controlled, double-blind, parallel study, Sweden **Burry 1992**: Multi-centre, double-blind study, USA **Cheng 2005**: Randomised, parallel, comparative study, Taiwan **Fedele 1989**: Randomised study, Italy **Fedele 1993**: Multicentre, randomised controlled study, Italy. **Fraser 1991**: Double-blind, double-dummy, randomised,	**Agarwal 1997**: <u>Relief of painful symptoms at 6 months:</u> **Pelvic tenderness:** • GnRHa (nafarelin) = 53/99 • GnRHa (LA depot) = 58/93 • RR=0.86 (0.67 to 1.09) **Pelvic induration:** • GnRHa (nafarelin) = 73/99 • GnRHa (LA depot) = 74/91 • RR=0.91 (0.78 to 1.06) **Bergqvist 1998**: <u>Relief of pelvic tenderness</u> GnRHa n=24 • Placebo group n=25 • RR 4.17 (95% CI 1.62 to 10.68, P=0.003) **Burry 1992**: <u>Quality of life</u> No data given, only reported that there were no between-group differences, however the nafarelin group showed significant (p<0.05, paired t-test) improvement from baseline in work productivity at all assessments, whereas there was no significant change in this measure in the danazol group. **Cheng 2005**:	**Agarwal 1997**: Adequate sequence generation? Low risk Allocation concealment? Unclear risk (No details) Blinding? Low risk Incomplete outcome data addressed? Low risk Free of selective reporting? Low risk **Bergqvist 1998**: Adequate sequence generation? Unclear risk Allocation concealment? Unclear risk Blinding? Low risk Incomplete outcome data addressed? Low risk Free of selective reporting? Low risk **Burry 1992**: Adequate sequence

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
2010 **Source of funding:** Internal sources Uiniversity of Auckland, New Zealand. Lead author AP (who is an undergraduate medical student) has been funded to complete the review. External sources No sources of support supplied	• Different treatment length of GnRHas for relieving painful symptoms associated with endometriosis and its related adverse effects • Different route of administration of GnRHas for relieving painful symptoms associated with endometriosis and its related adverse effects • Different GnRHas treatment regimes for relieving painful symptoms associated with endometriosis and its related adverse effects **Inclusion Criteria** **Agarwal 1997**: • 208 women were randomised, 192 were analysed • Laparoscopically diagnosed endometriosis within 18 months prior to study19-44 years old • Patients demonstrating clinical symptoms and signs • Bone mineral density within normal age range **Bergqvist 1998**: 49 women eligible; 49 were randomised and 46 were analysed; Age: mean of 31 years (19-44years); stage: most mild to moderate (IV n=1)	BD (400mcg/day) IN for 180 days (n=29) vs Danazol 200mg TID (600mg/day) PO for 180 days (n=30) **Fedele 1989**: Buserelin 400mcg TDS IN for 6 months (n=30) vs Danazol 200mg TDS PO for 6 months (n=32) **Fedele 1993**: Buserelin acetate 1200mcg daily IN for 6 months (n=19) vs expectant management (n=16) **Fraser 1991**: Nafarelin 200mcg BDS (400mcg/d) IN + placebo PO for 6 months (n=33) vs Danazol	parallel study, Australia/New Zealand **NEET 1992**: Multicentre, parallel, randomised, double-blind, double-dummy study **Petta 2005**: Randomised controlled trial, Brazillien **Wheeler 1992**: Double-blind, multi-centre, randomised trial	<u>Pelvic tenderness at 3 months</u> MD = -0.2 (-0.69 to 0.29)* <u>Pelvic tenderness at 6 months</u> MD = -0.2 (-0.66 to 0.26)* <u>Pelvic induration at 3 months</u> MD = -0.1 (-0.51 to 0.31)* <u>Pelvic induration at 6 months</u> MD = 0.2 (-0.21 to 0.61)* **Fedele 1989**: <u>Patients requiring surgery because of reappearance of symptoms and positive findings at pelvic examination at 6 months</u> • GnRHa = 4/11 • Danazol = 5/14 • RR = 1.02 (0.36 to 2.91)* **Fedele 1993**: <u>Relief of the pain of dysmenorrhoea associated with endometriosis</u> • GnRHa group n=19 • Expectant management group n=16 • RR 3.93 (95% CI 1.37 to 11.28, P=0.01). **Fraser 1991**: <u>Pelvic tenderness at 6 months</u> MD = -0.1 (-0.38 to 0.18) <u>Pelvic induration at 6 months</u> MD = 0.0 (-0.28 to 0.28)	generation? Unclear risk Allocation concealment? Unclear risk Blinding? Unclear risk Incomplete outcome data addressed? Low risk Free of selective reporting? Low risk **Fedele 1993**: Adequate sequence generation? Unclear risk Allocation concealment? Unclear risk Blinding? High risk Incomplete outcome data addressed? Low risk Free of selective reporting? Low risk **Fraser 1991**: Adequate sequence generation? Low risk

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	• The study population included women who were: • Menstruating regularly 3 months before study • Clinical symptoms of endometriosis • Not taken oral contraceptive or oral steroid therapy for 3 months • Not taken long acting depot gestagens or GnRHas within past 6 months • Not pregnant in prior 3 months • Not breastfeeding • No history of osteoporosis or coagulation disorders **Burry 1992**: • 169 women eligible; 169 were randomised and 147 analysed for efficacy • The study population included women who had laparoscopically diagnosed endometriosis **Cheng 2005**: • 59 women eligible; 59 were randomised and 41 were analysed for efficacy • Laparoscopically diagnosed within 3 months prior to study • Age 18-48 years • Barrier contraception	200mg TDS (600mg/d) PO + placebo IN for 6 months (n=16) **NEET 1992**: Nafarelin 200mcg BD IN + placebo PO for 6 months (n=206) vs Danazol 200mg TDS PO + placebo IN for 6 months (n=101) **Petta 2005**: LNG-IUS (Mirena) 20mcg/day 5 years IU for 6 months (n=40) vs Lupron 3.75mg every 28 days IM for 6 months (n=43) **Wheeler 1992**: Leuprolide 3.75mg monthly IM + placebo OD PO for 24 weeks (n=134) vs Danazol 800mg OD PO + placebo monthly		Pregnancies (infertile patients conceived within 12 months of completion of therapy • GnRHa (nafarelin) = 12/22 • Danazol = 6/14 • RR = 1.27 (0.62 to 2.60)* **NEET 1992**: Relief of painful symptoms at 6 months: **Pelvic tenderness** • GnRHa (nafarelin) = 50/65 • Danazol = 23/31 • RR=1.04 (0.81 to 1.33) **Pelvic induration** • GnRHa (nafarelin) = 59/65 • Danazol = 27/31 • RR=1.04 (0.89 to 1.22) **Petta 2005**: QoL (Psychological Well-Being index Questionnaire) at 6 months MD = -1.2 (-7.79 to 5.39)* **Wheeler 1992**: Pelvic tenderness • GnRHa=93/128 • Placebo=95/125 • RR=0.96 (0.83 to 1.11) *calculated by the 2016 NGA team	Allocation concealment? Unclear risk (No details) Blinding? Low risk Incomplete outcome data addressed? Unclear risk (No details on attrition) Free of selective reporting? Low risk **NEET 1992**: Adequate sequence generation? Unclear risk ("patients were randomised so that 2 were assigned to receive nafarelin for every 1 assigned to receive danazol") Allocation concealment? Unclear risk (No details) Blinding? Low risk Incomplete outcome data addressed? Low risk Free of selective reporting? Low risk

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Fedele 1989: • 62 women were randomised and analysed: • Laparoscopically diagnosed endometriosis within 3 months prior to study • No therapeutic intervention • stage: I or II • The study population included women who were: • Laparoscopically diagnosed endometriosis • One or more of dysmenorrhoea, pelvic pain and deep dyspareunia **Fraser 1991**: • 49 women were randomised and 45 were analysed, stage: I to III • Laparoscopically diagnosed endometriosis • Symptomatic • Regular menstrual cycle 24-36 days • Not pregnant • Negative pap smear • Barrier contraception **NEET 1992**: • 315 women were randomised, 307 were analysed for safety and 263 were analysed for efficacy	IM for 24 weeks (n=136)			**Wheeler 1992**: Adequate sequence generation? Unclear risk (No details) Allocation concealment? Unclear risk (No details) Blinding? Low risk Incomplete outcome data addressed? Low risk Free of selective reporting? Low risk Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	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 45-110 kg **Petta 2005:** • 83 women were randomised, 71 were analysed, stage: I to IV 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 pretreatment 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				

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	• Not taken any long acting progestins or GnRHa within 9 months prior to study • Not pregnant or breastfeeding 3 months prior to study • No osteoporosis, coagulation disorders or contra-indications **Wheeler 1992**: 270 women were randomised and 253 were analysed. Age: Leuprolide = 31.0 and Danazol = 29.8 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 GnRHa • Any other treatment completed at least 3 months prior to study **Exclusion Criteria** **Agarwal 1997**: • Conditions or drug therapies that may interfere with the study • Pregnant or lactating women				

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	• Danazol use within 6 months prior to study • GnRHa use within 12 months prior to study • OCP within 30 days prior to study treatment • Thyroid disease **Bergqvist 1998**: • Intraperitoneal adhesions making visual inspection and careful evaluation of the extension of endometriotic lesions difficult or impossible **Burry 1992**: not reported **Cheng 2005**: • Pregnancy • Breastfeeding • Menopause or post-menopausal • Use of oestrogen, progesterone or contraceptive steroids in previous 3 months • Impaired hepatic or renal function • Cardiovascular disease • AIDS or other sexually transmitted diseases **Fedele 1989**:				

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	• Bilateral tube occlusion or partner with severe dyspermia • Danazol or other sex hormone use within 6 months prior to study • Systemic or endocrine disease **Fedele 1993:** not reported **Fraser 1991**: • Concurrent disease which may interfere with drug • Surgical therapy within 6 months prior to study entry • Steroid therapy within 3 months prior to study entry **NEET 1992**: • Amenorrhoea • Concurrent disease which may interfere with endometriosis or contraindicate the use of androgenic therapy • Surgical treatment at baseline or within 6 months prior to study • Use of danazol, androgenic hormones, eostrogens, or progestogens within 3 months prior to study **Wheeler 1992**:				

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td>not reported</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Full citation
Brown, J., Kives, S., Akhtar, M., Progestagens and anti-progestagens for pain associated with endometriosis, Cochrane Database of Systematic Reviews, 3, CD002122, 2012
Ref Id
346707

Country/ies where the study was carried out
New Zealand, Canada, UK

Study type:
Cochrane systematic review

Aim of the study:
To determine the effectiveness and adverse effects of both progestagens and anti-progestagens in the treatment of painful symptoms</td><td>Sample size:
A total of 13 studies included in this 2011 Cochrane Review update. There were seven studies in the last published version from 2000.
The six newly included studies evaluated progestagens (comparisons with placebo, danazol, oral or subdermal contraceptive, oral contraceptive pill and danazol, gonadotrophin-releasing hormone (GnRH) analogue and other drugs). The remaining studies compared the anti-progestagen gestrinone with danazol, GnRH analogues or itself.

Characteristics
Only RCTs were included:
Bergvist 2001
Vercellini 1996

Inclusion Criteria
Bergvist 2001:
• 48 Swedish women 18-46 years.
• diagnosis of endometriosis by laparoscopy or laparotomy within 3 months regular menstruating and complaining of dysmenorrhoea,</td><td>Interventions
Bergvist 2001:
1. Nafarelin 200 µg intranasally (IN) BID and 'dummy' medroxyprogesterone tablets (23 women)
2. Medroxyprogesterone 15 mg PO BID and 'dummy' nafarelin nasal spray (25 women)
Duration of treatment: 6 months

Vercellini 1996:
1. Depot medroxyprogesterone acetate 150 mg every 90 days
2. Oral contraceptive pill (ethinyl estradiol 0.02 mg + desogestrel 0.15mg) plus 50 mg danazol daily for 21 days out of 28
Duration of</td><td>Details
Bergvist 2001:
Randomised single centre, double dummy parallel study.

Vercellini 1996:
Open randomised trial</td><td>Results
Bergvist 2001:
<u>Quality of life</u>
Means of scores for anxiety-depression, according to the short version of the General Health Questionnaire of Goldberg and disturbed sleep, according to Åkerstedt, for the nafarelin (n=17) and MPA (n=13) treated groups. Analysis of variance (ANOVA) for repeated measures (mixed model)
<table>
<tr><td></td><td>Before</td><td>6 months</td><td>12 months</td></tr>
<tr><td colspan="4">Disturbed sleep</td></tr>
<tr><td>Nafarelin</td><td>2.53</td><td>2.24</td><td>1.47</td></tr>
<tr><td>MPA</td><td>2.92</td><td>1.39</td><td>1.85</td></tr>
<tr><td colspan="4">F group=0.0003, p=0.95</td></tr>
<tr><td colspan="4">F time=4.32, p=0.02</td></tr>
<tr><td colspan="4">F interaction=1.72, p=0.19</td></tr>
<tr><td colspan="4">Anxiety-depression</td></tr>
<tr><td>Nafarelin</td><td>63.9</td><td>70.1</td><td>60.1</td></tr>
<tr><td>MPA</td><td>65.8</td><td>63.2</td><td>54.8</td></tr>
<tr><td colspan="4">F group=0.63, p=0.43</td></tr>
<tr><td colspan="4">F time=7.12, p=0.002</td></tr>
<tr><td colspan="4">F interaction=1.64, p=0.20</td></tr>
</table></td><td>Limitations
Bergvist 2001:
Random sequence generation (selection bias): Unclear risk (Method of randomisation not described)
Allocation concealment (selection bias): Unclear risk (No details)
Blinding (performance bias and detection bias): Unclear risk (Double dummy, no details and no details of blinding)
Incomplete outcome data (attrition bias): Low risk
Selective reporting (reporting bias): High risk (Main outcomes described, no details of side effects)
Selective reporting (reporting bias): Unclear risk (A priori outcomes</td></tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td>associated with endometriosis.

Study dates:
2011

Source of funding:
Internal sources
University of Cambridge, UK.
External sources
The Cambridge University Hospital's NHS Trust, UK.</td>
<td>dyspareunia and/or pelvic pain.

Vercellini 1996:
• first diagnosis of endometriosis at laparoscopy with attempt at implant reduction other than biopsy in the previous 3 months, pelvic pain of greater than 6 months duration.

Exclusion Criteria
Bergvist 2001:
• extensive adhesions,
• pelvic pain for other reasons
• no surgery within the last 12 months with the exception of removal of an endometrioma
• no use of laser or diathermy, steroid medication within 3 months or 1 month of diagnostic laparoscopy, previous use of any GnRH agonists, pregnant, breastfeeding or hysterectomy within 6 months prior to inclusion, use of concomitant contraceptive steroids, androgenic hormones, estrogens, progestagens, danazol,GnRh analogs, anxiolytics, cortizone and hypnotics,women with other concurrent disease either oncologic or psychiatric.</td>
<td>treatment: 12 months</td>
<td></td>
<td>Mean ranks for the different examinations and non-parametric variance tests (Friedman) for the nafarelin (n=16) and the MPA (n=13) treated groups concerning results from the Nottingham Health Profile (NHP) tests. Answers from one nafarelin treated patient are missing
<table>
<tr><td></td><td>Before</td><td>6 months</td><td>12 months</td><td>p</td></tr>
<tr><td colspan="5">Paid working li</td></tr>
<tr><td>Nafarelin</td><td>2</td><td>1.9</td><td>1.7</td><td>0.04</td></tr>
<tr><td>MPA</td><td>2.1</td><td>2</td><td>1.9</td><td>0.69</td></tr>
<tr><td>Total</td><td></td><td></td><td></td><td>0.06</td></tr>
<tr><td colspan="5">Household wor</td></tr>
<tr><td>Nafarelin</td><td>2.3</td><td>2</td><td>1.8</td><td>0.09</td></tr>
<tr><td>MPA</td><td>2.2</td><td>1.9</td><td>1.9</td><td>0.32</td></tr>
<tr><td>Total</td><td></td><td></td><td></td><td>0.04</td></tr>
</table>
Means of psychological and psychosocial variables according to the Nottingham Health Profile (NHP) for the nafarelin (n=16) and MPA (n=13) treated groups. Answers</td>
<td>reported but original protocol not sighted)

Vercellini 1996:
Random sequence generation (selection bias): Low risk
Allocation concealment (selection bias): Low
Blinding (performance bias and detection bias): High risk ('open label', subjects not blinded)
Incomplete outcome data (attrition bias): Unclear risk (4 MDPA withdrew (3 for prolonged bleeding and 1 for persistent pain); seven in the oral contraceptive pill (OCP) + danazol (3 for persistent pain, two for bloating and weight gain, 2 for personal reasons))</td>
</tr>
</table>

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td></td>
<td>Vercellini 1996:
Treatment for endometriosis other than non-steroidal anti-inflammatory drugs in preceding 3 months, contraindications to taking estrogens, progestagens or danazol, a desire to conceive in the next 2 years.</td>
<td></td>
<td></td>
<td>from one nafarelin treated patient are missing. Analysis of variance (ANOVA) for repeated measures (mixed model)
<table>
<tr><td></td><td>Before</td><td>6 months</td><td>12 months</td></tr>
<tr><td colspan="4">Vacation life</td></tr>
<tr><td>Nafarelin</td><td>0.38</td><td>0.19</td><td>0.19</td></tr>
<tr><td>MPA</td><td>0.31</td><td>0.15</td><td>0</td></tr>
<tr><td colspan="4">F group=0.99, p=0.33</td></tr>
<tr><td colspan="4">F time=3.15, p=0.05</td></tr>
<tr><td colspan="4">F interaction=0.33, p=0.72</td></tr>
<tr><td colspan="4">Leisure</td></tr>
<tr><td>Nafarelin</td><td>0.56</td><td>0.25</td><td>0.25</td></tr>
<tr><td>MPA</td><td>0.46</td><td>0.15</td><td>0.23</td></tr>
<tr><td colspan="4">F group=0.55, p=0.47</td></tr>
<tr><td colspan="4">F time=3.90, p=0.03</td></tr>
<tr><td colspan="4">F interaction=0.07, p=0.93</td></tr>
<tr><td colspan="4">Sexual life</td></tr>
<tr><td>Nafarelin</td><td>0.53</td><td>0.4</td><td>0.2</td></tr>
<tr><td>MPA</td><td>0.69</td><td>0.62</td><td>0.46</td></tr>
<tr><td colspan="4">F group=2.44, p=0.13</td></tr>
<tr><td colspan="4">F time=3.45, p=0.04</td></tr>
<tr><td colspan="4">F interaction=0.11, p=0.90</td></tr>
</table>
Vercellini 1996:
<u>MD in pain</u>:
At 6 months during treatment:</td>
<td>Selective reporting (reporting bias): Unclear risk (A priori outcomes reported but original protocol not sighted)</td>
</tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				Dysmenorrhea: MD=-1.8 (-2.23 to -1.45)* Dyspareunia: MD=-0.3 (-1.18 to 0.58)* Non menstrual pain: MD=0.6 (-0.09 to 1.29)* At the end of treatment (12 months): Dysmenorrhea: MD=-1.3 (-1.79 to -0.81)* Dyspareunia: MD=-0.3 (-1.41 to 0.81)* Non menstrual pain: MD=0.4 (-0.42 to 1.22)* * calculated by the 2016 NGA team Patient satisfaction with treatment (very satisfied/satisfied) at the end of the 12 month treatment period: • very satisfied/satisfied: 72.5% (n=29) in the medroxyprogesterone group • very satisfied/satisfied: 57.5 % (n=23) in the OCP + desogetrel group • OR=1.95 (0.76 to 4.97) [RR=1.26 (0.91 to 1.75)]	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				Other results: 2.5% very satisfied in the medroxyprogesterone group 70% satisfied in the medroxyprogesterone group 5% uncertain in the medroxyprogesterone group 20% dissatisfied in the medroxyprogesterone group 2.5% very dissatisfied in the medroxyprogesterone group 15% very satisfied in the OCP + desogetrel group 42.5% satisfied in the OCP + desogetrel group 10% uncertain in the OCP + desogetrel group 30% dissatisfied in the OCP + desogetrel group 2.5% very dissatisfied in the OCP + desogestrel group	
Full citation Davis, L., Kennedy, S. S., Moore, J., Prentice, A., Modern combined oral contraceptives for pain associated with endometriosis, Cochrane Database of Systematic	**Sample size:** **Vercellini 1993** N=57, stages I-IV n=29 in the goserelin group n=28 in the OC group **Characteristics** Women with laparoscopically diagnosed endometriosis and at least one moderate or severe pain symptom as judged by a verbal rating scale and a visual analogue scale. Included in the analysis:	**Interventions** Goserelin 3.6 mg subcutaneous depot formulation monthly for 6 months or cyclic low dose monophasic contraceptive pill, containing 0.02 mg ethinyl estradiol and 0.15 mg desogestrel	**Details** A randomisation list was used to allocate patients to a 6-month treatment with goserelin, 3.6 mg in a 28-day subcutaneous depot formulation or a cyclic low-dose monophasic OC containing ethinyl E2 (EE2), 0.02 mg and desogestrel 0.15 mg per pill. In the OC group, if spotting or breakthrough bleeding	**Results** **Pain at the end of treatment (6 months):** Dysmenorrhea: not reported Dyspareunia: MD -1.8 (-3.4 to -0.2) Non menstrual pain: MD 0.2 (-1.11 to 1.51) **Pain at 6 month after treatment:** Dysmenorrhea: MD 0.10 (-1.08 to 1.28)	**Limitations** Adequate sequence generation? Unclear risk (No details) Allocation concealment? Un clear risk (No details) Blinding? High risk ()No blinding of participants, investigators or assessors reported

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Reviews, CD001019, 2007 **Ref Id** 346744 **Country/ies where the study was carried out** New Zealand **Study type:** Cochrane Systematic Review **Aim of the study:** To assess the effects of the oral contraceptive pill (OCP) in comparison to other treatments for painful symptoms of endometriosis in women of reproductive age. **Study dates:** 2012 **Source of funding:** Internal sources	n=26 in the goserelin group n=24 in the OC group **Inclusion Criteria** • Women who had had a diagnostic laparoscopy with no attempts at endometriosis reduction other than biopsy within 3 months of study entry. **Exclusion Criteria** • Women who had received any treatment for endometriosis other than nonsteroidal anti-inflammatory drugs in the preceding 3 months • Women with the usual contraindications to OCs.	(dose increased to 0.03 mg ethinyl estradiol if spotting occurred)	occurred, patients could switch to a contraceptive with EE2, 0.03 mg and desogestrel 0.15 mg per pill.	Dyspareunia: MD -0.40 (-2.10 to 1.30) Non menstrual pain: MD 0.30 (-1.25 to 1.85)	Incomplete outcome data addressed? Low risk Free of selective reporting? Low risk

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
AP University of Cambridge, UK, Not specified. JM and SK University of Oxford, UK, UK. External sources LJD Peninsula Medical School Foundation Bursary, UK. LJD National Birthday Trust Fund, Wellbeing of Women, UK.					
Full citation Harada, T., Momoeda, M., Taketani, Y., Hoshiai, H., Terakawa, N., Low-dose oral contraceptive pill for dysmenorrhea associated with endometriosis: a placebo-controlled, double-blind, randomized trial, Fertility & Sterility, 90, 1583-8, 2008 **Ref Id** 338458	**Sample size:** Of 107 patients entered in the study, 7 were excluded before randomization because they had abnormal smear cytology (n = 3), Exclusion Criteria (n = 3), or positive antiphospholipid antibodies (n = 1). 100 patients were randomized to receive either OCP (n = 51) or placebo (n = 49). 1 patient in the OCP group did not take OCPs because she became pregnant after randomization. 1 patient in the OCP and two in the placebo group were lost to follow-up. n= 96 patients were included in at least one of the efficacy analyses.	**Interventions** Monophasic oral contraceptive pill (OCP) (ethinyl-estradiol 0.035mg plus norethisterone 1mg) for 21 days plus 7 days of placebo for 3 cycles (n=49) vs placebo for 28 days for 3 cycles (n=47).	**Details** This was a phase III, randomized, double-blind, placebocontrolled, multicenter trial of low-dose OCP versus placebo in 100 patients with endometriosis performed in 18 centers (13 clinics, 5 hospitals) in Japan. Subjects were randomly assigned in a ratio of 1:1 to receive monophasic OCP (ethinylestradiol 0.035 mg plus norethisterone 1 mg) for 21 days, plus 7 days of placebo or identical placebo for 28 days. The OCP and the	**Results** **Mean pain (VAS) at pre-treatment and at the end of treatment:** Dysmenorrhea: • Oral contraceptive group at pre-treatment =58.7 SD 18.6, at the end of treatment =27.6 SD 21.6, n=49 • Placebo group at pre-treatment =55.8 SD 17.5, at the end of treatment =46.2 SD 24.2, n=47 • Mean difference =-21.5 (95%CI -28.14 to -14.86)* Non-menstrual pelvic pain: • Oral contraceptive group at pre-treatment =27.5 SD 25.1, at the end of treatment =19.1 SD 22.9, n=49	**Limitations** Risk of bias (Cochrane Risk of Bias tool) Sequence generation: Low risk Allocation concealment: Low risk Blinding: Low risk Incomplete data: Low risk Selective reporting: Unclear risk Other: None **Other information** None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Country/ies where the study was carried out Japan **Study type:** A placebo-controlled, double-blind, randomized trial. **Aim of the study:** To evaluate the efficacy of a low-dose oral contraceptive pill (OCP) for patients with dysmenorrhea associated with endometriosis. **Study dates:** Not reported. **Source of funding:** All authors have received consulting fee from Nobelpharma Co., Ltd. Tokyo, Japan.	**Characteristics** Most patients (47 of 49 in the OCP group and 44 of 47 patients in the placebo group) had endometrioma. N=14 patients (seven OCP, seven placebo) discontinued the study. 4 of the OCP patients were discontinued because of adverse effects (one, rupture of ovarian cyst; one, nausea and headache; one, ovarian hemorrhagic cyst; one, edema), 2 patients were lost to follow-up, and 1 took a prohibited drug. 7 of the placebo patients terminated: 3 had adverse effects (one, edema and headache; one, ovarian hemorrhagic cyst; one, worsened dysmenorrhea), 3 were lost to follow-up, and 1 used a prohibited drug. Continuation rates were similar between the treatment groups, with 88% of patients receiving OCPs and 86% receiving placebo continuing in the study. **Inclusion Criteria** • women of 18 years and older; regular menstrual cycles; symptomatic endometriosis (diagnosed by laparoscopy orlaparotomy) or ovarian		placebo were prepared by the manufacturer in 28-day blister packs and appeared identical. The use of analgesic agents was allowed, but other hormonal treatments for pain or vaginal bleeding were prohibited. Randomization was done by the pharmaceutical company (Nobelpharma Co., Ltd. Tokyo, Japan), using the permuted block method. Allocation concealment was accomplished centrally by the company, not broken until after all data were collected. Both the patients and the doctors were blinded regarding the medication. Treatment began on the third day (2 days) of the menstrual cycle and continued for four cycles.	• Placebo group at pre-treatment =22.8 SD 24.5, at the end of treatment =21.0 SD 26.0, n=47 • Mean difference =-6.60 (95%CI -14.27 to 1.07)* <u>Induration identified:</u> • Oral contraceptive group at pre-treatment =32/49, at the end of treatment =21/49 • Placebo group at pre-treatment =33/47, at the end of treatment =14/47 • RR = 0.56 (95% CI 0.30 to 1.04)* *calculated by the 2016 NGA team	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	endometrioma (diagnosed by ultrasound or magnetic resonance imaging); normal cervical and endometrial smear cytology; moderate or severe dysmenorrhea (evaluated by a modified pain scale) and no medical or surgical treatment for endometriosis within 8 weeks before entry into the study, including hormonal agents, such as OCP, GnRHa, and danazol. • The study patients must have had moderate or severe dysmenorrhea, scoring higher than three points at the admission visit on a modified pain scale originally developed by Biberoglu et al. and Andersch et al. **Exclusion Criteria** Not reported.				
Full citation Hughes,E., Brown,J., Collins,J.J., Farquhar,C., Fedorkow,D.M., Vandekerckhove, P., Ovulation suppression for endometriosis, Cochrane Database of	**Sample size:** N=25 studies **Characteristics** All published, unpublished, and ongoing randomised controlled trials (RCTs) were included if they made the following comparisons for the treatment of endometriosis-associated subfertility. 1) An ovulation suppression	**Interventions** **Burry 1989** Danazol 800 mg daily (n=10) PO + placebo vs danazol 600 mg daily (n=8) PO + placebo vs nafarelin 800 µg	**Details** **Burry 1989** All patients were examined before the start of treatment and after 2, 4 and 6 months of therapy. A second laparoscopy was was performed during the last month of drug therapy for restaging of endometriosis.	**Results** **Burry 1989** <u>Clinical pregnancies for women randomised:</u> • GnRHa (nafaerlin)=15/35 • Danazol=2/18 • RR=3.86 (0.99 to 15.052) <u>Clinical pregnancies in infertile couples/those desiring pregnancy only:</u> • GnRHa (nafaerlin)=15/30	**Limitations** **Burry 1989** Adequate sequence generation? unclear risk (No details) Allocation concealment? Unclear risk (No details) Blinding? Low risk

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Systematic Reviews, #2007. **Ref Id** 68470 **Country/ies where the study was carried out** New Zealand **Study type:** Cochrane Systematic Review **Aim of the study:** To assess the effectiveness of ovulation suppression agents, including danazol, progestins and oral contraceptives, in the treatment of endometriosis-associated subfertility in improving pregnancy outcomes including live births. **Study dates:**	agent with placebo or no treatment. 2) Danazol with another ovulatory suppressive agent; where danazol was prospectively singled out for comparison with other agents because it has been considered the primary choice for medical suppression before the advent of gonadotropin-releasing hormone analogues (GnRHa). If newer agents were more effective than danazol, this comparison would demonstrate the extent of the improvement. 3) GnRH versus oral contraception. Quasi-randomised trials were excluded. If crossover design was used, only the first phase or stage would be extracted for analysis. <u>Types of participants</u> Women with visually diagnosed endometriosis, either by laparoscopy or laparotomy, who had failed to conceive after 12 or more months of unprotected intercourse. Trials where medical treatment was administered after surgical treatment for endometriosis were included. <u>Types of interventions</u> Interventions included danazol, medroxyprogesterone acetate (MPA), gestrinone, combined	daily (n=10) IN + placebo vs nafarelin 400 µg daily (n=25) IN + placebo.		• Danazol=2/14 • RR=3.50 (0.92 to 13.26)	Incomplete outcome data addressed? Low risk Free of selective reporting? high risk (Not followed up to live birth

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
2007 **Source of funding:** Internal sources No sources of support supplied External sources Royal Commission on New Reproductive Technologies, Not specified.	oral contraceptive pills (COC), GnRH analogues (GnRHa), and placebo. No dose ranges were specified. **Inclusion Criteria** **Burry 1989** Women complained of infertility, pain or both. **Exclusion Criteria** **Burry 1989** Women who received medical therapy for endometriosis within preceding 6 months.				
Full citation 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 **Ref Id** 338495 **Country/ies where the study was carried out** USA	**Sample size:** Of the 100 women who were randomized to treatment, 49 of 50 in the depot leuprolide group and 46 of 50 in the placebo group completed the study. **Characteristics** The mean age of women in the depot leuprolide group (32.3 years) was greater than that of women in the placebo group (29.4 years); this difference was statistically but not clinically significant (P 5 .036). Most patients were white (76%); others were black (17%) or Hispanic (7%). There were no clinically significant differences between treatment groups in laboratory test results, vital	**Interventions** Leuprolide acetate 3.75mg IM depot every 4 weeks on day 0, week 4 and week 8 (n=49) vs Placebo IM every 4 weeks on day 0, week 4 and week 8 (n=46).	**Details** Eligible women were assigned subject numbers in sequential order at each site and randomized to treatment with depot leuprolide (Lupron Depot 3.75 mg; TAP Pharmaceuticals, Deerfield, IL) or placebo, usually beginning treatment between days 1 and 4 of the menstrual cycle. The randomization schedules were prepared in random blocks of two and four, with treatment group assignment in a 1:1 ratio. Each group was	**Results** **Mean pain (VAS) at baseline and week 12:** Dysmenorrhea: • Depot leuprolide group at baseline =7.5, at week 12 =0.1, n=44 • Placebo group at baseline =8.0, at week 12 =6.4, n=44 • Mean difference =-6.3 (95%CI -9.93 to -2.67)* Pelvic pain: • Depot leuprolide group at baseline =7.7, at week 12 =2.2, n=44 • Placebo group at baseline =6.4, at week 12 =6.6, n=44 • Mean difference =-3.1 (95%CI -4.85 to -1.35)* Dyspareunia:	**Limitations** Cochrane risk of bias assessment tool Adequate sequence generation? Low risk (block randomization) Allocation concealment? Low risk (randomization schedule) Blinding? Unclear risk (no details given) Incomplete outcome data addressed? Low risk (details for attrition given)

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study type: Double-blind, randomized, parallel-group, placebo-controlled trial. **Aim of the study:** To assess and compare the safety and efficacy of depot leuprolide versus placebo in management of chronic pelvic pain in women with clinically suspected endometriosis. **Study dates:** The trial was conducted at 12 sites in the US between June 1995 and January 1997. **Source of funding:** This study was supported by a grant from TAP Holdings, Inc., which distributes depot leuprolide.	signs, or physical examination results at baseline. **Inclusion Criteria** • Women 18–45 years of age were eligible for enrollment if they had had moderate to severe chronic pelvic pain for at least 6 months, with severity being assessed by a physician using the four-point Biberoglu and Behrman scale (1 = none, 2 = mild, 3 = moderate, and 4 = severe), and that pain was unrelated to menstruation and incompletely relieved with nonsteroidal antiinflammatory drugs. Eligible patients also had to have had regular menstrual bleeding and menstrual cycles for 3 months before enrollment. **Exclusion Criteria** • Women were excluded if they had a previous diagnosis of endometriosis confirmed by laparoscopy, laparotomy, or histology; had received oral contraceptives (OCs) within the previous 3 months or GnRH agonists within the previous 6 months; or had undergone surgical treatment for endometriosis. Women whose chronic pelvic pain might be related to		represented once within each block of two and twice within each block of four. The schedules were prepared by an administrative staff member using a FORTRAN program to generate uniform random numbers. Study medication was packaged according to the randomization schedules and was sent to each site in sets of four, as needed. Patient numbers were sequential within each set. Patient number assignment started with the lowest available number for each site and proceeded in ascending order. Both depot leuprolide and placebo were administered IM three times at 4-week intervals: on day 0, during week 4, and during week 8. To preserve the double blind, active treatment and placebo intramuscular injections were prepared identically by mixing the formulation with a	• Depot leuprolide group at baseline =5.1, at week 12 =2.1, n=31 • Placebo group at baseline =5.2, at week 12 =5.1, n=30 • Mean difference =-4.4 (95%CI -4.40 to -1.87)* *calculated by the 2016 NGA team	Free of selective reporting? Low risk (All primary outcomes reported) Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	genitourinary disease or to chronic or recurrent gastrointestinal disease, including irritable bowel syndrome (defined as a disease characterized by pain relieved by defecation and irregular defecation patterns lasting at least 3 months), also were excluded, as were those with histories of alcohol use or other chronic tranquilizer or illicit drug use. Women who had not been sterilized surgically agreed to use barrier contraception during treatment and for 6 weeks thereafter.		diluent from a separate ampule.		
Full citation 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.	**Sample size:** N=102 n=47 in the gestodene 0.75 mg / ethinylestradiol 0.03 mg group n=55 in the triptorelin 3.75 mg group **Characteristics** Eligible women were randomly assigned treatment with E/P pill (gestroden 0.75 mg and ethinylestradiol 0.03 mg) for 12 months vs. triptorelin 3.75 mg slow release every 28 days for 4 months followed by E/P pill for 8 months. **Inclusion Criteria**	**Interventions** Gestodene 0.75 mg/ethinylestradiol 0.03 mg (E/P pill) for 12 months and triptorelin 3.75 mg slow release every 28 days for 4 months followed by E/P pill for 8 months.	**Details** Group allocation was done by telephone call to the randomization centre (1st Obstetric and Gynecology Clinic, University of Milan). Separate randomization lists for each participating centre were used. Whether or not treatment assigned was given, patients remained in the allocated group for intention to treat analysis. Additional treatment for relief of pain with	**Results** **Pain at 8 months during treatment:** <u>Dysmenorrhea:</u> MD=-1.9 (-2.54 to -1.26)* <u>Non menstrual pain:</u> MD=-2.5 (-3.0 to -2.0)* **Pain at the end of the treatment (12 months):** <u>Dysmenorrhea:</u> MD=-2.7 (-3.34 to -2.06)* <u>Non menstrual pain:</u> MD=0.8 (0.33 to 1.27)* * calculated by the 2016 NGA team	**Limitations** Adequate sequence generation?: Unclear risk (No details) Allocation concealment?: Unclear risk (No details) Blinding?: High risk (No blinding of study participants, investigators or assessors reported) Incomplete outcome data addressed?:

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
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 **Ref Id** 338537 **Country/ies where the study was carried out** Italy **Study type:** Multicentric randomised clinical trial. Eight collaborating centres. **Aim of the study:** To compare estroprogestin (E/P pill) given for 12 months vs. a	• Women with laparoscopically confirmed endometriosis and pelvic pain lasting 3-12 months after diagnosis. Only women who reported a score of >=3 for the multidimensional scale and/or >=5 for the analog scale for dysmenorrhea and/or non-menstrual pelvic pain were eligible. **Exclusion Criteria** • Women interested in pregnancy, those who had had previous therapy with GnRH-a or danazol and those who used E/P during the 6 months before the randomisation.		naproxen sodium as first line treatment was allowed, according to physicians and woman's judgment.		Unclear risk (No details on attrition) Free of selective reporting?: Low risk

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
GNRHa treatment given for 4 months followed by E/P pill treatment for 8 months in the relief of endometriosis related pelvic pain. **Study dates:** 1995 - 1996 **Source of funding:** Not reported.					
Full citation 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 **Ref Id** 338552	**Sample size:** A total of 274 patients. All patients received at least one dose of study medication and therefore were included in the ITT population. There was a dropout rate of 35.3% in the DMPA-SC 104 group (48/136) and of 26.1% in the leuprolide group (36/138) during the 6-month treatment period. The majority of these patients either actively withdrew from the study (DMPA-SC 104 21, leuprolide 9) or were lost to follow-up (14 and 11, respectively). Nine patients in each group (6.6% and 6.5% in the DMPA-SC 104 and leuprolide groups, respectively)	**Interventions** DMPA-SC 104 (104 mg/0.65 mL given by SC injection) vs leuprolide (11.25 mg given by IM injection)	**Details** Patients enrolled in this trial were randomized 1:1 to receive either DMPA-SC 104 (104 mg/0.65 mL given by SC injection) or leuprolide (11.25 mg given by IM injection). Both treatments were initiated within the first 5 days of a normal menstrual cycle at visit 1, and a second injection was given 3 months (91 7 days) later, for a total duration of 6 months of active treatment.	**Results** **Endometriosis impact diary** <u>Total hours of productivity lost at employment at 6 months</u> MD = 6.15 (-2.17 to 14.47)* <u>Total hours of productivity lost at employment at 18 months</u> MD = 6.38 (-1.94 to 14.70)* <u>Total hours of productivity lost at housework at 6 months</u> MD = -7.35 (-16.63 to 1.93)* <u>Total hours of productivity lost at housework at 18 months</u> MD = -3.64 (-12.92 to 5.64)* *calculated by the 2016 NGA team	**Limitations** Adequate sequence generation? Unclear (No details) Allocation concealment? Unclear (No details) Blinding of all outcomes? Low risk (The principal investigator and any designated subinvestigators and study coordinators at each center were blinded to the

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Country/ies where the study was carried out Canada/USA **Study type:** Phase 3, multicenter, randomised, evaluator-blinded, comparator-controlled clinical trial **Aim of the study:** The primary efficacy objective - to assess the equivalence of DMPA-SC 104, as compared with leuprolide acetate (2, 12, 13), in the reduction of endometriosis-associated pain. The primary safety objective - to evaluate differential effects of these treatments on bone mineral density (BMD) after 6 months of treatment relative to baseline and to	discontinued as a result of adverse side effects. Of those women who completed the 6 months of active treatment, 51 (58.0%) of 88 in the DMPA-SC 104 group and 58 (56.9%) of 102 in the leuprolide group left the study during the 12-month follow-up period. Th **Characteristics** 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. **Inclusion Criteria** • Patients included in this trial were 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				randomization of each patient) Incomplete outcome data addressed? Low risk (ITT, details given for attrition) Free of selective reporting? Low risk (All primary outcomes stated were reported on) **Other information** None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
assess BMD recovery after 12 months of post-treatment follow-up (month 18). **Study dates:** Not reported **Source of funding:** Not reported	must have persisted for a minimum of 3 months. **Exclusion Criteria** • Women were excluded if their baseline BMD at the lumbar spine and hip had a score that was less than 1.0 SD below the mean for peak adult bone mass. All sexually active women were advised to use nonhormonal contraception throughout the study.				

G.16 Review question: Non-pharmacological management

What is the effectiveness of non-pharmacological therapies (for example acupuncture) for managing pain associated with endometriosis?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation Chen, L, Lin, Y, Yuan, L, Huang, H, Abdominal acupuncture in treating 70 cases of endometriosis dysmenorrhea, International Journal of Clinical Acupuncture, 21, 100-2., 2012 **Ref Id** 437711	**Sample size** N=70 **Characteristics** Age range from 18 to 50, median age 38 y. Disease staging: • severe (13-15 scores): 30%, • moderate (8-12 scores): 43%, • mild (5-7 scores): 27%. Diagnosis was assessed by the Guidelines of Clincal Research in New Drug	**Interventions** Patiens were randomized to: abdominal acupuncture group (n=35) danazol group (n=35)	**Details** **Abdominal acupuncture** was given 7 days before menstruation, once a day on the first through the third days and the following days every other day until the 4th day of menstruation. They were given acupuncture roughly 7 times in each course of treatment. Patients were treated for a continuous 3 courses, after which they were	**Results** Cure (see definition in Methods section): • Acupuncture group = 3/35 • danazol group = 5/35 • RR = 0.60 (95%CI 0.16 to 2.32)* *calculate by the 2016 NGA team	**Limitations** Cochrane risk of bias assessment tool: Adequate sequence generation: Unclear risk (No details on randomisation) Allocation concealment: Unclear risk (No details given) Blinding: High risk (No details given)

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Country/ies where the study was carried out China **Study type** Randomised controled trial. **Aim of the study** To observe the therapeutic effects of abdominal acupuncture on endometriosis dysmenorrhea. **Study dates** Not reported. **Source of funding** Not reported.	Treatment of Traditional Chinese Medicine on Pelvic Endometriosis (subsidiary to the Guidelines of Clinical Research in New Drug Treatment of Traditional Chinese Medicine issued by the Ministry of Health in 2002: 1) progressive endometriosis, 2) discomfort in the lower abdomen and Lumbar sacral area during the menstrual period with gradual aggravation, 3) periodical symptoms of irritation of the rectum with gradual aggravation, 4) tenderness of the tubercle at the posterior fornix, uterosacral ligament and isthmus uteri, 5) adnexa uteri masses of adhesion with palpation of envelope tubercle, 6) obvious change of the size of the adnexa uteri masses before and after the menses. Patients represented with one of the manifestations in (1), (2) or (3) and one of the manifestations in (4), (5) or (6) were diagnosed with endometriosis. **Criteria for staging:** Lower abdominal pain during, before and after the menses, 5 scores (basal score); unbearable		observed in another 3 cycles of menstruation. Abdominal acupuncture: acupoints involved were Zhongwan (RN12), Xiawan (RN10), Qinai (RN6) and Guanyuan (RN4), which led Qi back to Yuan, and Zhongji (RN3), Wailing (ST26), bilateral Xiafengshi points. Wailing (ST26) was punctured of moderate depth, and the others were punctured to the lower 1/3 of the acupoints (Dibu), after which the needles were retained for 30 min. **Danazol group:** patients were administered with oral medication - Danazol capsules - 200mg twice a day, from the first day of menses for a continuous 3 periods. Criteria for therapeutic effects were assessed by standards on dysmennorhea in Guidelines of Clinical Research in New Drug Treatment of Traditional Chinese		Incomplete outcome data addressed: Low risk (No patient was lost during treatment or follow up) Free of selective reporting: Low risk (Outcomes introduced in the methods part were reported) Free of other bias: Unclear risk (Not clear where/how patients were enrolled) **Other information** None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	abdominal pain, 1 score, obvious abdominal pain, 0.5 score; restless, 1 score, shock, 2 scores, pale face, 0.5 score; dripping cold sweat, 1 score; needing to rest in bed, 1 score; affecting work and study, 1 score; no relief by common pain management, 1 score; temporary relief by common pain management, 0.5 score; accompanied by soresness in waist, 0.5 score; accompanied by nausea and vomiting, 0.5 score; accompanied by anus bulge, 1 score; pain <1 day, 0.5 score; pain >1 day, addition of 0.5 score/day. Severe: 13-15 scores, moderate: 8-12 scores, mild: 5-7 scores. **Inclusion criteria** • Women diagnosed with endometriosis dysmenorrhea meeting the criteria for diagnosis described in characteristics. **Exclusion criteria** • Patients accompanied by myoma of uterus, or serious diseass in cardiovascular and cerebrovascular systems, liver, kidney, hemopoietic		Medicine. Cure: complete relief of pain and other symptoms after medication (0 score) and no relapse in the next 3 menstrual cycles.		

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	system, or mental disease; also those allergic to the drugs in this study; pregnant women; patients failing to meet the Inclusion Criteria or failing to take medicine administered by the doctors, or failure in the therapeutic assessment and absence of complete data that might affect the assessment in the study.				
Full citation 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 **Ref Id** 338441 **Country/ies where the study was carried out** UK **Study type** Randomised controlled trial	**Sample size** N = 33 entered trial following randomisation* n = 15 active group n = 18 placebo group *40 women initially agreed to participate in the trial. 13 of these women were randomised to a "wait list control" group, and 27 were randomised to placebo/active treatment groups. After a 16 week period, women in the wait list control group were subsequently eligible for randomisation to the active/placebo treatment arms. However, the wait list control group was subsequently suspended in December 2007 due to high drop out (7/13). The 6 women who remained in the wait list control then entered a secondary randomisation	**Interventions** Women randomised to the **active treatment arm** received individualised formulations of between 10 and 15 herbs selected form the Chinese material medica with a daily dosage amounting to between 150g and 250g. Subjects allocated to the **placebo arm** were given packets identical in appearance to the active treatment arm, but which contained a decoction made from culinary herbs and dried foods.	**Details** Monthly consultations (lasting 20-30 minutes) were held with a practitioner of Chinese Herbal Medicine. A month's supply of herbs was soaked in 9L of water for 40 minutes, and then cooked for 1 hour. The precooked herbs were then dispensed into 180ml dosages in sealed plastic packets, to be taken twice daily. The duration of the trial was 16 weeks, with a four-week run in period to ensure stable and measurable levels of endometriosis pain. A group of Western herbal practitioners had previously agreed that the placebo decoction	**Results** **Pain scores using Visual Analogue Scores, change (from baseline) at week 16** Period pain mean change (10cm VAS) • CHM group = -2.36 (SD 2.22), n = 7 • Placebo group = -1.14 (SD 2.29), n = 5 • Adjusted mean difference between groups = -1.22 (95% CI -3.81 to 1.37)* Pain during intercourse mean change (10cm VAS) • CHM group = -2.98 (SD 1.56), n = 5 • Placebo group = -3.74 (SD 1.62), n = 3 • Adjusted mean difference between	**Limitations** Cochrane risk of bias assessment tool Adequate sequence generation: Low risk (Randomisation for allocation of the groups was generated through computer generated random numbers) Allocation concealment: Low risk (Allocation sequence was concealed through sealed, opaque envelopes) Blinding: Low risk (Practitioner and subjects were unaware of group

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr>
<td>Aim of the study
To test the feasibility of a novel methodology for inve stigating individualise d Chinese Herbal Medicine preparations rigorously, and to gather preliminary data on treatment effect for a larger, definitive trial.

Study dates
October 2006 to August 2008.

Source of funding
The post of one of the authors was funded by a grant from the Rufford Maurice Laing Foundation. No other Source of funding reported.</td>
<td>process to be allocated to either placebo or active treatment, resulting in N=33 total participants.

Characteristics
<table>
<tr><td>Charact eristics</td><td>Placebo group n = 15)</td><td>Active treatme nt group(n = 13)</td></tr>
<tr><td>Age, years, mean (SD)</td><td>35.7 (8)</td><td>33.2 (7.2)</td></tr>
<tr><td>Duration , years, mean (SD)</td><td>12.6 (8.9)</td><td>11.2 (5.8)</td></tr>
<tr><td colspan="3">Relationship status, n (%)</td></tr>
<tr><td>Single</td><td>7 (47%)</td><td>5 (38.5%)</td></tr>
<tr><td>Married/ co-habiting</td><td>6 (40%)</td><td>5 (38.5%)</td></tr>
<tr><td>Missing</td><td>2 (13%)</td><td>3 (23%)</td></tr>
<tr><td>Number using hormon al</td><td>2 (13%)</td><td>5 (38.5%)</td></tr>
</table>
</td>
<td></td>
<td>did not contain ingredients that had therapeutic action for endometriosis. Prior to the trial, a group of CHM naïve volunteers found the placebo to be as plausible as CHM in taste and appearance.
Four visual analogue scales (VAS) were used to measure weekly variations in menstrual pain, pain on intercourse, pain on bowel movement and daily pain. The Measure Your Own Medical Outcomes Profile (MYMOP) was completed once per month. This allowed participants to identify two symptoms that bothered them the most and an activity restricted by endometriosis, and to rate their level of wellbeing using a 1-7 point Likert scale. The Endometriosis Health Profile-30 (EHP-30) was completed at the start and end of the trial.</td>
<td>groups = 0.76 (95% CI -1.52 to 3.05)*
<u>Pain on bowel movement mean change (10 cm VAS)</u>
• CHM group = -0.88 (SD 2.51), n = 7
• Placebo group = -0.96 (SD 2.61), n = 5
• Adjusted mean difference between groups = 0.08 (95% CI -2.86 to 3.03)*
<u>Daily pain mean change (10 cm VAS)</u>
• CHM group = -0.83 (SD 2.32), n = 7
• Placebo group = -1.57 (SD 2.35), n = 6
• Adjusted mean difference between groups = 0.74 (95% CI -1.81 to 3.29)*

MYMOP scores change (from baseline) at week 16 (7-point Likert scale)
<u>Mean change in symptom 1 of MYMOP score</u>
• CHM group = -2.15 (SD 1.97), n = 8
• Placebo group = -1.57 (SD 1.96), n = 10
• Adjusted mean difference between</td>
<td>allocation, and placebo/active treatments were provided in identical plastic packets.)
Incomplete outcome data addressed: High risk (There were 2 dropouts and 2 mid-trial dropouts in the active group. There were 3 dropouts and 2 mid-trial dropouts in the placebo group)
Selective reporting: Low risk (outcomes adequately reported compared with the descriptions in the methods)
Free of other bias: Unclear risk (Selection bias is likely, as recruitment to the trial was extremely difficult through NHS sources, so participants all self-referred to the study organisers)</td>
</tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	<table><tr><td>medication, n (%)</td><td></td><td></td></tr><tr><td colspan="3">Pretreatment scores, mean (SD) [number of respondents</td></tr><tr><td>Period pain</td><td>6.8 (1.9) [12]</td><td>6.6 (2.4) [11]</td></tr><tr><td>Pain during sex</td><td>3.1 (2.65) [7]</td><td>5.2 (2.9) [6]</td></tr><tr><td>Pain with bowel movement</td><td>3.2 (2.3) [12]</td><td>4.9 (3.4) [9]</td></tr><tr><td>Daily pain</td><td>4.0 (2.2) [13]</td><td>4.9 (2.3) [10]</td></tr><tr><td colspan="3">Number of women with severe pain before treatment, n (%)</td></tr><tr><td>Period pain VAS >7</td><td>9 (60%)</td><td>9 (69.2%)</td></tr><tr><td>Pain during sex VAS >5</td><td>2 (13.3%)</td><td>4 (30.7%)</td></tr><tr><td>Pain with</td><td>3 (20%)</td><td>5 (38.5%)</td></tr></table>		A computer generated random numbers table was used for both phases of randomization to produce an irregular block allocation sequence. Codes for each group allocation (treatment or wait list control) were transferred to sealed opaque envelopes and this information was relayed to the practitioner. An additional randomisation took place at the dispensary using opaque brown envelopes that divided participants into either active or placebo arms. This information was not presented to the practitioner or participants until after the conclusion of the whole trial.	groups = -0.58 (95% CI -2.41 to 1.25)* Mean change in symptom 2 of MYMOP score • CHM group = -2.41 (SD 1.93), n = 8 • Placebo group = -1.51 (SD 1.90), n = 10 • Adjusted mean difference between groups = -0.90 (-2.68 to 0.88)* Mean change in limitation of activity due to endometriosis on MYMOP score • CHM group = -2.19 (SD 1.71), n = 8 • Placebo group = -1.50 (SD 1.69), n = 9 • Adjusted mean difference between groups = -0.69 (95% CI -2.31 to 0.93)* Mean change in well-being on MYMOP score • CHM group = -2.01 (SD 1.97), n = 7 • Placebo group = -0.95 (SD 1.93), n = 10 • Adjusted mean difference between groups = -1.06 (-2.94 to 0.82)*	**Other information** None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	<table><tr><td>bowel movement >5</td><td></td><td></td></tr><tr><td>Daily pain >5</td><td>3 (20%)</td><td>6 (46.2%)</td></tr></table> SD standard deviation, VAS visual analogue scale **Inclusion criteria** • Women with a laparoscopically confirmed diagnosis of endometriosis, with relatively stable and measurable symptoms of disease, who were naïve to Chinese Herbal Medicine (therefore unable to distinguish between active and placebo preparations). **Exclusion criteria** • Women who had received surgery, started conventional medical treatment in the past three months, reported other conditions associated with pelvic pain, who had hepatic or renal complications, who were pregnant or taking any drugs known to interact with Chinese Herbal Medicine.			**EHP-30 scores change (from baseline) at week 16** Mean change in pain scores • CHM group = -6.43 (SD 10.1), n = 11 • Placebo group = -6.11 (SD 10.3), n = 7 • Adjusted mean difference between groups = -0.32 (-10.01 to 9.37)* Mean change in control and powerlessness scores • CHM group = -7.49 (SD 5.83), n = 11 • Placebo group = -5.76 (SD 5.99), n = 7 • Adjusted mean difference between groups = -1.73 (-7.35 to 3.89)* Mean change in emotional well-being • CHM group = -4.49 (SD 4.16), n = 11 • Placebo group = -4.12 (SD 4.28), n = 7 • Adjusted mean difference between groups = -0.37 (-4.38 to 3.64)* Mean change in social support	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				• CHM group = -4.19 (SD 4.52), n = 11 • Placebo group = -1.48 (SD 4.69), n = 7 • Adjusted mean difference between groups = -2.71 (-7.09 to 1.67)* <u>Mean change in self-image</u> • CHM group = -2.57 (SD 2.79), n = 11 • Placebo group = -3.03 (SD 2.86), n = 7 • Adjusted mean difference between groups = 0.46 (-2.22 to 3.14)* * Calculated by the 2016 NGA team	
Full citation Flower, A., Liu, J. P., Lewith, G., Little, P., Li, Q., Chinese herbal medicine for endometriosis, Cochrane Database of Systematic Reviews, 5, CD006568, 2012 **Ref Id** 346769 **Country/ies where the study was carried out**	**Sample size** 58 cases of endometriosis, confirmed by laparoscopy. **Characteristics** Experimental group 1: 16 Experimental group 2: 24 Control group: 18 Drop-out rate: 0 **Inclusion criteria** Not reported. **Exclusion criteria**	**Interventions** **Experimental group** 1: Nei Yi pills (10g twice daily) Experimental group 2: Nei Yi pills (10g twice daily) plus Nei Yi enema (70ml daily) **Control group**: danazol (400mg/day)	**Details** Chinese validated outcomes (CAITWN 1991) used and divided responses to treatment into four categories: 'symptomatic relief' described a complete resolution of all symptoms and signs and included pregnancy, when desired, within three years of stopping treatment; 'significant improvement'	**Results** **Chinese herbal medicine (CHM) (oral) vs danazol**: <u>Symptomatic relief:</u> RR (95%CI) = 5.06 [1.28 to 20.05] <u>Dysmenorrhea score</u>: RR (95%CI) = -1.01 [-3.11, 1.09] <u>Lumbosacral pain relief:</u> RR (95%CI) = 1.21 [0.86, 1.70] <u>Rectal irritation relief:</u> RR (95%CI) = 1.67 [0.90, 3.10]	**Limitations** <u>Cochrane risk of bias assessment tool</u> Adequate sequence generation: Low risk (Randomisation for allocation of three groups was generated through random number table) Allocation concealment: Low

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
China **Study type** Parallel randomised controlled trial. **Aim of the study** To review the effectiveness and safety of Chinese herbal medicine (CHM) in alleviating endometriosis-related pain and infertility. **Study dates** December 1999 to October 2003. **Source of funding** Funding source declared.	Not reported.	Nei Yi pills consisted of: Dan Shen (Salviae multiorrhizae Radix), Xue Jie (Draconis Sanguis), San Leng (Sparganii Rhizoma), E Zhu (Curcumae Rhizoma), Tao Ren (Persicae Semen), San Qi (Notoginseng Radix), Dang Gui (Angelica sinensis), Gui Zhi (Cinnamomi Ramulus), Xiang Fu (Cyperi Rhizoma), Niu Xi (Achyranthis bidentate Radix) Nei Yi enema consisted of: Dan Shen (Salviae multiorrhizae Radix), Xue Jie (Draconis Sanguis), Chi Shao (Paeonia rubra Radix), Hu Zhang (Radix et Rhizoma Polygoni Cuspidati), San Leng (Sparganii Rhizoma), E Zhu (Curcumae Rhizoma), Tao Ren (Persicae Semen) Treatment duration: 3 months	described when most symptoms resolved and pelvic masses were reduced in size; 'improvement' described symptomatic improvement and no worsening of symptoms within three months of stopping the treatment but only minor or no change in pelvic masses; and finally 'no effect' was where symptoms either remained unchanged or worsened during the intervention.	Tenderness of vaginal nodules in posterior fornix: RR (95%CI) = 1.31 [0.87, 1.97] Adnexal masses disappearance or shrinkage: RR (95%CI) = 1.41 [0.79, 2.50] **Chinese herbal medicine (oral + enema) vs danazol** Symptomatic relief: RR (95%CI) = 5.63 [1.47, 21.54] Dysmenorrhea score: RR (95%CI) = -2.9 [-4.55, -1.25] Lumbosacral pain relief: RR (95%CI) = 1.15 [0.82, 1.62] Rectal irritation relief: RR (95%CI) = 1.78 [0.99, 3.20] Tenderness of vaginal nodules in posterior fornix: RR (95%CI) = 1.26 [0.84, 1.90] Adnexal masses disappearance or shrinkage: RR (95%CI) = 1.70 [1.04, 2.78]	risk (Allocation sequence was concealed through numbered, sealed, opaque envelopes) Blinding: High risk (Although described as patient and assessor blinded (and confirmed with author) there is no description of an attempt to match the herbal enema with an inert control, so it is very unlikely patients were not aware of which group they were allocated to) Incomplete outcome data addressed: Low risk (No patient was lost during treatment or follow up) Free of selective reporting: Low risk (Identified outcomes adequately reported compared with the descriptions in the methods)

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				Chinese herbal medicine (oral+ enema) vs Chinese herbal medicine (oral) Symptomatic relief: RR (95%CI) = 1.11 [0.65, 1.89] Dysmenorrhea score: RR (95%CI) = -1.89 [-3.89, 0.11] Lumbosacral pain relief: RR (95%CI) = 0.95 [0.74, 1.23] Rectal irritation relief: RR (95%CI) = 1.07 [0.79, 1.44] Tenderness of vaginal nodules in posterior fornix: RR (95%CI) = 0.96 [0.74, 1.25] Adnexal masses disappearance or shrinkage: RR (95%CI) = 1.21 [0.85, 1.72]	Free of other bias:Low risk (No source of other bias) **Other information** None
Full citation 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	**Sample size** N=22 women with deep endometriosis. **Characteristics** Women with deep endometriosis diagnosed in the cul-de-sac and intestinal loop who sustained pelvic pain and/or deep	**Interventions** **Group 1** – acupuncture-like TENS (Dualpex 9611) (n = 11) **Group 2** – self-applied TENS (Tanyx1) (n = 11)	**Details** **Acupuncture-like TENS:** Frequency: 8 Hz Pulse duration: 250μs and VIF (variation in intensity and frequency of 1ms) Intensity: adjusted according to the	**Results** Mean scores for quality of life (EHP-30; the better the quality of life the lower the total score): • Acupuncture-like TENS: pre treatment =47.98 SD 11.18, post treatment =32.09 SD 8.65, n=11	**Limitations** Cochrane risk of bias assessment tool Adequate sequence generation: Unclear risk (Randomisation for allocation of two groups was

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Electrical Nerve Stimulation (TENS): randomized controlled trial, European Journal of Obstetrics, Gynecology, & Reproductive Biology, 194, 1-6, 2015 **Ref Id** 437773 **Country/ies where the study was carried out** Brazil **Study type** Non-blind, randomized clinical trial, randomized controlled trial. **Aim of the study** To primarily evaluating the effectiveness of electrotherapy with TENS as a complementary treatment of pelvic pain and/or deep dyspareunia, as well its impact on quality of life of women suffering from deep	dyspareunia, despite continuous clinical medication. All women were undergoing hormone therapy with continuous progestin alone or combined oral contraceptives for at least three months, reporting pelvic pain and/or deep dyspareunia persistence, associated or not with other pain complaints (dysmenorrhea, dyschezia and dysuria). **Inclusion criteria** • Women at menacme, ranging from 18 to 50 years-old, diagnosed with deep endometriosis in the cul-de-sac and/or intestinal loop using imaging tests with ultrasonography after bowel preparation. **Exclusion criteria** • Women with decreased skin sensitivity, implanted with a pacemaker, skin hypersensitivity (allergic reactions to gel or electrodes), epilepsy, heart disease (cardiac arrhythmia), osteosynthesis in the region of application, full-thickness defects of the	TENS was applied at the S3–S4 region for both groups.	woman (‘‘strong, but comfortable’’) without any motor stimulation. Application site: sacral region (S3–S4). Method: A dual-channel TENS unit was used, equipped with four rubber electrodes (5 cm to 3 cm) and neutral aqueous gel lubricant, attached to the skin with adhesive tape crossed in an ‘‘X’’ pattern. Time: 30 min and sessions were performed once a week, for a period of 8 weeks. **Self-applied TENS:** Frequency: 85 Hz Pulse duration: 75μs Intensity: adjustable in three options: 10, 20 or 30mA. Women were instructed to choose the intensity that was ‘‘strong, but comfortable’’ Application site: sacral region (S3–S4) Method: The correct placement of the	• Self-applied TENS: pre treatment =61.18 SD 9.32, post treatment =46.88 SD 13.91, n=11 • MD = 1.59 (95%CI -6.45 to 9.63)* • (using a calculator of 0.7 to calculate SD; mean difference in QoL from baseline (EHP-30): acupuncture-like TENS = -15.98 SD 0.3, n=11 • self-applied TENS = -14.5 SD 9.94, n=11) *calculated by the 2016 NGA team	generated by a computer program, no details given) Allocation concealment: Unclear risk (Allocation was done through opaque, sealed envelopes, not reported in what sequence) Blinding: High risk (non-blind, randomized clinical trial) Incomplete outcome data addressed: Low risk (No patient was lost during treatment or follow up) Free of selective reporting: Low risk (Identified outcomes adequately reported compared with the descriptions in the methods) Free of other bias:Low risk (No source of other bias)

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
endometriosis with persistent pain complaints, despite the use of hormone therapy. **Study dates** November 2013 to June 2014. **Source of funding** Study was partially funded by the Research Support Foundation of the State of Sa˜o Paulo (FAPESP), process n 2013/ 11790-2.	skin, malignant tumors, acute inflammatory disease, and cognitive deficiency.		device was initially explained and demonstrated on the patient during evaluation, and doubts were dispelled by the researcher. TENS application was performed at home by the patient herself. She could follow instructions from a didactic illustration showing the exposed sacral region of a supine woman next to another illustration of the same woman with the equipment in place. Time: Twice a day, 20 min per application, setting an interval of 12 h between applications. A return visit was scheduled after four weeks of treatment for followup of the use of the device. A final reassessment was carried out after 8 weeks.		Other information
Full citation Sesti, F., Capozzolo, T., Pietropolli, A., Marziali, M., Bollea, M. R., Piccione, E., Recurrence rate of endometrioma after	**Sample size** N=259 Of 264 women selected as eligible subjects to enter the trial, 5 were excluded because they refused to participate. The remaining	**Interventions** The patients were randomly allocated to one of four post-operative management arms: • **placebo** (n = 65)	**Details** **Surgical treatment:** The laparoscopic removal of endometrioma was performed as follows. As first step, pelvis,	**Results** <u>Recurrence of endometrioma (n (%)):</u> • Placebo = 10 (16.6%) n = 60	**Limitations** <u>Cochrane risk of bias assessment tool</u> Adequate sequence generation: Low

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
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 **Ref Id** 338560 **Country/ies where the study was carried out** Italy **Study type** Randomised comparative trial. **Aim of the study** To assess the endometrioma recurrence rate after laparoscopic cystectomy plus hormonal suppression treatment or plus dietary therapy	259 women underwent laparoscopic cystectomy. placebo (randomized n=65, analyzed n = 60) GnRH-a (randomized n=65, analyzed n = 58) continuous low-dose monophasic oral contraceptives (randomized n=64, analyzed n = 64) dietary therapy (randomized n=65, analyzed n = 62) (see Interventions) **Characteristics** The study population was selected from women who were referred to Endometriosis Center, Section of Gynecology, Tor Vergata University Hospital, Rome, between January 2004 and August 2006. No women were attempting to conceive at the time of study entry. **Inclusion criteria** • Reproductive age, up 40 years of age at the time of surgery; • ultrasonographic evidence of endometrioma; • moderateto-severe endometriosis-related painful symptoms (graded	• **GnRH-a** (tryptorelin or leuprorelin, 3.75 mg every 28 days) (n = 65) • **continuous low-dose monophasic oral contraceptives** (ethynilestradiol, 0.03 mg plus gestoden, 0.75 mg) (n = 64) • **dietary therapy** (n = 65) for 6 months Laparoscopic cystectomy plus placebo group was used as control. Dietary therapy was a protocol consisting of nutritional intake additioned to vitamins (B6, A, C, E), mineral salts (Ca, Mg, Se, Zn, Fe), lactic ferments VSL3 (Bifidobacterium breve, Bifidobacterium longum, Bifidobacterium infantis, Lactobacillus acidophilus, Lactobacillus casei, Lactobacillus bulgaricus,	abdomen, uterus and tubo-ovarian structures were inspected for possible evidence of disease. If necessary, lysis of adhesions was performed to fully mobilize the ovaries. A sharp cortical incision was made, and a cleavage plane was developed by sharp dissection. The entire cyst was enucleated and stripped from the normal ovarian tissue, using bilateral traction. Hemostasis was achieved with bipolar forceps, avoiding contact with the external ovarian surface for preventing adhesion formation and cortical damage. The ovarian cysts were removed from the abdomen into the trocars, or using a disposable endobag. All areas of superficial active endometriosis involving the ovaries or the pelvic peritoneum were treated by bipolar coagulation. Radicality of the procedures was defined as complete excision of all evident	• GnRH-a = 6 (10.3%) n = 58 • Estroprogestin = 9 (15%) n = 60 • Dietary therapy = 11 (17.8%) n = 62 • RR diet vs placebo = 1.06 (95%CI 0.49 to 2.32)* • RR diet vs GnRHa = 1.72 (95%CI 0.68 to 4.34)* • RR diet vs Estroprogestin = 1.18 (95%CI 0.53 to 2.65)* *calculated by t he 2016 NGA team	risk (Randomisation for allocation of three groups was generated through a computer randomisation sequence) Allocation concealment: Low risk (Allocation sequence was concealed through serially numbered, opaque, sealed envelopes) Blinding: Low risk (Neither the surgeons nor the ultrasonography operator nor the patients were aware of the regimen prescribed) Incomplete outcome data addressed: Unclear risk (19 women withdrew) Free of selective reporting: Low risk (Identified outcomes adequately reported compared with the

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
compared to post-operative placebo. **Study dates** January 2004 to August 2006. **Source of funding** Not reported.	as 4 on a 10-point by visual analogue scale) (VAS); • laparoscopic diagnosis of endometrioma staged according to American Fertility Society Classification of Endometriosis; • first laparoscopic surgery for endometriosis, and conservative treatment with retention of uterus and ovaries; • complete excision of all evident ovarian and peritoneal disease; ultrasonographic and clinical follow-up after surgery. **Exclusion criteria** • Patients who had received 6 months estrogen-suppressing drugs before first surgery were excluded from the study. Other Exclusion Criteria were: usual contraindications to estrogens and progestins; previous surgical treatment for endometriosis; surgical findings of concomitant deeply infiltranting endometriosis.	Streptococcus thermophilus), omega-3 and omega-6 fatty acids (fish oil), which secured nutritional rate between 1600 and 2000 calories.	ovarian and peritoneal disease. Seven days after surgery, when a definitive histological diagnosis of endometriosis was available, randomization was performed according to a computer-generated randomization sequence using serially numbered, opaque, sealed envelopes. At 18 months' follow-up, 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. When the cyst was indistinguishable from a transient corpus luteum cyst or an intraovarian haematoma, the diagnosis of recurrence was made only when the cyst had not disappeared after 30 days. Second-look		descriptions in the methods) Free of other bias: Low risk (No source of other bias) **Other information** Nonr

<table>
<tr><th>Study details</th><th>Participants</th><th>Interventions</th><th>Methods</th><th>Outcomes and Results</th><th>Comments</th></tr>
<tr><td></td><td></td><td></td><td>laparoscopy was performed in patients with ultrasonographic scan suggesting recurrent endometrioma.
The outcome was the endometrioma recurrence rate after post-operative hormonal suppression treatment or dietary therapy compared to control-group.</td><td></td><td></td></tr>
<tr><td>Full citation
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</td><td>Sample size
N = 18

Characteristics
<table>
<tr><th>Characteristics</th><th>Active group n = 10</th><th>Sham group n = 8</th></tr>
<tr><td>Age, years, mean (SD)</td><td>17.8 (2.1)</td><td>17.0 (2.1)</td></tr>
<tr><td>Sexually active</td><td>50%</td><td>50%</td></tr>
<tr><td>Mean pain score (SD)</td><td>7.7 (2.4)</td><td>7.4 (0.9)</td></tr>
<tr><td>Time since surgery, months</td><td>7.4 (8.9)</td><td>9.5 (15.9)</td></tr>
</table></td><td>Interventions
Participants were assigned to either acupuncture intervention, or sham acupuncture. Both groups underwent 2 acupuncture treatments per week for 8 consecutive weeks (a total of 16 treatments).
Active acupuncture treatments followed guidelines defined and written in a treatment manual, developed by three senior practitioners. Treatments were individually tailored according to the participants' symptoms.</td><td>Details
The study used a style of Japanese acupuncture following the Japanese acupuncture training curriculum at the New England School of Acupuncture. This uses smaller needles, inserts needles less deeply and with less manipulation than traditional Chinese medicine acupuncture. Treatments were administered by licensed acupuncturists with formal training, who also underwent a specific 6-hour training session to learn the specific active and sham acupuncture</td><td>Results
Pain scores, measured with Visual Analogue Scale (0-10)
<u>Change (from baseline) in pain during the last four weeks, measured at 4 weeks</u>
• Acupuncture group = -4.8 (SD 2.4), n = 9
• Sham group = -1.4 (SD 2.1), n = 5
• Mean difference = -3.4 (95% CI -5.82 to -0.98)*
<u>Change (from baseline) in pain during the last four weeks, measured at 8 weeks</u>
• Acupuncture group = -4.3 (SD 3.6), n = 9
• Sham group = -3.8 (SD 1.7), n = 6</td><td>Limitations
<u>Cochrane risk of bias assessment tool</u>
Adequate sequence generation: Unclear risk (no details are provided regarding sequence generation)
Allocation concealment: Unclear risk (no details are provided regarding allocation concealment)
Blinding: Low risk (sham-acupuncture control was used, and the degree to</td></tr>
</table>

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments

Ref Id
424789

Country/ies where the study was carried out
USA

Study type
Randomised sham-controlled trial

Aim of the study
To assess feasibility and collect preliminary data for a subsequent trial to evaluate Japanese-style acupuncture for reducing chronic pelvic pain and improving health-related quality of life in adolescents with endometriosis.

Study dates
Not reported.

Source of funding
A grant from the National Center for Complementary and Alternative Medicine.

mean, (SD)		
Stage of endometriosis		
Stage 1	100%	100%
EHP-30 score, mean (SD)	36.5 (20.2)	44.9 (16.5)
Pediatric QoL inventory score, mean (SD)	65.1 (14.4)	61.9 (13.0)
Activity scale, mean (SD)	6.6 (2.3)	6.3 (2.5)
Perceived Stress Scale mean (SD)	1.6 (0.7)	1.8 (0.6)

SD standard deviation

Inclusion criteria

- Women aged 13-22 with a diagnosis of stage I, II or III endometriosis determined by laparoscopic surgery within the past 5 years

protocols employed in this study.

Treatment protocols included:

1. needling 8-12 points to activate and balance Extraordinary and Divergent acupuncture channels
2. burning of small threads of a 'warming' herb (moxibustion) on both back shu acupuncture points and sacral areas that affect the pelvic region
3. electro-stimulation of reactive auricular acupuncture points using the Hibiki-7 device

Sham acupuncture was designed to mimic active treatments, while being minimally active. A validated, sham-acupuncture device that does not penetrate the skin was used.

All outcome measures were assessed at baseline, and at 4 weeks, 8 weeks and 6 months following the start of treatment. The main treatment outcome was change in pelvic pain not

- Mean difference = -0.5 (95% CI -3.22 to 2.22)*

Change (from baseline) in pain during the last four weeks, measured at 6 months

- Acupuncture group = -3.6 (SD 3.0), n = 9
- Sham group = -2.8 (SD 3.8), n = 5
- Mean difference = -0.8 (95% CI -4.66 to 3.06)*

EHP-30 total scores (range 0-100)

Change (from baseline) in scores, measured at 4 weeks

- Acupuncture group = -17.2 (SD 18.3), n = 9
- Sham group = 4.3 (SD 15.0), n = 5
- Mean difference = -21.50 (-39.27 to -3.73)*

Change (from baseline) in scores, measured at 8 weeks

- Acupuncture group = -16.6 (SD 24.8), n = 9
- Sham group = 3.1 (SD 13.4), n = 6
- Mean difference = -19.70 (95% CI -39.13 to -0.27)*

which patients were blinded to their allocation did not differ between groups)

Incomplete outcome data addressed: High risk (There was 1 dropout in the acupuncture group and 3 dropouts in the sham group)

Selective reporting: Low risk (outcomes adequately reported compared with the descriptions in the methods)

Free of other bias: Low risk

Other information
None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	• Persisting pelvic pain with an intensity between 2 and 8 on a 1-point numerical scale • Post menarchal, intact uterus and at least one ovary • A candidate for, or already using, combination hormonal therapy (oral contraceptive pill, contraceptive patch or contraceptive vaginal ring) • No prior experience with acupuncture • Living within 2 hours of the Boston metropolitan area. **Exclusion criteria** • pregnant or lactating • history of drug or alcohol abuse • use of a GnRH analogue within the 6 months prior to their participation in the study • co-existing disabling physical or psychiatric conditions that the study physician believed might interfere with participation in the study		associated with menses and sexual activity, and was assessed after 8 weeks of treatment. A numerical analogue scale was used to rate pain severity during the past 4 weeks from 0 to 10. Secondary outcomes associated with health related quality of life (HRQOL) were assessed with the Endometriosis Health Profile-30 (EHP-30) - scores range from 0-100; a lower score reflects fewer symptoms and better HRQOL the Pediatric Quality of Life Inventory - scores range from 0-100; a higher score indicates better HRQOL a participant generated list of 3 activities made difficult due to pelvic pain - rated on a score of 0-10; higher scores indicate the activity is more difficult to perform	<u>Change (from baseline) in scores, measured at 6 months</u> • Acupuncture group = -17.9 (SD 21.9), n = 9 • Sham group = 3.0 (SD 10.8), n = 5 • Mean difference = -20.90 (95% CI -38.06 to -3.74)* **Pediatric Quality of Life Inventory scores (range 0-100)** <u>Change (from baseline) in scores, measured at 4 weeks</u> • Acupuncture group = 6.6 (SD 16), n = 9 • Sham group = -3.5 (SD 9.5), n = 5 • Mean difference = 10.10 (95% CI -3.26 to 23.46)* <u>Change (from baseline) in scores, measured at 8 weeks</u> • Acupuncture group = 11.1 (SD 19.9), n = 9 • Sham group = -3.1 (SD 9.7), n = 6 • Mean difference = 14.20 (95% CI -0.94 to 29.34)* <u>Change (from baseline) in scores, measured at 6 months</u>	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				• Acupuncture group = 15.1 (SD 18.2), n = 9 • Sham group = 0.2 (SD 7.8), n = 5 • Mean difference = 14.90 (95% CI 1.18 to 28.62)* **3-activity scale (range 0-10)** <u>Change (from baseline) in scores, measured at 4 weeks</u> • Acupuncture group = -3.4 (SD 2.2), n = 9 • Sham group = -0.5 (SD 1.5), n = 5 • Mean difference = -2.90 (95% CI -4.85 to -0.95)* <u>Change (from baseline) in scores, measured at 8 weeks</u> • Acupuncture group = -2.6 (SD 3.2), n = 9 • Sham group -0.8 (SD 2.1), n = 6 • Mean difference = -1.80 (95% CI -4.48 to 0.88)* <u>Change (from baseline) in scores, measured at 6 months</u> • Acupuncture group = -3.6 (SD 2.6), n = 9 • Sham group = -1.9 (SD 3.5), n = 5	

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				• Mean difference = -1.70 (95% CI -5.21 to 1.81)* *Calculated by the 2016 NGA team	
Full citation Xia, T, Effect of Acupuncture and Traditional Chinese Herbal Medicine in Treating Endometriosis, International Journal of Clinical Acupuncture, 15, 145-50., 2006 **Ref Id** 437769 **Country/ies where the study was carried out** China **Study type** Randomised controlled study. **Aim of the study** To compare the clinical effect of acupuncture and Chinese herbal medicine with danazol in treating endometriosis.	**Sample size** N=78 **Characteristics** 78 women with confirmed endometriosis according to the Diagnostic and Treatment Criteria of Endometriosis by Integrative Chinese-Western Medicine, revised at the 3rd Academic Conference of Speciality Committee of Gynecology, China Association of Integrative Chinese-Western Medicine in 1991. Patients were randomly divided into a treatment group (n=40) and a control group (n=38). In the treatment group the disease duration was 0.5-14 (mean 5.4) years, in the control group the disease duration was 0.7-13 (mean 36.2) years. **Inclusion criteria** • Women with confirmed endometriosis according to	**Interventions** **Intervention group:** • **Acupuncture:** the points included: Sanjiajiu (Ex), Zhongji (CV3), bilateral Shangliao (UB31), Ciliao (UB32), Zhongliao (UB33), Xialiao (Ub34), Sanyinjiao (SP6). 20 to 30 min. of moderate moxibustion with a moxa stick was performed on Sanjiaojiu (Ex) and the heat sensation was regulated to the patients' tolerance. Zhongji (CV3) was punctured 1.5-2.5 cun sensation was regulated to the patients' tolerance. Zhongji (CV3) was punctured 1.5-2.5 cun perpendiculalrly and stimulated with a reducing	**Details** Therapeutic effect criteria were developed according to the Diagnostic and Treatment Criteria of Endometriosis by Integrative Chinese-Western Medicine, revised in the 3rd Academic Conference of the Speciality Committee of Gynecology, China Association of Integrative Chinese Western Medicine in 1991. Clinical recovery: all of the symptoms disappeared, the local signs of pelvic nodules basically disappeared and the infertile patients got pregnant within 3 days. Markedly effective: the symptoms basically disappeared and the pelvic nodules shrank by more than half and the infertility patients	**Results** Therapeutic effect in both comparison groups **Cessation of signs and symptoms:** <u>Dysmenorrhea:</u> • intervention group = 16/40 • control group = 13/38, • RR (95%CI) = 1.28 (95%CI 0.51 to 3.22)* <u>Lumbo-sacral pain:</u> • intervention group = 15/40 • control group = 12/38, • RR (95%CI) = 1.30 (95%CI 0.51 to 3.32)* <u>Dyspareunia:</u> • intervention group = 5/40 • control group = 2/38, • RR (95%CI) = 2.57 (95%CI 0.47 to 14.14)* *calculated by the 2016 NGA team	**Limitations** <u>Cochrane risk of bias assessment tool</u> Adequate sequence generation: Unclear risk (No details on randomisation) Allocation concealment: Unclear risk (No details given) Blinding: High risk (No details given) Incomplete outcome data addressed: Low risk (No patient was lost during treatment or follow up) Free of selective reporting: Low risk (Outcomes introduced in the methods were reported) Free of other bias: Unclear risk (Not clear where/how

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Study dates Not reported. **Source of funding** Not reported.	the criteria described in Characteristics. **Exclusion criteria** Not reported.	manipulation by rotation, for 1 min. every 5 min. during the 15-20 min. needle retention period. Shangliao (UB31), Ciliao (UB32), Zhongliao (UB33) and Xialiao (UB34) were treated first by performing 20-30 min. of moxibustion with a moxa box that covered the four-point area and then by moderate tapping with a plum-blossom needle intil the local area was slightly bleeding. Sanyinjiao (SP6) was punctured 1.5-2 cun perpendicularly with a reinforcing manipulation by rotation and manipulated 1 min. every 5 min. during the 15-20 min. needle retention period. The acupuncture therapy started 9 days before the period and was	were able to conceive despite the existence of local symptoms. Effective: the symptoms were alleviated, the pelvic nodule shrank by more than 1/3 and the symptoms remained stable for 3 months after discontinuing the treatment. Failure: the major symptoms remained unchanged or turned worse and the local signs deteriorated.		patients were enrolled) **Other information** None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
		discontinued during the period. • **-Chinese herbal medicine (CHM):** Gui-zhi-fu-ling-wan: Ramulus Cinnamomi-10g, Poria - 15g, Radix Paeoniae Rubra-15g, Semen Persicae-10g, Cortex Moutan-15g. The medicine was taken for 3 menstrual cycles. **Control group:** 200 mg danazol was administered twice a day. For both groups one treatment course consisted of 3 consecutive months of treatment.			
Full citation Xiang, D., Situ, Y., Liang, X., Cheng, L., Zhang, G., Ear acupuncture therapy for 37 cases of dysmenorrhea due to endometriosis, Journal of Traditional Chinese Medicine, 22, 282-5, 2002	**Sample size** n=67 **Characteristics** 67 women ages 22-47 years. Diagnostic criteria met for endometriosis (Guideline for Clinical Research on New Chinese Drugs for Treatment of Pelvic Endometriosis, 1993). Participants were	**Interventions** **Ear acupuncture therapy (EAT):** Ting Zong (centre of cymba auriculae), Pi Zhi Xia (hypo-cortex), Nei Fen Mi (endocrine), Jiao Gan (sympathetic) and Nei Sheng Zhi Qi (internal genitals).	**Details** n=37 cases in the group of ear acupuncture therapy and n=30 cases in the group of Chinese drugs. Pain scores were defined according to the 15-point Guideline for Clinical Research	**Results** <u>Dysmenorrhea score (mean) (max score 15):</u> • EAT group pre-treatment = 12.19 SD 2.42, post-treatment = 5.53 SD 2.17, n=37 • CD group pre-treatment = 11.22 SD 3.11, post-treatment = 10.34 SD 3.51, n=30	**Limitations** <u>Cochrane risk of bias assessment tool</u> Adequate sequence generation? Uncle ar risk (not reported) Allocation concealment? Uncl

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Ref Id 338616 **Country/ies where the study was carried out** China **Study type** Randomised, active-controlled study comparing auricular acupuncture with Chinese herbal medicine. **Aim of the study** Not stated. **Study dates** May 1997 to August 1999. **Source of funding** Financed by Administration of Traditional Chinese Medicine of Guangdong Province (97Y203).	diagnosed by peritoneoscopy and operative pathology. Baseline severity of pain: Acupuncture group: n=6 mild, n=12 moderate, n=9 severe; Herbal medicine group: n=12 mild, n=10 moderate, n=8 severe. **Inclusion criteria** • Women who met diagnostic criteria for endometriosis and the grading criteria for dysmenorrhea according to the Guideline for Clinical Research on New Chinese Medicine for Treatment of Pelvic Endometriosis, 1993. Endometriosis was confirmed by peritoneoscopy and operative pathology. **Exclusion criteria** Not reported.	Acupuncture treatment began 5 days before menstruation and was given four times every other day. **Chinese herbal medicine:** a decoction of Dan Shen Radix Salviae Miltiorrhizae, ChiShao Radix Paeoniae Rubra, San Leng Rhizoma Sparganii, E Zhu Rhizoma Curcumae, Zhi Qiao Fructus Aurantii and Xiang Fu Rhizoma Cyperi was given 5 days before menstruation; one dose for 7 days. Both therapeutic courses constituted 3 menstrual cycles.	on New Chinese Medicine for Treatment of Pelvic Endometriosis scale (Zhu et al. 2011, Acupuncture for pain in endometriosis, Cochrane Library) Dysmenorrhea scores (according to Zhu et al. 2011, Acupuncture for pain in endometriosis, Cochrane Library): Dysmenorhea symptoms: score: Pain in the lower abdomen prior to and during menstruation: 5 Unbearable abdominal pain: 1 Pronounced abdominal pain: 0.5 Restless: 1 Pass out (loss of consciousness): 2 Pale complexion: 0.5 Perspiration: 1 Cool extremities: 1 Required bed resting: 1 Interfering with daily activity: 1 No relief from common used analgesic: 1 Relief from common used analgesic: 0.5 Lower back pain: 0.5 Nausea, vomiting: 0.5	• MD = -4.81 (95%CI -6.25 to -3.37)* <u>Effect of the therapeutic effect (cure):</u> • EAT group 11/37 • CD group 3/30 • RR (95%CI) = 2.97 (0.91 to 9.70)* *calculated by the NGA 2016 team	ear risk (not reported) Blinding? High risk (not reported) Incomplete outcome data addressed? Low risk (All participants who were randomized were analysed) Free of selective reporting? Unclear risk (The outcomes of interest were not described in the Methods) Free of other bias: Unclear risk (Not reported where/how patient were enrolled) **Other information** None

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
			Distension and sore in the anus: 1 Pain within a day: 1 Pain occurs on each additional day: 0.5		
Full citation 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 **Ref Id** 338626 **Country/ies where the study was carried out** China **Study type** Prospective, randomized controlled trial.	**Sample size** Group A n=52 Grou B n=52 Group C n=52 (see Intervention) **Characteristics** The study population was infertile women with minimal or mild endometriosis confirmed by laparoscopy, according to the revised American Fertility Society (r-AFS) classification (r-AFS score < 16). All participants completed their one-month visit after surgery, where their menstrual status was noted and their recovery was ensured. Then, they were regularly followed up via the phone or outpatient visits every three months for 12 months in Group C and 14 months in complementary medical treatment Group A and B. **Inclusion criteria**	**Interventions** After the operation, the patients were randomly allocated to three groups: **Group A**: an OC (Marvelon: 30 μg ethinyl estradiol and 150 μg desogestrel/tablet) was administered one tablet continuously for 63 days, **Group B**: the OC was administered one tablet continuously for 63 days and the Dan'e mixture (manufactured by DIHON Medicine, Yunnan Province, China) was administered at 30 g/day for the latter 30 days, **Group C:** no medical treatment was given.	**Details** All patients underwent laparoscopy under general anesthesia. All apparent endometriosis lesions, including superficial endometriomas and implant lesions, were excised or cauterized by monopolar or bipolar electrocauterization. The pelvic and fallopian adhesions were detected and lysed to restore normal anatomy. The random allocation was conducted using a computer-generated list of random numbers. The codes A, B, and C were placed separately in three sealed envelopes; they were sequentially numbered and then chronologically opened in the ward only after an eligible patient was identified.	**Results** Within 12 months of follow-up: Pregancy rate n (%) • Group A = 20 (38.5%) n=52 • Group B = 16 (30.8%) n=52 • Group C = 24 (46.2%) n=52 • RR group B vs C = 0.67 (95%CI 0.40 to 1.10)* • RR group B vs A = 0.80 (95%CI 0.47 to 1.36)* Live birth n (%) • Group A = 14 (70.0%) n=52 • Group B = 13 (81.3%) n=52 • Group C = 19 (79.2%) n=52 • RR group B vs C = 1.03 (95%CI 0.75 to 1.40)* • RR group B vs A = 1.16 (95%CI 0.80 to 1.68)* Miscarriage (<28 weeks) n (%): • Group A = 20 (20.0%) n=52	**Limitations** Cochrane risk of bias assessment tool Adequate sequence generation: Low risk (Randomisation for allocation of three groups was conducted using a computer-generated list of random numbers) Allocation concealment: Low risk (Allocation sequence was concealed through numbered, sealed envelopes) Blinding: Unclear risk (It was not possible to blind participants to treatment allocation since the treatment involved the patients themselves taking medication at

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Aim of the study To compare laparoscopy alone with laparoscopy followed by treatment with oral contraceptive OCs or a combination of OCs and the Dan'e mixture in the treatment of minimal/mild endometriosis, primarily with regard to improvement of fecundity and alleviation of pelvic pain. **Study dates** February 2011 to May 2013. **Source of funding** Not reported.	• Women aged 20 to 40 years who wished to conceive and had failed to get pregnant after at least 12 months of unprotected intercourse. **Exclusion criteria** • Women were excluded if they had previously undergone medical or surgical treatments for endometriosis; if their infertility resulted from problems with the ovary, fallopian tube, or uterus, or other causes such as adenomyosis, ovarian endometrioma or deep endometriosis; or if the male partner had abnormal sperm cells or was suspected to have any gynecologic malignancies. Women with contraindications for OCs such as severe diabetes and hypertension, hepatic or renal dysfunction, and idiopathic vagina bleeding were excluded.	The patients in Group C were prepared to conceive after their one-month visit, and the patients in Group A and Group B were prepared to conceive after they experienced withdrawal bleeding at the end of medical treatment.		• Group B = 3 (81.25%) n=52 • Group C = 19 (79.16%) n=52 • RR group B vs C = 1.50 (95%CI 0.34 to 6.52)* • RR group B vs A = 0.94 (95%CI 0.24 to 3.60)* <u>Median in pelvic pain at baseline and 6 months after treatment (VAS scale from 0 to 10):</u> • Group A = baseline 38.5 (IQR 0-63), at 6 months 15 (IQR 0-46) n=52 • Group B = baseline 35 (IQR 0-82), at 6 months 19 (IQR 0-52) n=52 • Group C = baseline 28 (IQR 0-61), at 6 months 29 (IQR 0-56) n=52 *calculated by the 2016 NGA team	home and the control group received no intervention) Incomplete outcome data addressed: Unclear risk (3 patients were lost to follow-up) Free of selective reporting: Low risk (Identified outcomes adequately reported compared with the descriptions in the methods) Free of other bias:Low risk (No source of other bias) **Other information** None
Full citation de Sousa, Tatiane Regina, de Souza, Bruna Cruz, Zomkowisk, Kamilla, da Rosa, Priscila	**Sample size** GROUP A n=20 GROUP B n=22 (see Intervention) **Characteristics**	**Interventions** **Group A:** experimental treatment of acupuncture - five sessions of	**Details** Women were recruited from the Department of Pelvic Pain at the de São Thiago University Hospital, Federal	**Results** Pain scores, measured with Visual Analogue Scale (0-10)	**Limitations** <u>Cochrane risk of bias assessment tool</u> Adequate sequence

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
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 **Ref Id** 557680 **Country/ies where the study was carried out** Brazil **Study type** Prospective, randomized controlled trial. **Aim of the study** To investigate the effect of acupuncture in chronic pelvic pain, dyspareunia, and quality of life in women with endometriosis **Study dates**	Mean age (SD), years: 30.5(5.9) (GROUP A); 31.1 (6.9) (GROUP B) Mean duration of endometriosis (SD), years: 11.7 (1.3) (GROUP A); 11.7 (1.3) (GROUP B) Etnicity (%): Caucasian: 80 (GROUP A); 91 (GROUP B) Black: 20 (GROUP A); 9 (GROUP B) **Inclusion criteria** • positive diagnosis for endometriosis for at least 1 year, • age between 18 and 45 years, • waiting list to undergo a videolaparoscopy or had already undergone this procedure during the previous 3 years. • continuous use of contraceptives and the complaint of chronic pelvic pain (VAS cutoff = 4) and dyspareunia (VAS cutoff = 4) **Exclusion criteria** • fearing needles • using analgesics or anti-inflammatory drugs in the 1	acupuncture, during which 19 Dong Bang® needles were inserted (0.25 × 0.30 cm). The therapy was performed once per week, at an interval of 6–8 days. Between preparation, insertion, and needle withdrawal, the sessions lasted on average 40 min **Group B**: placebo group (sham acupuncture) - therapy consisted of placing the same number of needles and following the same time of insertion as for the EG, over a course of 5 weeks.	University of Santa Catarina. Randomization was carried out with the aid of Clinical Trials Management System (CTMS) software. The allocation sequence was performed by a laboratory assistant, and hidden to the team conducting the project and responsible for collecting the information. Survey data were collected by two previously trained researchers. A different physiotherapist specialist conducted all therapy sessions. Women were blinded as to their assigned group.	**Change (from baseline) in pain during the last 2 months,** chronic pelvic pain • Acupuncture group = -3.7 (SD 1.2)*, n = 20 • Sham group = -0.41 (SD 1.02)*, n = 22 • Mean difference = -3.29 (95% CI -3.97 to -2.61)* dyspareunia • Acupuncture group = -3.85 (SD 1.21)*, n = 20 • Sham group = -0.09 (SD 1.41)*, 22 • Mean difference = -3.76 (95% CI -4.55 to -2.97)* *Calculated by the 2016 NGA team	generation: Low risk (Randomisation for allocation of three groups was conducted using Clinical Trials Management System (CTMS) software) Allocation concealment: Low risk (The allocation sequence was performed by a laboratory assistant, and hidden to the team conducting the project and responsible for collecting the information) Blinding: unclear risk (participants were blinded to the intervention, unclear masking of outcome assessors for the measures of interest) Incomplete outcome data addressed: Unclear risk (no information given in the text to

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
December 2014 to December 2015. **Source of funding** None	month before and during data collection.				ascertain this criteria.) Free of selective reporting: Low risk (Identified outcomes adequately reported compared with the descriptions in the methods) Free of other bias:Low risk (No source of other bias) **Other information** None

1

G.17 Review question: Surgical management and combinations of treatment

3 **What is the effectiveness of pharmacological therapy before or after surgery compared with surgery alone?**

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation 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	Sample size N=144 Characteristics Infertile patients referred to private and university infertility clinics with laparoscopical and histological diagnosis of endometriosis who were infertile at least for 12	Interventions Surgery Laparoscopy was performed under general anesthesia, using a subumbilical incision and two or three lower part incisions. After evaluation of the abdomino-pelvic structures and peritoneal surface, adhesionolysis by	Details Follow up: at 3-month intervals for 1 year after restoration of menstruation cycles. Only those patients who completed their follow-up periods were included.	Results Pain recurrence at 12 months Hormonal treatment group: 5/87 No treatment group: 3/57 RR 1.09 (0.27 - 4.39) Endometriosis at 12 months	Limitations Random sequence generation (selection bias) Low risk Authors reported the use of computer-generated randomisation. Allocation concealment

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
on pregnancy rate and symptom and sign recurrence after laparoscopic treatment of endometriosis, Archives of Gynecology and Obstetrics, 284, 105-110, 2011 Ref Id 155113 Country/ies where the study was carried out Iran Study type RCT - Please note that there is an error in cataloguing and the first author in this study is Alborzi S Aim of the study To compare the role of an aromatase inhibitor (letrozole) with a GnRH agonist (triptorelin) versus no hormonal treatment following surgery on the pregnancy rate and recurrence of symptoms and signs in patients with endometriosis. Study dates	months and some of whom had symptoms such as dysmenorrhea, dyspareunia and pelvic pain. There were no statistically significant differences regarding the mean age, type of infertility, duration of infertility, prevalence of different stages of endometriosis, score of the disease and preoperative prevalence of the symptoms such as pelvic pain, dysmenorrhea, and dyspareunia among three groups. Inclusion criteria Women were entered into the study only if endometriosis was shown histologically. Exclusion criteria Those with severe male factor infertility requiring intra-cytoplasmic sperm injection (ICSI) or those who had preoperative medication were excluded	sharp dissection was done to fully mobilize the ovaries and other pelvic structures. Pharmacological treatment Group 1: women were prescribed an aromatase inhibitor, letrozole, one tablet 2.5 mg/day for 2 months Group 2: women were administered GnRH analogue, triptorelin, Amp 3.75 mg (IM) every 4 weeks, for 2 months Group 3: women did not receive any medication	At each follow up visit, the patients were asked about their symptoms and transvaginal sonography was performed. Before and after surgery each patient was asked to record the presence and severity of pelvic pain on a 10-cm linear analog scale. Recurrence of symptoms and signs was defined when dysmenorrhea, dyspareunia and pelvic pain returned. Score of 1–4: mild pain and was not included in this study because of similarities between endometriosis and non-endometriotic pain. Score of 5–7: moderate pain Score 8–10: severe pain.	Hormonal treatment group: 12/87 No treatment group: 0/57 RR 16.48 (0.99 - 272.92)	(selection bias) Unclear risk. No details reported. Blinding of participants and personnel (performance bias) All outcomes Unclear risk No placebo used Incomplete outcome data (attrition bias) All outcomes High risk 18% withdrawal overall 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) Selective reporting (reporting bias) Low risk Protocol was not available but outcomes in methods and results are similar. Other bias Low risk Authors reported that the groups were similar at baseline.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
June 2004 - January 2007 Source of funding Not reported although there were no conflicts of interest					Other information
Full citation 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 Ref Id 359851 Country/ies where the study was carried out Germany Study type RCT Aim of the study To evaluate three different treatment strategies (hormonal medication, surgical, or combined treatment) and discusses the influence of endometriosis on the	Sample size N=450 women randomised into 3 treatment groups. 2 groups of 150 women are reported here n=410 women at follow up. Characteristics Groups were similar at baseline for EEC stage. No further baseline characteristics are reported. Across groups women with different stages were EEC stage 0 n=0, EEC stage I n=185, EEC stage II n=127, EEC stage III n=85 Inclusion criteria Women with symptomatic endometriosis (18-44 years old) in whom 2 consecutive laparoscopic interventions were to be assessed. Exclusion criteria Previous surgery or hormone therapy for endometriosis was exclusion criterion, as was	Interventions Surgery: Laparoscopic excision of endometrial foci, removal of adhesions and restoration of normal reproductive anatomy. Ureter and superficial bowel lesions were removed. For infertility patients, tubal patency was checked and chromoperturbation was performed at the second-look laparoscopy Pharmacological comparison: Leuprorelin depot subcutaneously injected monthly over a 3 month period with subsequent second-look laparoscopy 1-2 months after conclusion of the hormonal therapy or no treatment with subsequent second-look laparoscopy at 5-6 months post-surgery.	Details The same team of physicians performed the primary and secondary intervention For women receiving leuprorelin, a second-look laparoscopy was performed 1-2 months after hormonal therapy and, for women receiving no hormonal therapy, 5 to 6 months after surgical endometriosis treatment. After the second-look laparoscopy, patients were monitored over a period of 2 years and completed an extensive questionnaire to determine their recurrence of symptoms, new endometriotic lesions determined laparoscopically, and	Results Pain recurrence (questionnaire based) at 12 months post treatment completion Abdominal pain Leuprorelin group: 25/62 No treatment group: 33/58 RR 0.71 (0.49 - 1.03) Dysmenorrhoea Leuprorelin group: 24/80 No treatment group: 27/78 RR 0.87 (0.55 - 1.36) Dyspareunia Leuprorelin group: 12/75 No treatment group: 21/69 RR 0.53 (0.28 - 0.99) Disease recurrence at 5-6 months Leuprorelin group: 59/148	Limitations Random sequence generation (selection bias) Unclear risk Not described although a flow chart is presented and the authors state that "All patients were allocated exactly according to the random principle" and ethics committee approval was given Allocation concealment (selection bias) Unclear risk Not described although a flow chart is presented and the authors state that "All patients were allocated exactly according to the random principle" and eth Blinding of participants and

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
cure of this disease and pain relief. Study dates Not reported Source of funding Not reported although there were no conflicts of interest	deep infiltrating endometriosis with bladder or rectum excision.		confirmed pregnancy rates.	No treatment group: 55/137 RR 0.99 (0.75 - 1.32)	personnel (performance bias) All outcomes Unclear risk No placebo used Incomplete outcome data (attrition bias) Pain outcomes Unclear risk 40/450 women were lost to follow up. 13 were in the surgery only group and 2 were in the combined treatment group.9 more women in the surgery only group declined to participate and 2 more were lost to follow up compared to the combined group Selective reporting (reporting bias) Low risk Protocol was not available but outcomes in methods and results are similar. Other bias Low risk Authors only report that the groups were similar at baseline for EEC staging Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation 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 Ref Id 346669 Country/ies where the study was carried out Canada Study type Cochrane systematic review Aim of the study To determine if the levonorgestrel-releasing intrauterine device (LNG-IUD), also known as the levonorgestrel intrauterine system (LNG-IUS), improves pain symptoms associated with menstruation and reduces recurrence of endometriosis when	Sample size N= 3 RCTs of which 2 are relevant (Tanmahasamut 2012 and Vercellini 2003) Characteristics Trials comparing insertion of the LNG-IUD versus no postoperative treatment, placebo (inert IUD), or any other active systemic treatment in women undergoing surgery for endometriosis. Inclusion criteria Trials were included if they compared women undergoing surgical treatment for endometriosis with uterine preservation and then randomised within three months to LNG-IUD insertion versus no postoperative treatment, placebo (inert IUD), or other treatment. Tanmahasamut 2012 Participants: Women (n=55) with moderate to severe dysmenorrhea, chronic pelvic pain, or both for more than 6 months and who were scheduled for laparoscopic surgery. Using ASRM staging. 10	Interventions Tanmahasamut 2012 Randomisation to immediate LNG-IUD insertion or no postoperative treatment (expectant management) after laparoscopic treatment of endometriotic lesions. Vercellini 2003 Randomisation to immediate LNG-IUD insertion or no postoperative treatment (expectant management) after laparoscopic treatment of endometriotic lesions.	Details Tanmahasamut 2012 Design: double-blind, parallel-group, randomised controlled trial Follow-up: 12 months Setting: Single centre Gynecologic Endocrinology Unit (University setting). Vercellini 2003 Design: open-label, parallel-group, randomised controlled trial. Follow-up: 12 months Setting: a tertiary care and referral centre for women with endometriosis.	Results Tanmahasamut 2012 Dysmenorrhea recurrence at 12 m LNG-IUD group: 2/28 No treatment: 9/27 RR 0.21 (0.05 - 0.90) Patient satisfaction at 12 m log RR: 0.193125 SE 0.24634 RR 1.21 (0.75 - 1.97) Vercellini 2003 Dysmenorrhea recurrence at 12 m LNG-IUD group: 2/20 No treatment: 9/20 RR 0.22 (0.05 - 0.90) Patient satisfaction at 12 m log RR: 0.176091 SE 0.39188 RR 1.19 (0.55 - 2.57)	Limitations Abou Setta 2013 AMSTAR 9/11 Low risk of bias Tanmahasamut 2012:Risk of bias Random sequence generation (selection bias) Low risk Authors reported the use of computer-generated randomisation sequence. Allocation concealment (selection bias) Low risk Authors reported that "the codes were individually contained in a sealed opaque envelope, which was sequentially numbered and then chronologically opened in the operating room only after an eligible patient was identified". Blinding of participants and personnel

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
inserted postoperatively in women undergoing surgery for endometriosis. The LNG-IUD was to be compared with no postoperative treatment, postoperative placebo (inert IUD), or postoperative systemic treatment. Study dates Updated Issue 1 Cochrane Library 2013 Source of funding None	women stage 1, 7 women stage 2, 8 women stage 3 and 29 women stage 4 Vercellini 2003 Participants: Parous women (n=40) with moderate to severe dysmenorrhea undergoing first-line operative laparoscopy for symptomatic endometriosis. Women were AFS stages I - IV Exclusion criteria The use of diagnostic laparoscopy alone was not considered suitable treatment for trials to be included into the systematic review.				(performance bias) All outcomes Unclear risk Authors reported that "the patients and assessor nurse were blinded to the treatment groups" but not clear how patients were prevented from physically feeling the vaginally placed IUD strings. Blinding of outcome assessment (detection bias) All outcomes Low risk Authors reported that "the patients and assessor nurse were blinded to the treatment groups". Incomplete outcome data (attrition bias) All outcomes Low risk Authors reported that one patient in the LNG-IUD group was lost to follow-up as compared with three in the control group. Also one patient was removed from the study due to a protocol violation. The authors analysed all

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					the randomised patients with the exception of the patient with the protocol violation (e.g. 54/55) using last evaluation carried forward method. Selective reporting (reporting bias) Low risk Protocol was not available but outcomes in methods and results are similar. Other bias Low risk Authors reported that "the two groups were comparable in age, weight, body mass index, obstetric history, and baseline pain scores" and provided statistical evidence of similarity. Vercellini 2003: Risk of bias Random sequence generation (selection bias) Low risk Authors reported the use of computer-generated randomisation

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					sequence. Allocation concealment (selection bias) Low risk Authors reported using serially numbered, opaque, sealed envelopes. Blinding of participants and personnel (performance bias) All outcomes High risk Reported as open-label study (i.e. no blinding of participants and personnel). Blinding of outcome assessment (detection bias) All outcomes High risk Reported as open-label study (i.e. no blinding of outcome assessors). Incomplete outcome data (attrition bias) All outcomes Low risk Authors reported that "In one patient the LNG-IUD was expelled after five months. One subject

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					in each group was lost to follow-up". Intention-to-treat analysis used for all analyses. Selective reporting (reporting bias) Low risk Protocol was not available, but outcomes described in the methods section and results section match. Other bias Unclear risk The authors reported that "the distribution of the study variables was similar in both groups" without providing any statistical support. No other biases were evident from the trial report Other information Tanmahasamut 2012: Authors reported that the trial was "supported by the research fund of the Gynecologic Endocrinology Unit, Faculty of Medicine

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Siriraj Hospital, Mahidol University, Thailand" and that "Bayer Schering Pharma Company provided the levonorgestrel-releasing intrauterine system"
Full citation 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 Ref Id 338558 Country/ies where the study was carried out Italy Study type RCT Aim of the study To evaluate long-term cyclic and continuous	Sample size N=239 Characteristics Similar across groups at baseline for age, AFS stage (AFS stage III n=99 and AFS stage IV n=118), mean cyst diameter, incidence of bilateral cysts, associated implants. associated adhesions, length of follow up (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 fo endometriosis and no receipt of oral contraceptives for at least 6 months prior to surgery. Exclusion criteria	Interventions Surgery: Laparoscopic excision of ovarian endometriomas using the classic stripping technique. Pharmacological comparison: Group 1: no pharmacological treatment for 24 months Group 2: low dose monophasic oral contraceptives cyclic therapy (daily for 21 days followed by a 7 day interval) for 24 months Group 3: continuous low dose monophasic oral contraceptives for 24 months	Details Women were randomised into 3 treatment groups after surgery which started on the day of discharge and continued for 24 months. All women underwent clinical and TV US examination every 6months to assess possible endometrioma recurrence. Recurrence was defined as the presence of a cyst with a minimum diameter of 1.5cm with a typical aspect detected by TV US. All scans were performed by experiences operators who were blinded to study	Results Endometrioma recurrence at 12 months post treatment completion (24 months) OC group (continuous and cyclic): 17/148 No treatment group: 20/69 RR 0.40 (0.22 - 0.71)	Limitations Random sequence generation (selection bias) Low risk Computer generated randomisation Allocation concealment (selection bias) Low risk Opaque sealed envelopes used Blinding of participants and personnel (performance bias) Unclear risk No placebo used although outcome assessors were blinded to treatment group Incomplete outcome data (attrition bias) Low risk 22/239 women were lost to follow up. 10

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
administration of oral contraceptive pills (OCP) in preventing ovarian endometrioma recurrence after laparoscopic cystectomy. Study dates Not reported Source of funding Not reported	Women who refused to be randomised to treatment were excluded from the study from outset. Patients having contraindications to OC therapy, unwillingness to tolerate the absence of menstruation, or the lack of desire to postpone pregnancy for at least 2 years after surgery.		allocation.2 months after detection of a recurrent cyst, additional US examination was performed to confirm the diagnosis.		were in the no treatment group (4 became pregnant and 6 received OCs for dysmenhorroea) and 12 were in the OC groups (4 for reasons unrelated to the study and 8 for side effects related to OC use) Selective reporting (reporting bias) Low risk Protocol was not available but outcomes in methods and results are similar. Other bias Low risk Authors reported that the groups were similar at baseline Other information
Full citation Furness,Susan, Yap,Christine, Farquhar,Cindy, Cheong,Ying C., Pre and post-operative medical therapy for endometriosis surgery, Cochrane Database of Systematic Reviews, - , 2011	Sample size N=16 trials examining 4 comparisons. One comparison is relevant here and eight trials included outcomes relevant to this protocol Characteristics Trials were included if they were randomised controlled trials comparing medical therapies for	Interventions Medical hormonal suppression therapies used post-surgery for endometriosis compared with surgery alone or surgery and placebo. Bianchi 1999 Post-surgical medical therapy 1. Danazol oral 600 mg	Details Bianchi 1999 No. of centres: 1 Location: University of Milan, Italy Recruitment period: July 1994 to October 1996 Busacca 2001 Location: University of Milan, Italy No. of centres: 1	Results Bianchi 1999 Pain recurrence <=12 months Hormonal treatment group: 7/31 Control group: 9/29 RR 0.73 [0.31, 1.70] Disease recurrence at 12 months Hormonal treatment group: 3/36	Limitations Furness 2011 AMSTAR 9/11 Low risk of bias Bianchi 1999 Random sequence generation (selection bias) Low risk "Randomization was done according to a computer generated list"

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Ref Id 106969 Country/ies where the study was carried out UK Study type Cochrane systematic review Aim of the study To determine the effectiveness of medical therapies for hormonal suppression before or after surgery for endometriosis for improving painful symptoms, reducing disease recurrence and increasing pregnancy rates. Study dates Updated in Issue 10 Cochrane Library 2011 Source of funding Singhealth Research, Singapore General Hospital (internal source of support). No external sources of support	hormonal suppression before or after or before and after, surgery for endometriosis. All randomised controlled trials of the use of medical hormonal suppression therapies used: •pre-surgery for endometriosis compared with surgery alone or placebo prior to surgery for the treatment of endometriosis; •post-surgery for endometriosis compared with surgery alone or surgery and placebo; •pre and post-surgery for endometriosis compared with surgery alone or surgery and placebo; •pre-surgery for endometriosis compared with medical therapies used post-surgery for endometriosis. The highlighted comparison is the comparison of interest in this review. Studies included in the remaining 3 comparisons were excluded (See excluded studies table) Inclusion criteria	daily x 3/12 (n = 36) 2. No treatment (n = 41) Busacca 2001 Post-surgical medical therapy Gr A (n=44): leuprolide acetate SC 3.5 mg 4 weekly x 3 doses Gr B (n=45): no treatment Loverro 2008 Post-operative triptorelin versus placebo Gr A (n=29): triptorelin 3.75 mg depot monthly on day 20 of cycle for 3 months Gr B (n=25): placebo monthly on day 20 of cycle for 3 months Muzii 2000 Post-surgical medical therapy Gr A (n=35): cyclic monophasic oral contraceptive pill (ethinyl estradiol 0.03 mg, gestodene 0.075 mg) for 21 days with 7 pill free days x 6/12 Gr B (n=35): no treatment Parazzini 1994 Post-surgical medical therapy Gr A (n=36): nafarelin nasal 400 µg daily x 3/12 Gr B (n=39): placebo	Recruitment period: July 1997 to December 1999 Loverro 2008 Location: Italy No. of centres: one Recruitment period: January 1998 to January 1999 Muzii 2000 Location: University departments, Rome, Italy No. of centres: 2 Recruitment period: January 1994 to June 1997 Parazzini 1994 Location: University centres in Italy No. of centres: 6 Recruitment period: January 1990 to July 1991 Sesti 2007 Location: Rome, Italy No. of centres: one Recruitment period: January 1999 to May 2005 Tsai 2004 Location: Taiwan No. of centres: one Recruitment period: June 1988 to December 2001	Control group: 6/41 RR 0.57 [0.15, 2.11] Reoperation* Hormonal treatment group: 0/31 Control group: 1/29 RR 0.31 [0.01, 7.38] Busacca 2001 Pain recurrence 13-24 months Hormonal treatment group: 10/44 Control group: 11/45 RR 0.93 [0.44, 1.97] Disease recurrence at 12 months Hormonal treatment group: 4/44 Control group: 4/45 RR 1.02 [0.27, 3.84] Reoperation* Hormonal treatment group: 2/44 Control group: 0/45 RR 5.11 [0.25, 103.53] Loverro 2008 Pain recurrence <=12 months Hormonal treatment group: 15/33 Control group: 13/29 RR 1.01 [0.58, 1.76] Pain recurrence at 5 years Hormonal treatment group: 13/29	Allocation concealment (selection bias) Unclear risk not mentioned Blinding (performance bias and detection bias) All outcomes High risk not mentioned, no placebo Incomplete outcome data (attrition bias) All outcomes Low risk all randomised patients included in analysis Selective reporting (reporting bias) Low risk important outcomes - recurrence of endometriosis pain, Other bias Low risk groups appear comparable at baseline Busacca 2001 Random sequence generation (selection bias) Low risk "randomization was performed according to a computer generated list unknown to the physicians" Allocation

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Furness 2011: The study population included women of reproductive age who were undergoing surgery for endometriosis. The diagnosis of endometriosis could have been made provisionally by clinical examination and confirmed during the surgery, or could have been confirmed endometriosis where women were undergoing second or subsequent surgery. They would have further medical treatment either before or after surgery. Studies in the hospital care setting were considered. Bianchi 1999 Inclusion criteria: < 40 yrs No. randomised: 77 No. analysed: 77 Busacca 2001 Inclusion criteria: < 40 yrs, laparoscopic diagnosis of endometriosis stage III-IV No. randomised: 89 No. analysed: 89 Loverro 2008 Inclusion criteria: women of reproductive age with stage III - IV endometriosis, associated with chronic pelvic pain,adnexial mass or	Sesti 2007 Gr A (n=115): placebo for 6 months Gr B (n=119): post-operative medical or dietary therapy. Patients received either triptorelin or leuprorelin 3.75 mg depot monthly for 6 months (n=42), continuous low dose monophasic oral contraceptives for 6 months, (ethinlyestradiol 0.03 mg + gestoden 0.75 mg) (n=40) or (not included here) dietary therapy for 6 months (vitamins, mineral salts, lactic ferments and omega 3 and omega 6 fatty acids together with individually tailored diet) (n=37) Tsai 2004 Post-operative medical therapy (either danazol or GNRH analogue) Gr A (n=15): either 3 months 400 mg danazol orally, twice daily for 3 months or 3.75 mg leuprolide acetate depot SC every 28 days for 3 months Gr B (n= 30): no post-operative medical treatment	Vercellini 1999 Location: Italy No. of centres: 19 Recruitment period: February 1992 to June 1994	Control group: 12/25 RR 0.93 [0.53, 1.66] Disease recurrence at 5 years Hormonal treatment group: 4/19 Control group: 2/16 RR 1.68 [0.35, 8.03] Muzii 2000 Pain recurrence 13-24 months Hormonal treatment group: 3/33 Control group: 6/35 RR 0.53 [0.14, 1.95] Endometrioma recurrence at 13-36 months* Hormonal treatment group: 2/33 Control group: 1/35 RR 2.12 [0.20, 22.31] Parazzini 1994 Pelvic pain at 12 months* Hormonal treatment group: Mean 3.6 SD 2.9 N=24 Control group: Mean 4.0 SD 3.6 N=29 MD -0.40 [-2.15, 1.35] Sesti 2007 Pelvic Pain at 12 months (VAS) Hormonal treatment	concealment (selection bias) Unclear risk not described Blinding (performance bias and detection bias) All outcomes High risk not mentioned, no placebo Incomplete outcome data (attrition bias) All outcomes Low risk all randomised patients included in the analysis Selective reporting (reporting bias) Low risk important outcomes of recurrence of endometriosis and pain reported Other bias Low risk groups appear comparable at baseline Loverro 2008 Random sequence generation (selection bias) Low risk "using a computer generated randomization table" Allocation concealment (selection bias) Unclear risk

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	infertility, who had undergone complete laparoscopic excision, had rAFS score > 15 and no previous hormonal treatment No. randomised: 60 No. analysed: 54 Muzii 2000 Inclusion criteria: 20-35 yrs, moderate to severe dysmenorrhoea and/or chronic pelvic pain, not desiring fertility No. randomised: 70 No. analysed: 68 Parazzini 1994 Inclusion criteria: age < 38 yrs, normal medical examination, unexplained infertility for at least 1 year, with/without chronic pelvic pain, endometriosis stage III-IV, partners with normal sperm analysis and post-coital tests No. randomised: 75 No. analysed: 75 (pregnancy rates), 68 (pain scores) Sesti 2007 Inclusion criteria: women of reproductive age <40, with endometriosis related symptoms (dysmenorrhoea, pelvic pain, deep dyspareunia), laparoscopic diagnosis of	Vercellini 1999 Post-surgical medical therapy Gr A (n= 133): goserelin SC 3.6 mg every 4 weeks x 6 months Gr B (n=134): no treatment		group: Mean 5.0 SD 0.95 N=77 Control group: Mean 6.2 SD 0.9 N=110 MD -1.20 [-1.47, -0.93] Dysmenhorroea at 12 months (VAS) Hormonal treatment group: Mean 5.7 SD 1.07 N= 77 Control group: Mean 6.4 SD 1.3 N=110 MD -0.70 [-1.04, -0.36] Dyspareunia at 12 months (VAS) Hormonal treatment group: Mean 4.4 SD 1.25 N=77 Control group: Mean 4.8 SD 1.2 N=110 MD -0.40 [-0.76, -0.04] Short form 36 general health survey:* Improvement of scores in all domains at 12 months in both treatment and control groups Tsai 2004 Disease recurrence	not mentioned Blinding (performance bias and detection bias) All outcomes Low risk patients were blinded to treatment allocation. placebo injections used Incomplete outcome data (attrition bias) All outcomes Unclear risk 1 and 5 patients lost to follow up from triptorelin and no treatment groups respectively. Possibility of bias Selective reporting (reporting bias) Low risk pain, relapse and pregnancy reported (for those who desired pregnancy) Other bias Low risk groups appear similar at baseline Muzii 2000 Random sequence generation (selection bias) Low risk "randomly allocated to one of two management arms on the basis of a computer generated sequence" Allocation

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	St III -IV endometriosis, desiring pregnancy, nulliparous No. randomised: 234 No. analysed: 222 Tsai 2004 Inclusion criteria: women of reproductive age with infertility and stage III or IV endometriosis planning to undergo controlled ovarian hyperstimulation and intrauterine insemination or in vitro fertilisation and embryo transfer. All had surgery for endometriosis - either laparotomy or laparoscopy for cystectomy, adhesiolysis, ablation of endometriosis No. randomised: 45 No. analysed: 41 Vercellini 1999 Inclusion criteria: pre-menopausal, endometriosis score >/= 4 points, chronic pelvic pain No. randomised: 269 No. analysed: 210 Exclusion criteria Bianchi 1999 Exclusion criteria: medical or surgical treatment for endometriosis, concurrent disease that might affect fertility or cause pelvic pain, women without pain			at 24 months Hormonal treatment group: 0/15 Control group: 4/30 RR 0.22 [0.01, 3.75] Vercellini 1999 Pain recurrence <=12 months Hormonal treatment group: 14/107 Control group: 22/103 RR 0.61 [0.33, 1.13] Pain recurrence 13-24 months Hormonal treatment group: 3/33 Control group: 6/35 RR 0.53 [0.14, 1.95] *additional outcomes reported in the full text of the paper but not in the Furness review	concealment (selection bias) Unclear risk not described Blinding (performance bias and detection bias) All outcomes High risk not mentioned, no placebo Incomplete outcome data (attrition bias) All outcomes Low risk two post-randomisation withdrawals. Unlikely to have introduced a bias Selective reporting (reporting bias) Low risk important outcomes reported - recurrence of endometriosis, pain, AFS scores. Patients not desiring pregnancy Other bias Unclear risk no information of the baseline characteristics of the groups reported Parazzini 1994 Random sequence generation (selection bias) Low risk "computer generated

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	symptoms, women not seeking pregnancy, liver or endocrine disease Busacca 2001 Exclusion criteria: previous medical or surgical therapy for endometriosis, other diseases that might affect fertility or cause pelvic pain; liver, endocrine or neoplastic disease Loverro 2008 Exclusion criteria: NS Muzii 2000 Exclusion criteria: treatment for endometriosis in previous 6 months Parazzini 1994 Exclusion criteria: previous laparoscopic/clinical diagnosis of endometriosis, other diseases that might cause infertility or pelvic pain, previous treatment for endometriosis or infertility Sesti 2007 Exclusion criteria: concurrent disease, such as cancer or pelvic inflammatory disease, previous surgery for endometriosis, contraindications to estrogens/progestins				randomization list" Allocation concealment (selection bias) Low risk assigned by telephone call 7 days from surgery Blinding (performance bias and detection bias) All outcomes Low risk double blind but authors acknowledge that adverse effects of treatment make maintaining blinding difficult Incomplete outcome data (attrition bias) All outcomes Low risk no losses to follow up, all randomised patients included in analyses Selective reporting (reporting bias) Low risk pregnancy rate and pelvic pain reported Other bias Low risk groups appear comparable at baseline Sesti 2007 Random sequence generation (selection bias) Low risk

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Tsai 2004 Exclusion criteria: NS Vercellini 1999 Exclusion criteria: NS				"randomized according to a computer generated randomization sequence" Allocation concealment (selection bias) Low risk allocated by serially numbered opaque sealed envelopes Blinding (performance bias and detection bias) All outcomes Unclear risk "neither the surgeons not the patients were aware of the regimen prescribed during the study period". However placebo not described and it seems unlikely that blinding of patients could be maintained when treatments are either SC, oral medication or diet plus supplements Incomplete outcome data (attrition bias) All outcomes Unclear risk 5 and 3 lost to follow up from placebo and GNRHa groups and reasons given. 2 lost to follow up from each

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					of OCP and diet groups but reasons not given. 222 evaluated Selective reporting (reporting bias) Unclear risk pain and health related quality of life reported. No pregnancy outcome in a group of women desiring pregnancy Other bias Low risk groups appear comparable at baseline Tsai 2004 Random sequence generation (selection bias) Low risk "simple randomisation with a computer generated list unknown to physicians" Allocation concealment (selection bias) Low risk list "unknown to physicians" Blinding (performance bias and detection bias) All outcomes High risk not mentioned, no placebo

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Incomplete outcome data (attrition bias) All outcomes High risk 4 lost to follow up from Gr A (27%) Selective reporting (reporting bias) Low risk pregnancy and recurrence reported Other bias Unclear risk 13 years of recruitment - ? associated changes in surgical techniques over this time Vercellini 1999 Random sequence generation (selection bias) Low risk "randomised in a proportion of 1:1 ... in accordance with a computer-generated randomisation sequence" Allocation concealment (selection bias) Low risk centralised randomisation, allocation obtained by phone call Blinding (performance bias and detection bias) All outcomes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					High risk not mentioned, no placebo Incomplete outcome data (attrition bias) All outcomes Unclear risk 269 patients randomised, 2 excluded because case record forms not completed, 26 & 31patients (22%) withdrew from treatment and control groups respectively for reasons other than symptom recurrence or were excluded due to major protocol violations. Reasons for exclusion similar in each group- may have introduced bias Selective reporting (reporting bias) Low risk important outcomes of recurrence, dysmenorrhoea and pregnancy reported Other bias Low risk groups appear comparable at baseline Other information
Full citation	Sample size	Interventions	Details	Results	Limitations

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
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 Ref Id 338560 Country/ies where the study was carried out Study type RCT Aim of the study To assess the recurrence rate of endometrioma after laparoscopic cystectomy plus hormal suppression treatment or plus dietary therapy	N=259 N=240/259 completed the study Characteristics Across groups, women were similar at baseline for age, disease stage, uni/bilateral ovarian endometriosis, diameter of endometrioma, presence of uterine myoma, non-menstrual pain, deep dyspareunia. Significantly fewer women in the GNRH-a group had dysmenorrhoea compared to the placebo, estroprogestin (and dietary) groups14/58 vs 33/60, 32/60 (and 30/62) respectively p=0.003 Inclusion criteria Reproductive age, up to 40 years at time of surgery, US evidence of endometrioma, moderate to severe endometriosis-related painful symptoms (=>4 on 10 point VAS), laparoscopic diagnosis of endometrioma staged by AFS classification, first laparoscopic surgery for endometriosis and conservative treatment with retention of the uterus and ovaries, complete	Surgery: 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	Seven days after laparoscopic cystectomy surgery for endometrioma, 259 consecutive women were randomly allocated to one of four post-operative management arms (placebo (n=65) or gonadotrophin-releasing hormone analogue (tryptorelin or leuprorelin, 3.75 mg every 28 days) (n=65) or continuous low-dose monophasic oral contraceptives (ethynilestradiol, 0.03 mg plus gestoden, 0.75 mg) (n=64) or dietary therapy (not reported here) (n=65)) for 6 months. At 18 months' follow-up after surgery, all patients were monitored with a clinical gynaecologic examination, and a transvaginal ultrasonography for possible evidence of endometrioma recurrence. Recurrence was defined as the presence of a cyst, detected by TVUS	Reoperation Hormonal treatment group: 6/118 Control group: 3/60 RR 1.02 [0.26, 3.93] Endometrioma recurrence at 13-36 months Hormonal treatment group: 15/118 Control group: 10/60 RR 0.76 [0.36, 1.59]	Random sequence generation (selection bias) Low risk Computer generated randomisation Allocation concealment (selection bias) Low risk Opaque envelopes used Blinding (performance bias and detection bias) All outcomes Low risk placebo used Incomplete outcome data (attrition bias) All outcomes Low risk 240/259 women who underwent surgical laparoscopy completed the study Selective reporting (reporting bias) Low risk important outcomes - reported Other bias Low risk groups appear comparable at baseline Other information

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
compared to post-operative placebo Study dates Jan 2004 – Aug 2006 Source of funding Not reported	excision of all evidnece peritoneal and ovarian disease, US and clinical follow-up after surgery. No women were attempting to conceive at the time of study entry. Exclusion criteria Women who received 6 months estrogen-suppressing drugs before first surgery, usual contradictions to estrogens and progestins, previous surgical treatment for endometriosis, surgical findings of concomitant deeply infiltrating endometriosis		with a pattern suggesting an endometrioma of more than 20mm in diameter		

What is the effectiveness of the following treatments for endometriosis, including recurrent and asymptomatic endometriosis: hysterectomy, with or without oophorectomy?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Full citation 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, 2008	Sample size N=240 n=120 in hysterectomy group (selected from the clinic) n=120 in laparoscopy group	Interventions Hysterectomy with or without bilateral oophorectomy. Laparascopic excision of endometriotic lesions.	Details Identification of participants Participants identified through electronic medical records for women who had undergone gynaecological surgery at the	Results Health related quality of life Not reported Rate of success (disease recurrence and subsequent re-operation rate) Re-operation	Limitations CASP checklist for cohort studies 1. Did the study address a clearly focussed issue? (Issue could be in terms of population, risk factors, outcomes considered, is it clear if the study clearly tried to detect a beneficial or harmful effect?) Yes/Unclear/No: yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
Ref Id 370275 Country/ies where the study was carried out USA Study type Retrospective cohort study. Aim of the study To investigate the need for further surgery after laparascopic excision of endometriosis or hysterectomy. Study dates January 1995 to December 2003 Source of funding Not reported	Hysterectomy divided into two subgroups: Group 1: Hysterectomy with ovarian preservation (at least one ovary preserved), n=47 Group 2: Hysterectomy without ovarian preservation (both ovaries removed), n=50 Characteristics Surgery age (years, n) 19-29: hysterectomy=5; laparoscopy=36 30-39: hysterectomy=43; laparoscopy=50 40 and older: hysterectomy=49; laparoscopy=23 Race (n) Other: hysterectomy=22; laparoscopy=15 White: hysterectomy=75; laparoscopy=94		clinic with diagnosis of endometriosis. Following surgery, women were contacted by post about the study and how to participate via telephone survey (questionnaire about any re-operation, pain clinic visit, medical treatment, level of satisfaction). Follow-up information was obtained from computerised medical records (operative reports, pathology reports, outpatient charts, telephone survey). A second letter was sent to those women who were not contactable in the first round. Index surgery defined as first surgery performed at the Cleveland clinic for pelvic pain.	Hysterectomy without oophorectomy group: 9/47 required further surgery Hysterectomy with oophorectomy group: 4/50 required further surgery Hazards ratios within the hysterectomy subgroups and ovarian preservation on re-operation-free survival Hysterectomy with bilateral oophorectomy: Reference 1.00 Hysterectomy with unilateral oophorectomy: HR 2.53 (95%CI 0.63-10.11) Hysterectomy without oophorectomy: HR 2.44 (95%CI 0.65-9.10)	2. Was the cohort recruited in an acceptable way? HINT: Look for selection bias which might compromise the generalisibility of the findings: Was the cohort representative of a defined population? yes, but from medical records Was there something special about the cohort? only women who had surgery for chronic pelvic pain with histological confirmation of endometriosis were included. Was everybody included who should have been included? yes Yes/Unclear/No: Yes Risk of bias: Low 3. Was the exposure measured accurately to minimise bias? HINT: Look for measurement or classification bias: Did they use subjective or objective measurements? The telephone survey may have been subjective, as it consisted of a survey/questionnaire about reoperation, pain clinic visits, medical treatments, and level of satisfaction (recall by patients). Scales were not used to address these issues. Do the measurements truly reflect what you want them to (have they been validated)? Yes/unclear/No: Unclear. Although standardised approaches were used for surgical techniques, it is not apparent

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Disease stage (n) Stage I: hysterectomy=16; laparoscopy=16 Stage II: hysterectomy=28; laparoscopy=35 Stage III: hysterectomy=21; laparoscopy=12 Stage IV: hysterectomy=32; laparoscopy=46 Ovary involvement (n) No: hysterectomy=48; laparoscopy=36 Yes: hysterectomy=49; laparoscopy=73 Ovary preservation (n) No: hysterectomy=50; laparoscopy=2 Yes: hysterectomy=47; laparoscopy=107 Re-intervention (n) None: hysterectomy=82; laparoscopy=43		Previous surgery defined as procedure before the index surgery. Surgery was performed only if medical management with GnRH agonists or other medical suppressive therapies were refused or failed to control symptoms. Recurrence was defined as pelvic pain necessitating further surgical treatment. Time to recurrence was measured as the time (years) from index surgery until additional surgery. For time to re-operation, survival methods were used, estimates of re-operation free survival at 2, 5 and 7 years were calculated using Kaplan-Meier methods and log-rank tests. Estimates of risk (HR) were	Pain relief Not reported Unintended effects from treatment Not reported Participant satisfaction with treatment Not reported	how well the surgeon performed the surgery, and authors did not report any scales used to assess level of pain experienced by the patients. 4. Were all the subjects classified into exposure groups using the same procedure Yes/Unclear/No: No. The exposure group was selected from electronic medical records, those who had gynaecological surgery. The comparator group was randomly selected from electronic records. 5. Was the outcome measured accurately to minimise bias? HINT: Look for measurement or classification bias: Did they use subjective or objective measurements? Subjective (recurrence of pelvic pain requiring re-operation) Do the measures truly reflect what you want them to (have they been validated)? Unclear Has a reliable system been established for detecting all the cases (for measuring disease occurrence)? Yes Were the measurement methods similar in the different groups? Yes Were the subjects and/or the outcome assessor blinded to exposure (does this matter)? No. The assessors/subjects were not blinded to exposure due to the type of intervention. Yes/Unclear/No: Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	Re-operation: hysterectomy=13; laparoscopy=62 Pain clinic: hysterectomy=2; laparoscopy=4 Prior surgeries (n) None: hysterectomy=47; laparoscopy=48 1-2 surgeries: hysterectomy=30; laparoscopy=48 3 or more surgeries: hysterectomy=20; laparoscopy=13 Inclusion criteria Diagnosis of endometriosis Women who underwent surgery for chronic pelvic pain with histological confirmation of endometriosis Exclusion criteria Women who underwent surgery for infertility or menorrhagia as the primary indication		computed using Cox proportional hazards methods. A significance level of 0.05 was assumed for all tests. Sample size: allowed for 90% power to detect decrease in 3 year re-operation rate of 60% in the hysterectomy group as compared with the laparoscopic group if the historical rate of 3-year re-operation rate of 25% was observed in the laparascopic group. Sample size calculations were based on log-rank test with significance of 0.05.		Risk of bias: Medium 6. Have authors identified all important confounding factors? List the ones that you think may be important, that the authors have missed Yes/unclear/No: Yes 7. Have the authors taken account of confounding factors in the design and/or analyses? HINT: Look for restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors Yes/Unclear/No: Yes. Cox proportional hazards models were performed. 8. Was the follow up of subject complete enough? Yes/Unclear/No: Yes 9. Was the follow up of subjects long enough? HINT: Consider The good or bad effects should have had long enough to reveal themselves The persons that are lost to follow-up may have different outcomes than those available for assessment In an open or dynamic cohort, was there anything special about the outcome of the people leaving, or the exposure of the people entering the cohort?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Yes/Unclear/No: Yes Risk of bias: low 10. What are the results of this study? HINT: Consider What are the bottom line results? Have they reported the rate or the proportion between the exposed/unexposed, the ratio/the rate difference? The authors report hazard ratios between hysterectomy plus oophorectomy and hysterectomy without oophorectomy. Hysterectomy+ bilateral oophorectomy: Reference: 1.00; hysterectomy only: HR 2.44 (95%CI 0.65-9.10) How strong is the association between exposure and outcome? Preservation of both ovaries increased the risk of re-operation by 2.44 times (regardless of age), but the result did not reach statistical significance (P=0.18). What is the absolute risk (AR)? N/A 11. How precise are the results? HINT: Look for the range of the confidence intervals, if given. The results are not precise as the confidence intervals are wide. 12. Do you believe the results? HINT: Consider Big effect is hard to ignore! Can it be due to bias, chance or confounding?

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Are the design and methods of this study sufficiently flawed to make the results unreliable? Bradford Hills criteria (e.g. time sequence, dose-response gradient, biological plausibility, consistency) The results do reflect what is expected to happen, that there would be fewer re-operation events for women who have hysterectomy+oophorectomy as ovaries are removed. Although the result is clinically important, the result is not significant, which could be due to the small sample size of the population. Yes/unclear/no: Unclear Risk of bias: medium 13. Can the results be applied to the local population? HINT: Consider whether A cohort study was the appropriate method to answer this question The subjects covered in this study could be sufficiently different from your population to cause concern Your local setting is likely to differ much from that of the study You can quantify the local benefits and harms Yes/unclear/no: Unclear. The result shows clinical benefit for hysterectomy+oophorectomy, but as the results are not statistically significant.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					14. Do the results of this study fit with other available evidence? Yes/unclear/no: Unclear (no other sources of evidence identified) 15. What are the implications of this study for practice? HINT: Consider One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making For certain questions observational studies provide the only evidence Recommendations from observational studies are always stronger when supported by other evidence The direction of effect of re-operation favours women who have hysterectomy and oophorectomy over 7 years but there is imprecision around the estimate of effect as the confidence intervals are wide, which would suggest that there is variation which could be due to the stage of endometriosis and also the age of the patients. The authors do report hazards ratios for re-operation stratified by age, but the comparison of hysterectomy + or - oophorectomy is made with laparoscopy, which is an intervention that is not a criterion of the review protocol.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Other information
Full citation 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 Ref Id 370996 Country/ies where the study was carried out USA Study type Retrospective cohort study. Aim of the study To determine the incidence of symptom recurrence and reoperation after hysterectomy for endometriosis, with and without ovarian conservation and to	Sample size N = 138 women Group A (some ovarian tissue preserved) = 29 women Group B (all ovarian tissue removed during hysterectomy) = 109 women Mean length of follow-up was 58 months and was not statistically different between the two groups using the Student's t-test Characteristics Age at time of hysterectomy (years) Group A: 33 (24 to 45) Group B: 35 (22 to 44)P = 0.03	Interventions Hysterectomy with some ovarian tissue preserved. Hysterectomy with removal of all ovarian tissue.	Details A computer search identified 182 women who underwent hysterectomy with the diagnosis of endometriosis. Inpatient charts were reviewed to collect information regarding demographics, previous therapy for endometriosis, surgery performed, surgical findings, and pathology report. Outpatient charts were reviewed to collect follow-up information including symptom recurrence, need for further medical or surgical therapy, findings at subsequent	Results Health related quality of life Not reported Rate of success (disease recurrence and subsequent re-operation rate) Re-operation Hysterectomy without oophorectomy group: 31.0 % (9/29) required reoperation Hysterectomy with oophorectomy group: 3.7% (4/109) required reoperation Cox proportional hazards model: confirmed the crude observation of increased risk of	Limitations CASP checklist for cohort studies 1. Did the study address a clearly focussed issue? (Issue could be in terms of population, risk factors, outcomes considered, is it clear if the study clearly tried to detect a beneficial or harmful effect?) Yes/Unclear/No: yes (To determine the incidence of symptom recurrence and reoperation after hysterectomy for endometriosis, with and without ovarian conservation) 2. Was the cohort recruited in an acceptable way? HINT: Look for selection bias which might compromise the generalisibility of the findings: Was the cohort representative of a defined population? unclear, the participants were recruited from medical records but the authors noted that referral to the centre had meant they are likely to have failed medical and possibly surgical treatment so they may have been more affected than many women with endometriosis.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments

evaluate the effect of HRT on symptom recurrence in patients after hysterectomy with bilateral oophorectomy.
Study dates
1979 to 1991
Source of funding
No information.
	(younger in group with some ovarian tissue preservation)
	Time from diagnosis to hysterectomy (months)Group A: 47.1 (0 to 192)
	Group B: 52 (0 to 216)
	P = not significant
	Parity
	Group A: 1.3 (0 to 2)
	Group B: 0.8 (0 to 4)P = 0.004
	(women with some preservation of ovarian tissue had given birth to more children per woman than those with all ovarian tissue removed)
	Length of medical treatment (months)Group A: 19 (0 to 89)
	Group B: 15 (0 to 84)
	P = not significant
	No of previous diagnostic laparoscopies
	Group A: 1 (0 to 4)
	Group B: 1 (0 to 4)
	P = not significant
			surgery, and timing and dose of HRT.
			When follow-up information was not available from outpatient charts, telephone questionnaires were used to obtain that information. Written questionnaires were sent if the patient could not be reached by telephone.
			Patients who had ovarian tissue conserved at the time of hysterectomy were compared with those who had bilateral oophorectomy.
			Analysis methods
			The X2 test was used to assess the significant association of risk factors with pain recurrence and subsequent surgery.
			The time between total abdominal
				reoperation (P = 0.0023). The relative risk for reoperation in patients with ovarian conservation was 8.1 (95% CI 2.1 to 31.2) compared with patients with oophorectomy adjusting for revised AFS classification of endometriosis stage, previous medical therapy, and age at time of hysterectomy.
				The non-significant covariates with their respective RRs, 95% CIs, and P values are as follows:
				revised AFS stage III versus I, II (RR = 0.2; 95% CI 0.2 to 4.6; P = 0.89);
				revised AFS stage IV versus I, II (RR = 0.9; 95% CI 0.2 to 3.2; P = 0.84);
					Women over the age of 45 were excluded.
					Was there something special about the cohort? no, all women underwent hysterectomy for endometriosis. 138/182 (75.8%) of women undergoing hysterectomy were included. The paper gives clear reasons for exclusions and provides the baseline characteristics for the women not included where possible. They paper makes statements about the population not included being similar to those included.
					Was everybody included who should have been included? this search.
					Yes/Unclear/No: Unclear, it says the computer search identified 182 cases, but it is not clear if there are records that would not have been retrieved from
					Risk of bias: Low
					3. Was the exposure measured accurately to minimise bias?
					HINT: Look for measurement or classification bias:
					Did they use subjective or objective measurements? The exposure (type of surgery e.g hysterectomy +/- oophorectomy) was collected from the medical records, this is unlikely to be biased.
					Do the measurements truly reflect what you want them to (have they been validated)?
					Yes/unclear/No: Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
	No or previous therapeutic surgeries Group A: 1 (0 to 3) Group B: 1 (0 to 4) P = not significant Stage at time of hysterectomy - AFS revised classification of endometriosis (%)Group A: Stages I, II: 51.8; Stage III: 20.7; Stage IV: 27.5 Group B: Stages I, II: 18.3; Stage III: 13.8; Stage IV: 67.8 P = 0.0002 (women with some ovarian tissue preserved were had endometriosis classified as lower stages on the AFS classification compared with women who had all ovarian tissue removed during hysterectomy Inclusion criteria Women who underwent hysterectomy with the diagnosis of endometriosis at		hysterectomy with or without oophorectomy and pain recurrence and/or reoperation was analyzed with the Kaplan-Meier technique, and differences in curves were tested with the Wilcoxon and the log-rank analyses. Cox proportional hazards models were used to allow for adjustment for covariates. The covariates included The American Fertility Society (AFS) revised classification of endometriosis stage at the time of hysterectomy, previous medical therapy for endometriosis, previous surgical therapy for endometriosis, and age at the time of hysterectomy. The relative risk (RR) between	previous medical therapy (RR = 4.4; 95% CI 1.0 to 20.7; P = 0.06); and age at time of hysterectomy (age > 35 versus <35 years): RR = 1.4; 95% CI 0.4 to 4.6; P = 0.57). Pain relief Hysterectomy without oophorectomy group: 62% (18/29) had recurrent symptoms Hysterectomy with oophorectomy group: 10.1% (11/106) had recurrent symptoms Cox proportional hazards model: confirmed the crude observation of increased risk of pain recurrence (P = 0.0001). Adjusting for revised AFS classification of	4. Were all the subjects classified into exposure groups using the same procedure Yes/Unclear/No: Unclear, procedures took place over a period of 12 years in which time the techniques are likely to have changed quite a bit. Also no indication of when in time the oophorectomies took place (i.e. were they all in 1979, for example?). 5. Was the outcome measured accurately to minimise bias? HINT: Look for measurement or classification bias: Did they use subjective or objective measurements? Subjective (pain); Objective (reoperation) Do the measures truly reflect what you want them to (have they been validated)? Unclear for pain. Likely to be a 'yes' or 'no' outcome. Unclear, for pain. They women were called by telephone or written questionnaire. Has a reliable system been established for detecting all the cases (for measuring disease occurrence)? May be difficult for pain, easier for reoperation. Were the measurement methods similar in the different groups? Yes Were the subjects and/or the outcome assessor blinded to exposure (does this matter)? Unclear. People conducting telephone surveys may have known the exposure status of the patient. Yes/Unclear/No: Yes

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments

	the Johns Hopkins Hospital between 1979 and 1991.
	Exclusion criteria
	Patients were excluded if:
	medical records describing the hysterectomy were not available (n = 8),
	follow-up information was unobtainable (n = 23)
	women> 45 years of age at the time of their hysterectomy (n = 13) [so that followup would not be clouded by menopausal changes].
			each independent variable and the outcome variable (pain recurrence or reoperation) was determined. A P value of <0.05 was considered to be significant. Computerized data were analyzed using the Statistical Analysis System.	endometriosis stage, previous medical therapy, previous surgical therapy, and age at time of hysterectomy, the relative risk for pain recurrence in patients with ovarian conservation was 6.1 (95% CI 2.5 to 14.6) compared with patients with oophorectomy.
				The nonsignificant covariates with their respective RRs, 95% CIs, and P values are as follows:
				revised AFS stage III versus I, II (RR = 1.1; 95% CI 0.4 to 3.0; P = 0.79);
				revised AFS stage IV versus I, II (RR = 0.4; 95% CI 0.2 to 1.1; P = 0.08);
				previous medical therapy (RR =
					Risk of bias: Medium (reoperation), High (pain)
					6. Have authors identified all important confounding factors?
					List the ones that you think may be important, that the authors have missed
					Yes/unclear/No: Yes
					
					7. Have the authors taken account of confounding factors in the design and/or analyses?
					HINT: Look for restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors
					Yes/Unclear/No: Yes. Cox proportional hazards models were performed. Models to adjust for classification of disease, previous medical or surgical failure and age at time of hysterectomy.
					
					8. Was the follow up of subject complete enough?
					Yes/Unclear/No: Yes. Reasons were given for all those not completing and some discussion on background characteristics and results where possible.
					
					9. Was the follow up of subjects long enough?
					HINT: Consider The good or bad effects should have had long enough to reveal themselves

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
				2.0; 95% CI 0.8 to 5.0; P = 0.12); previous surgical therapy (RR = 2.8; 95% CI 0.8 to 9.6; P = 0.10); and age at time of hysterectomy (age> 35 versus :535 years: RR = 0.8; 95% CI 0.4 to 1.8; P = 0.66). Unintended effects from treatment Not reported Participant satisfaction with treatment Not reported	The persons that are lost to follow-up may have different outcomes than those available for assessment. 23/182 people were unable to be followed up (12.6%) which seems reasonable for a study spanning a mean of nearly 5 years. The baseline characteristics of people who were lost to follow up are provided in the paper. In an open or dynamic cohort, was there anything special about the outcome of the people leaving, or the exposure of the people entering the cohort? The mean duration of follow up was 58 months. A longer duration may have had different rates. Yes/Unclear/No: Yes Risk of bias: low 10. What are the results of this study? HINT: Consider What are the bottom line results? How strong is the association between exposure and outcome? There is an increased risk in requirement for reoperation and recurrence of pain associated with preservation of ovarian tissue compared with removal of ovarian tissue at the time of hysterectomy. What is the absolute risk (AR)? 11. How precise are the results? HINT: Look for the range of the confidence intervals, if given.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					The results are not precise as the confidence intervals are wide, but they are statistically significant. 12. Do you believe the results? HINT: Consider Big effect is hard to ignore! Can it be due to bias, chance or confounding? Are the design and methods of this study sufficiently flawed to make the results unreliable? Bradford Hills criteria (e.g. time sequence, dose-response gradient, biological plausibility, consistency) The results do reflect what is expected to happen, that there would be fewer re-operation events for women who have hysterectomy+oophorectomy as ovaries are removed. There is a large difference in the size of population who underwent oophorectomy (n=29) and those who didn't (n=109). Yes/unclear/no: Unclear Risk of bias: medium 13. Can the results be applied to the local population? HINT: Consider whether A cohort study was the appropriate method to answer this question The subjects covered in this study could be sufficiently different from your population to cause concern Your local setting is likely to differ much from that of the study

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					You can quantify the local benefits and harms Yes/unclear/no: Unclear. The result shows clinical benefit for hysterectomy+oophorectomy, but as the results are not statistically significant. Results are for patients undergoing surgery between 1979 and 1991, which may not represent the same techniques as surgery today. 14. Do the results of this study fit with other available evidence? Yes/unclear/no: Yes, to a certain extent. The other paper did not have significant results but it did have results suggestive of the same pattern. 15. What are the implications of this study for practice? HINT: Consider One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making For certain questions observational studies provide the only evidence Recommendations from observational studies are always stronger when supported by other evidence The direction of effect of re-operation favours women who have hysterectomy and oophorectomy over 5 years but there is imprecision around the estimate of effect as the confidence intervals are wide.

Study details	Participants	Interventions	Methods	Outcomes and Results	Comments
					Other information The paper also looks at the number of women who were prescribed Hormone Replacement Therapy (HRT) and the timing of this intervention.

G.18 Review question: Pharmacological, non-pharmacological, surgical and combination management strategies - if fertility is a priority Management strategies to improve spontaneous pregnancy rates

No evidence tables were prepared for studies included in the NMA analysis

G.19 Economic Evidence

Study	Limitations	Applicability	Other comments	Costs	Effects	ICER	Uncertainty
Araujo 2011	Costs only Six month time horizon	Limited applicability (Brazilian study)	Goserelin acetate for all vs goseralin acetate for thiose with confirmed deep endometriosis only Costs obtained from Ambulatory and Hospital Information System and Price Database of Brazilian	Treating all USD$1662 cheaper	N/A	N/A	None described

Study	Limitations	Applicability	Other comments	Costs	Effects	ICER	Uncertainty
			Ministry of Health				
Avxentyeva 2013	Costs only, abstract only Unclear if modelling or direct clinical evidence Six month time horizon	Limited applicability (Russian study)		Triprorelin = €1102 Leuprorelin = €1118 Buserelin = €340 Dydrogesterone = €369 Dienogest = €295	"Literature search did not reveal clinically significant differences", otherwise none reported	N/A	None described
Bodner 1996	Costs obtained from interviews with clinical managers, not standard reference sources Did not account for indirect costs Population had comorbid infertility Dated	Partially applicability (Scottish study)	Cohorting very imperfect – control arm much healthier to begin with 6% discount rate	Medical arm £645.02 Expectant management arm £387.29	SF-36 score Medical arm 61 (21.1) to 61.4 (29.9) Expectant management arm 76.4 (18.2) to 75.3 (22.)	N/A	Three univariate sensitivity analyses presented. Most significant is increasing length of stay in hospital
Lalchandani 2005	Small population Did not account for indirect costs	Directly applicable (UK study)	GnHR limited to six months because of bone mineral density risk but time horizon standard 12 months	Surgical arm £323.29 Medical arm £918.12	Medical arm 3/18 symptom free, 11/17 required surgical treatment	N/A	Univariate and multivariate sensitivity analysis undertaken

Study	Limitations	Applicability	Other comments	Costs	Effects	ICER	Uncertainty
	Source of direct costs unclear; much lower than values in NHS Reference Costs				Surgical arm 9/17 symptom free, 3/17 required surgical treatment		
Lukac 2005a	Source of direct costs "Published price lists, clinical guidelines, product labels and expert opinion" and therefore applicability unclear 5% discount rate and SF-36 QoL instrument used so not in keeping with NICE Reference Case	Partial applicability (Slovakian study)	Markov chain design Part of AU19 trial	GnHR €1248 Dienogest €969	SF-36 Dienogest gains 0.002 QALY, but unclear what control arm got	Dienogest dominates	CEAC considered; found in 69% of cases Dienogest was below 18,000 E / QALY (which is the Slovakian threshold)
Lukac 2005b	Source of direct costs "Published price lists, clinical guidelines, product labels and expert opinion" and therefore applicability unclear	Partial applicability (Slovakian study)	Markov chain design Part of AU19 trial Appears to be re-analysis of Lukac 2005a with longer time horizon (5 years vs 2 years)	No direct costs given Dienogest saves €426	SF-36 Dienogest gains 0.069 QALY, but unclear what control arm got	Dienogest dominates	CEAC considered; found in 79% of cases Dienogest was below 18,000 E / QALY (which is the Slovakian threshold)

Study	Limitations	Applicability	Other comments	Costs	Effects	ICER	Uncertainty
	5% discount rate and SF-36 QoL instrument used so not in keeping with NICE Reference Case						
Romero 2012	Costs only Unclear why arms have different treatment lengths – possibly to do with side effects of GnRHa Cross-national groups not randomised – some patients in Argentina were given local schedule of treatment	Limited applicability (Columbian study)		Colombia - Diogenest US$986.16 vs GnHR US$2855.57 Argentinia Schedule 1 - Dienogest US$490.75 vs GnRH US$812.21 Argentinia Schedule 2 - Diengest US$490.75 vs GnHR $1386.21	N/A	N/A	None described
Tuletova 2014	Quality of life measure not NICE standard and does not appear to be used anywhere but this study, making comparison difficult	Limited applicability (Kazakhstani study)		Direct medical expenses Endometriosis surgery 143298 KT (Kazakhstani Tenge)	'Efficacy index' Endometriosis surgery 66.7% Hormonal treatment 70.0%	N/A	No sensitivity analysis undertaken

Study	Limitations	Applicability	Other comments	Costs	Effects	ICER	Uncertainty
				Hormonal treatment 92428 KT Combined treatment 115718 KT	Combined treatment 91.7%		
Wasiak 2013	Based on data from Cardiff and Vale Trust only Nonrandomised	Directly applicable (UK study)	Retrospective Cohort Design	Surgical £871 cost per visit, 1.4 (1.4) GP visits in previous 6 weeks, length of stay 0.4 (0.7) Clinical £1525.20 cost per visit, 2.0 (2.9) GP visits in previous 6 weeks, length of stay 2.2 (3.4)	EQ-5D Surgical arm 0.70 (0.32) Clinical arm 0.71 (0.27)	N/A	No sensitivity analysis described
Prast 2013	Nonrandomised Small population	Partially applicable (Austrian study)	Costs only	Surgical costs €3466.60 (3712.42) Medical costs €116.90 (293.94)	N/A	N/A	N/A
Simoens 2012	Nonrandomised	Partially applicable (ten countries, including the UK)	Costs only Part of EndoCost consortium	Direct costs €3281.0 (13336.40) Indirect costs (not relevant to	N/A	N/A	N/A

Study	Limitations	Applicability	Other comments	Costs	Effects	ICER	Uncertainty
				NICE methodology) €6298.30 (7262.60)			
Schwartz 1994	Costs only Nonrandomised Very unusual trial design which would not normally be considered in NICE evidence evaluation	Partially applicable (US study)	Time horizon 10.9 months	Costs are 10.9 months before MRI (10.9 months after MRI) for entire cohort All surgery $157,630 ($106,878) Abdominal surgery $147,363 ($76,169) Medical treatment $17,676 ($64,488)	N/A	N/A	No sensitivity analysis described
Sanghera 2016	No discount rate specified Expert elicitation used to identify QALY values, with substantially non-intuitive results not explained in text	Partial (UK study but modelling approach only)	Time horizon 36 months	DMPA £622.56 LNG-IUS £650.94 COCP £599.93 No treatment £371.34	QALY values DMPA 1.92 LNG-IUS 1.88 COCP 1.92 No treatment 2.27	No treatment dominates	Probabilistic uncertainty analysis undertaken with no major changes to results

Study	Limitations	Applicability	Other comments	Costs	Effects	ICER	Uncertainty
Zalis'ka 2014	Costs only No discount rate specified, source of cost data unclear, short follow up (six months)	Limited applicability(Ukranian study)		Dydogesterone = USD $345 Dienogest = USD $1347 triptorelin = USD $1347	N/A	N/A	N/A
Zhao 1998	Costs only Short follow-up (six months) Unusual study design – descriptive analysis of retrospective cohort	Partially applicable (US study)	Source of cost data Medstat MarketScan database	Data given is USD geometric mean Nafarelin (log SD) / geometric mean Leuprolide (log SD) Drug cost 692.9 (0.31) / 953.8 (0.27) Other drugs 127.6 (0.96) / 112.5 (0.89) Outpatient services 733.8 (0.70) / 816.1 (0.67) Endometriosis-related inpatient admissions 364.2 (0.16) / 362.8 (0.11)	N/A	N/A	None described, but uncertainty intervals carefully chosen to reflect uncertainty

1

2