Cochrane Library

Cochrane Database of Systematic Reviews

Blood biomarkers for the non-invasive diagnosis of endometriosis (Review)

Nisenblat V, Bossuyt PMM, Shaikh R, Farquhar C, Jordan V, Scheffers CS, Mol BWJ, Johnson N, Hull ML

Nisenblat V, Bossuyt PMM, Shaikh R, Farquhar C, Jordan V, Scheffers CS, Mol BWJ, Johnson N, Hull ML.
Blood biomarkers for the non-invasive diagnosis of endometriosis.
Cochrane Database of Systematic Reviews 2016, Issue 5. Art. No.: CD012179.
DOI: 10.1002/14651858.CD012179.

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Blood biomarkers for the non-invasive diagnosis of endometriosis

Vicki Nisenblat[1], Patrick MM Bossuyt[2], Rabia Shaikh[3], Cindy Farquhar[4], Vanessa Jordan[4], Carola S Scheffers[5], Ben Willem J Mol[1], Neil Johnson[1], M Louise Hull[1]

[1]Discipline of Obstetrics and Gynaecology, School of Medicine, Robinson Research Institute, The University of Adelaide, Adelaide, Australia. [2]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands. [3]Westmead Fertility Clinic, IVF Australia, Sydney, Australia. [4]Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand. [5]University of Groningen, Groningen, Netherlands

Contact: Vicki Nisenblat, Discipline of Obstetrics and Gynaecology, School of Medicine, Robinson Research Institute, The University of Adelaide, Level 6, Medical School North,, Frome Rd, Adelaide, SA, 5005, Australia. vnisenblat@gmail.com.

Editorial group: Cochrane Gynaecology and Fertility Group.
Publication status and date: New, published in Issue 5, 2016.

Citation: Nisenblat V, Bossuyt PMM, Shaikh R, Farquhar C, Jordan V, Scheffers CS, Mol BWJ, Johnson N, Hull ML. Blood biomarkers for the non-invasive diagnosis of endometriosis. *Cochrane Database of Systematic Reviews* 2016, Issue 5. Art. No.: CD012179. DOI: 10.1002/14651858.CD012179.

ABSTRACT

Background

About 10% of reproductive-aged women suffer from endometriosis, a costly chronic disease causing pelvic pain and subfertility. Laparoscopy is the gold standard diagnostic test for endometriosis, but is expensive and carries surgical risks. Currently, there are no non-invasive or minimally invasive tests available in clinical practice to accurately diagnose endometriosis. Although other reviews have assessed the ability of blood tests to diagnose endometriosis, this is the first review to use Cochrane methods, providing an update on the rapidly expanding literature in this field.

Objectives

To evaluate blood biomarkers as replacement tests for diagnostic surgery and as triage tests to inform decisions on surgery for endometriosis. Specific objectives include:

1. To provide summary estimates of the diagnostic accuracy of blood biomarkers for the diagnosis of peritoneal, ovarian and deep infiltrating pelvic endometriosis, compared to surgical diagnosis as a reference standard.

2. To assess the diagnostic utility of biomarkers that could differentiate ovarian endometrioma from other ovarian masses.

Search methods

We did not restrict the searches to particular study designs, language or publication dates. We searched CENTRAL to July 2015, MEDLINE and EMBASE to May 2015, as well as these databases to 20 April 2015: CINAHL, PsycINFO, Web of Science, LILACS, OAIster, TRIP, ClinicalTrials.gov, DARE and PubMed.

Selection criteria

We considered published, peer-reviewed, randomised controlled or cross-sectional studies of any size, including prospectively collected samples from any population of reproductive-aged women suspected of having one or more of the following target conditions: ovarian, peritoneal or deep infiltrating endometriosis (DIE). We included studies comparing the diagnostic test accuracy of one or more blood biomarkers with the findings of surgical visualisation of endometriotic lesions.

Data collection and analysis

Two authors independently collected and performed a quality assessment of data from each study. For each diagnostic test, we classified the data as positive or negative for the surgical detection of endometriosis, and we calculated sensitivity and specificity estimates. We used the bivariate model to obtain pooled estimates of sensitivity and specificity whenever sufficient datasets were available. The predetermined criteria for a clinically useful blood test to replace diagnostic surgery were a sensitivity of 0.94 and a specificity of 0.79 to detect endometriosis. We set the criteria for triage tests at a sensitivity of ≥ 0.95 and a specificity of ≥ 0.50, which 'rules out' the diagnosis with high accuracy if there is a negative test result (SnOUT test), or a sensitivity of ≥ 0.50 and a specificity of ≥ 0.95, which 'rules in' the diagnosis with high accuracy if there is a positive result (SpIN test).

Main results

We included 141 studies that involved 15,141 participants and evaluated 122 blood biomarkers. All the studies were of poor methodological quality. Studies evaluated the blood biomarkers either in a specific phase of the menstrual cycle or irrespective of the cycle phase, and they tested for them in serum, plasma or whole blood. Included women were a selected population with a high frequency of endometriosis (10% to 85%), in which surgery was indicated for endometriosis, infertility work-up or ovarian mass. Seventy studies evaluated the diagnostic performance of 47 blood biomarkers for endometriosis (44 single-marker tests and 30 combined tests of two to six blood biomarkers). These were angiogenesis/growth factors, apoptosis markers, cell adhesion molecules, high-throughput markers, hormonal markers, immune system/inflammatory markers, oxidative stress markers, microRNAs, tumour markers and other proteins. Most of these biomarkers were assessed in small individual studies, often using different cut-off thresholds, and we could only perform meta-analyses on the data sets for anti-endometrial antibodies, interleukin-6 (IL-6), cancer antigen-19.9 (CA-19.9) and CA-125. Diagnostic estimates varied significantly between studies for each of these biomarkers, and CA-125 was the only marker with sufficient data to reliably assess sources of heterogeneity.

The mean sensitivities and specificities of anti-endometrial antibodies (4 studies, 759 women) were 0.81 (95% confidence interval (CI) 0.76 to 0.87) and 0.75 (95% CI 0.46 to 1.00). For IL-6, with a cut-off value of > 1.90 to 2.00 pg/ml (3 studies, 309 women), sensitivity was 0.63 (95% CI 0.52 to 0.75) and specificity was 0.69 (95% CI 0.57 to 0.82). For CA-19.9, with a cut-off value of > 37.0 IU/ml (3 studies, 330 women), sensitivity was 0.36 (95% CI 0.26 to 0.45) and specificity was 0.87 (95% CI 0.75 to 0.99).

Studies assessed CA-125 at different thresholds, demonstrating the following mean sensitivities and specificities: for cut-off > 10.0 to 14.7 U/ml: 0.70 (95% CI 0.63 to 0.77) and 0.64 (95% CI 0.47 to 0.82); for cut-off > 16.0 to 17.6 U/ml: 0.56 (95% CI 0.24, 0.88) and 0.91 (95% CI 0.75, 1.00); for cut-off > 20.0 U/ml: 0.67 (95% CI 0.50 to 0.85) and 0.69 (95% CI 0.58 to 0.80); for cut-off > 25.0 to 26.0 U/ml: 0.73 (95% CI 0.67 to 0.79) and 0.70 (95% CI 0.63 to 0.77); for cut-off > 30.0 to 33.0 U/ml: 0.62 (95% CI 0.45 to 0.79) and 0.76 (95% CI 0.53 to 1.00); and for cut-off > 35.0 to 36.0 U/ml: 0.40 (95% CI 0.32 to 0.49) and 0.91 (95% CI 0.88 to 0.94).

We could not statistically evaluate other biomarkers meaningfully, including biomarkers that were assessed for their ability to differentiate endometrioma from other benign ovarian cysts.

Eighty-two studies evaluated 97 biomarkers that did not differentiate women with endometriosis from disease-free controls. Of these, 22 biomarkers demonstrated conflicting results, with some studies showing differential expression and others no evidence of a difference between the endometriosis and control groups.

Authors' conclusions

Of the biomarkers that were subjected to meta-analysis, none consistently met the criteria for a replacement or triage diagnostic test. A subset of blood biomarkers could prove useful either for detecting pelvic endometriosis or for differentiating ovarian endometrioma from other benign ovarian masses, but there was insufficient evidence to draw meaningful conclusions. Overall, none of the biomarkers displayed enough accuracy to be used clinically outside a research setting. We also identified blood biomarkers that demonstrated no diagnostic value in endometriosis and recommend focusing research resources on evaluating other more clinically useful biomarkers.

PLAIN LANGUAGE SUMMARY

Blood biomarkers for the non-invasive diagnosis of endometriosis

Review Question

How accurate are blood tests in detecting endometriosis? Can any blood test be accurate enough to replace or reduce the need for surgery in the diagnosis of endometriosis?

Background

Women with endometriosis have endometrial tissue (the tissue that lines the womb and is shed during menstruation) growing outside the womb within the pelvic cavity. This tissue responds to reproductive hormones, causing painful periods, chronic lower abdominal pain and difficulty conceiving. Currently, the only reliable way of diagnosing endometriosis is to perform keyhole surgery and visualise the endometrial deposits inside the abdomen. Because surgery is risky and expensive, we evaluated whether the results of blood tests

(blood biomarkers) can help to detect endometriosis non-invasively. An accurate blood test could lead to the diagnosis of endometriosis without the need for surgery, or it could reduce the need for diagnostic surgery to a group of women who were most likely to have endometriosis. Separate Cochrane reviews from this series evaluate other non-invasive ways of diagnosing endometriosis using urine, imaging, endometrial and combination tests.

Study characteristics

The evidence included in this review is current to July 2015. We included 141 studies involving 15,141 participants. All studies evaluated reproductive-aged women who were undertaking diagnostic surgery because they were suspected of having one or more of the following target conditions: ovarian, peritoneal or deep infiltrating endometriosis (DIE). Cancer antigen-125 (CA-125) was the most common blood biomarker studied. Seventy studies evaluated 47 blood biomarkers that were expressed differently in women with and without endometriosis, and 82 studies identified 97 biomarkers that did not distinguish between the two groups. Twenty-two biomarkers were in both categories.

Key results

Only four of the assessed biomarkers (anti-endometrial Abs (anti-endometrial autoantibodies), interleukin-6 (IL-6), CA-19.9 and CA-125) were evaluated by enough studies to provide a meaningful assessment of test accuracy. None of these tests was accurate enough to replace diagnostic surgery. Several studies identified biomarkers that might be of value in diagnosing endometriosis, but there are too few reports to be sure of their diagnostic benefit. Overall, there is not enough evidence to recommend testing for any blood biomarker in clinical practice to diagnose endometriosis.

Quality of the evidence

Generally, the reports were of low methodological quality, and most blood tests were only assessed by a single or a small number of studies. When the same biomarker was studied, there were significant differences in how studies were conducted, the group of women studied and the cut-offs used to determine a positive result.

Future research

More high quality research trials are necessary to accurately assess the diagnostic potential of certain blood biomarkers, whose diagnostic value for endometriosis was suggested by a limited number of studies.

SUMMARY OF FINDINGS

Summary of findings 1. Biomarkers evaluated as a diagnostic test for endometriosis

Review question	What is the diagnostic accuracy of the blood biomarkers in detecting pelvic endometriosis (peritoneal endometriosis, endometrioma, deep infiltrating endometriosis)?	
Importance	A simple and reliable non-invasive test for endometriosis with the potential to either replace laparoscopy or to triage patients in order to reduce surgery, would minimise surgical risk and reduce diagnostic delay	
Patients	Reproductive aged women with suspected endometriosis or persistent ovarian mass, or women undergoing infertility work-up or gynaecological laparoscopy	
Settings	Hospitals (public or private of any level), outpatient clinics (general gynaecology, reproductive medicine, pelvic pain) or research laboratories	
Reference standard	Visualisation of endometriosis at surgery (laparoscopy or laparotomy) with or without histological confirmation	
Study design	Cross-sectional of a single-gate design (N = 25) or a two-gate design (N = 44); unable to determine if single- or two-gate design for 1 study; prospective enrolment; a single study could assess more than one test	
Risk of bias	Overall judgement	Poor quality of most of the studies (no study had a 'low risk' assessment in all 4 domains)
	Patient selection bias	High risk: 31 studies; unclear risk: 32 studies; low risk: 7 studies
	Index test interpretation bias	High risk: 56 studies; unclear risk: 12 studies; low risk: 2 studies
	Reference standard interpretation bias	High risk: 0 studies; unclear risk: 30 studies; low risk: 40 studies
	Flow and timing selection bias	High risk: 11 studies; unclear risk: 3 studies; low risk: 56 studies
Applicability concerns	Concerns regarding patient selection	High concern: 32 studies; unclear concern: 5 studies; low concern: 33 studies
	Concerns regarding index test	High concern: 0 studies; unclear concern: 4 studies; low concern: 66 studies
	Concerns regarding reference standard	High concern: 0 studies; unclear concern: 0 studies; low concern: 70 studies
Diagnostic criteria	Replacement test: sensitivity ≥ 94 and specificity ≥ 79 SnOUT triage test: sensitivity ≥ 95 and specificity ≥ 50 SpIN triage test: sensitivity ≥ 50 and specificity ≥ 95	

Test	N participants (studies)	Outcomes				Diagnostic estimates (95% CI)	Implications
		True positives (endometriosis)	False positives (incorrectly classified as endometriosis)	True negatives (disease-free)	False negatives (incorrectly classified as disease-free)		
1. Angiogenesis and growth factors and their receptors							
Glycodelin-A cut-off threshold > 2.07 ng/ml follicular or luteal cycle phase rASRM stage I-IV[c]	99 (1)	47	9	33	10	Sens = 0.82 (0.70 to 0.91); spec = 0.79 (0.63 to 0.90)	Insufficient evidence to draw meaningful conclusions
Glycodelin[a#] cut-off threshold > 9.0 ng/ml follicular cycle phase rASRM stage I-IV[b]	45 (1)	20	11	6	8	Sens = 0.71 (0.51 to 0.87); spec = 0.35 (0.14 to 0.62)	Insufficient evidence to draw meaningful conclusions
Glycodelin[a#] cut-off threshold > 18 ng/ml any cycle phase rASRM stage I-IV[b]	99 (1)	36	23	18	22	Sens = 0.62 (0.48 to 0.74); spec = 0.44 (0.28 to 0.60)	Insufficient evidence to draw meaningful conclusions
IGFBP-3 (insulin-like growth factor-binding protein-3)[a*] cut-off threshold > 200 ng/ml follicular cycle phase rASRM I-IV[b]	45 (1)	20	12	5	8	Sens = 0.71 (0.51 to 0.87); spec = 0.29 (0.10 to 0.56)	Insufficient evidence to draw meaningful conclusions
IGFBP-3 (insulin-like growth factor-binding protein-3)[a*] cut-off threshold > 210 ng/ml	99 (1)	32	23	18	26	Sens = 0.55 (0.42 to 0.68);	Insufficient evidence to draw meaningful conclusions

any cycle phase rASRM I-IV[b]						spec = 0.44 (0.28 to 0.60)	
VEGF (vascular endothelial growth factor) cut-off threshold > 1.5 pg/ml any cycle phase rASRM I-IV[b]	99 (1)	29	16	25	29	Sens = 0.50 (0.37 to 0.63); spec = 0.61 (0.45 to 0.76)	Insufficient evidence to draw meaningful conclusions
VEGF (vascular endothelial growth factor) cut-off threshold > 236 pg/ml follicular cycle phase rASRM I-IV	95 (1)	60	7	23	5	Sens = 0.92 (0.83 to 0.97); spec = 0.77 (0.58 to 0.90)	Insufficient evidence to draw meaningful conclusions; approaches the criteria for a replacement and SnOUT triage test; further diagnostic test accuracy studies recommended
VEGF (vascular endothelial growth factor) cut-off threshold > 680 pg/ml follicular cycle phase rASRM III-IV	60 (1)	28	1	29	2	Sens = 0.93 (0.78 to 0.99); spec = 0.97 (0.83 to 1.00)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test and approaches criteria for a replacement test; further diagnostic test accuracy studies recommended
Urocortin[a&] cut-off threshold > 29 pg/ml cycle phase not specified rASRM III-IV[d]	80 (1)	39	6	34	1	Sens = 0.97 (0.87 to 1.00); spec = 0.85 (0.70 to 0.94)	Insufficient evidence to draw meaningful conclusions; meets criteria for a replacement and SnOUT triage test; further diagnostic test accuracy studies recommended
Urocortin[a&] cut-off threshold > 33 pg/ml cycle phase not specified rASRM III-IV[d]	80 (1)	35	4	36	5	Sens = 0.88 (0.73 to 0.96); spec = 0.90 (0.76 to 0.97)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SpIN triage test; further diagnostic test accuracy studies recommended

Urocortin cut-off threshold > 41.6 pg/ml follicular cycle phase rASRM III-IV[d]	88 (1)	32	25	21	10	Sens = 0.76 (0.61 to 0.88); spec = 0.46 (0.31 to 0.61)	Insufficient evidence to draw meaningful conclusions
2. Apoptosis markers							
Survivin cut-off threshold not reported follicular cycle phase rASRM stage not reported[e]	60 (1)	3	2	18	37	Sens = 0.07 (0.02 to 0.20); spec = 0.90 (0.68 to 0.99)	Insufficient evidence to draw meaningful conclusions
3. Cell adhesion molecules and other matrix-related proteins							
sICAM-1 (soluble form of intercellular-adhesion molecule-1)[a#] cut-off threshold < 243 ng/ml any cycle phase rASRM I-IV[b]	99 (1)	32	21	21	26	Sens = 0.55 (0.42 to 0.68); spec = 0.50 (0.34 to 0.66)	Insufficient evidence to draw meaningful conclusions
sICAM-1 (soluble form of intercellular-adhesion molecule-1)[a#] cut-off threshold < 254.6 ng/ml menstrual cycle phase rASRM I-IV[b]	28 (1)	8	12	5	3	Sens = 0.73 (0.39 to 0.94); spec = 0.29 (0.10 to 0.56)	Insufficient evidence to draw meaningful conclusions
sICAM-1 (soluble form of intercellular-adhesion molecule-1) cut-off threshold > 241.46 µg/ml cycle phase not specified	60 (1)	18	4	26	12	Sens = 0.60 (0.41 to 0.77); spec = 0.87 (0.69 to 0.96)	Insufficient evidence to draw meaningful conclusions

rASRM I-IV							
LN-1 (laminin-1) cut-off threshold > 1110.0 pg/ml cycle phase not specified rASRM II-IV	73 (1)	38	6	14	15	Sens = 0.72 (0.58 to 0.83); spec = 0.70 (0.46 to 0.88)	Insufficient evidence to draw meaningful conclusions
4. High-throughput molecular markers							
Metabolome by ESI-MS/MS (SMOH C16:1 + PCaa C36:2/ PCae C34:2) any cycle phase rASRM III-IV[e] age/body mass index-adjusted	92 (1)	36	8	44	4	Sens = 0.90 (0.76 to 0.97); spec = 0.85 (0.72 to 0.93)	Insufficient evidence to draw meaningful conclusions; approaches criteria of a replacement and SnOUT triage test; further diagnostic test accuracy studies recommended
Proteome by SELDI-TOF-MS (3 peaks with the MW 3956.00, 11,710.00 and 6986.00 Da) cycle phase not specified rASRM I-IV	31 (1)	14	3	12	2	Sens = 0.88 (0.62 to 0.98); spec = 0.80 (0.52 to 0.96)	Insufficient evidence to draw meaningful conclusions; further diagnostic test accuracy studies using standardised methodology is recommended
Proteome by SELDI-TOF MS (5 peaks with MW 4159.00, 5264.00, 5603.00, 9861.00 and 10,533.00 Da) follicular/ luteal cycle phase rASRM I-IV	90 (1)	40	16	23	11	Sens = 0.78 (0.65 to 0.89); spec = 0.59 (0.42 to 0.74)	Insufficient evidence to draw meaningful conclusions; further diagnostic test accuracy studies using standardised methodology is recommended
Proteome by SELDI-TOF MS (5 peaks with MW 9926.31, 10,072.2, 6753.04, 4302.67, 9328.49 Da) menstrual cycle phase	67 (1)	18	4	18	27	Sens = 0.40 (0.26 to 0.56); spec = 0.82 (0.60 to 0.95)	Insufficient evidence to draw meaningful conclusions; further diagnostic test accuracy studies using standardised methodology is recommended

rASRM I-IV							
Proteome by SELDI-TOF MS (5 peaks with MW 2831.02, 7554.66, 4241.29, 2953.25, 9927.73 Da) follicular cycle phase rASRM I-IV	98 (1)	25	5	28	40	Sens = 0.38 (0.27 to 0.51); spec = 0.85 (0.68 to 0.95)	Insufficient evidence to draw meaningful conclusions; further diagnostic test accuracy studies using standardised methodology is recommended
Proteome by SELDI-TOF MS (5 peaks with MW 11,366.3, 5712.69, 10,070.7, 3017.68, 3824.44 Da) luteal cycle phase rASRM I-IV	88 (1)	29	6	27	26	Sens = 0.53 (0.39to 0.66); spec = 0.82 (0.65 to 0.93)	Insufficient evidence to draw meaningful conclusions; further diagnostic test accuracy studies using standardised methodology is recommended
Proteome by SELDI-TOF-MS (6 peaks with MW 1629, 3047, 3526, 3774, 5046 and 5068 Da) any cycle phase rASRM II-IV	139 (1)	40	1	77	21	Sens = 0.66 (0.52 to 0.77); spec = 0.99 (0.93 to 1.00)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test; further diagnostic test accuracy studies using standardised methodology is recommended
5. Hormonal markers							
Prolactin[1a^] cut-off threshold > 14.8 ng/ml luteal cycle phase rASRM I-IV[c]	97 (1)	28	2	32	35	Sens = 0.44 (0.32 to 0.58); spec = 0.94 (0.80 to 0.99)	Insufficient evidence to draw meaningful conclusions
Prolactin[1a^] cut-off threshold > 20 ng/ml luteal cycle phase rASRM I-IV[c]	97 (1)	13	0	34	50	Sens = 0.21 (0.11 to 0.33); spec = 1.00 (0.90 to 1.00)	Insufficient evidence to draw meaningful conclusions

6. Immune system and inflammatory markers

Anti-endometrial Abs cut-off threshold - definitions for positive result varied cycle phase varied (not specified in 2 studies) rASRM I-IV in 3 studies; not reported in 1 study	759 (4)	359	48	276	76	Sens = 0.81 (0.76 to 0.87); spec = 0.75 (0.46 to 1.00)	Summary estimates did not meet the predetermined criteria for triage or replacement test; varying methodologies and populations across the studies
Anti-endometrial Abs (MW of 26/34/42 kd) cut-off threshold: dark band in the blot for at least 1 Ab cycle phase not specified rASRM I-IV	36 (1)	18	11	7	0	Sens = 1.00 (0.81 to 1.00); spec = 0.39 (0.17 to 0.64)	Insufficient evidence to draw meaningful conclusions
Anti-laminin auto Abs cut-off threshold > 1 U/ml cycle phase not specified rASRM I-IV	68 (1)	17	3	23	25	Sens = 0.40 (0.26 to 0.57); spec = 0.88 (0.70 to 0.98)	Insufficient evidence to draw meaningful conclusions
sCD23 (soluble CD23) cut-off threshold: absorbance value of ELISA > control mean ± 2SD (standard deviations) follicular or luteal cycle phase rASRM I-IV	97 (1)	14	3	37	43	Sens = 0.25 (0.14 to 0.38); spec = 0.93 (0.80 to 0.98)	Insufficient evidence to draw meaningful conclusions
MCP-1 (monocyte chemotactic protein-1) cut-off threshold > 100 pg/ml menstrual cycle phase	101 (1)	37	17	27	20	Sens = 0.65 (0.51 to 0.77); spec = 0.61 (0.45 to 0.76)	Insufficient evidence to draw meaningful conclusions

rASRM I-IV							
Copeptin cut-off threshold > 251.2 pg/ml cycle phase not specified rASRM I-IV	87 (1)	33	15	21	18	Sens = 0.65 (0.50 to 0.78); spec = 0.58 (0.41 to 0.74)	Insufficient evidence to draw meaningful conclusions
hs-CRP (high sensitive C-reactive protein)[a$] cut-off threshold > 0.62 mg/l any cycle phase rASRM I-IV	295 (1)	126	40	51	78	Sens = 0.62 (0.55 to 0.68); spec = 0.56 (0.45 to 0.66)	Insufficient evidence to draw meaningful conclusions
hs-CRP (high sensitive C-reactive protein)[a$] cut-off threshold > 0.73 mg/l menstrual cycle phase rASRM I-IV	60 (1)	28	10	9	13	Sens = 0.68 (0.52 to 0.82); spec = 0.47 (0.24 to 0.71)	Insufficient evidence to draw meaningful conclusions
hs-CRP (high sensitive C-reactive protein)[a$] cut-off threshold > 0.61 mg/l follicular cycle phase rASRM I-IV	119 (1)	45	18	18	38	Sens = 0.54 (0.43 to 0.65); spec = 0.50 (0.33 to 0.67)	Insufficient evidence to draw meaningful conclusions
hs-CRP (high sensitive C-reactive protein) cut-off threshold >438 µg/ml follicular cycle phase rASRM I-IV	95 (1)	54	4	26	11	Sens = 0.83 (0.72 to 0.91); spec = 0.87 (0.69 to 0.96)	Insufficient evidence to draw meaningful conclusions
hs-CRP (high sensitive C-reactive protein)[a$]	116 (1)	47	13	23	33	Sens = 0.59 (0.47 to 0.70);	Insufficient evidence to draw meaningful conclusions

cut-off threshold > 0.70 mg/l luteal cycle phase rASRM I-IV						spec = 0.64 (0.46 to 0.79)	
hs-CRP (high sensitive C-reactive protein)[a$] cut-off threshold not specified luteal cycle phase rASRM I-IV	116 (1)	32	11	27	46	Sens = 0.41 (0.30 to 0.53); spec = 0.71 (0.54 to 0.85)	Insufficient evidence to draw meaningful conclusions
IFN-γ (interferon-gamma) cut-off threshold < 76 pg/ml follicular cycle phase rASRM I-IV[b]	45 (1)	19	6	11	9	Sens = 0.68 (0.48 to 0.84); spec = 0.65 (0.38 to 0.86)	Insufficient evidence to draw meaningful conclusions
MIF (macrophage migration inhibitory factor) cut-off threshold > 0.57 ng/ml follicular or luteal cycle phase rASRM I-IV	93 (1)	36	13	25	19	Sens = 0.65 (0.51 to 0.78); spec = 0.66 (0.49 to 0.80)	Insufficient evidence to draw meaningful conclusions
TNF-α (tumour necrosis factor alpha) cut-off threshold >12.45 pg/ml follicular cycle phase rASRM I-IV	95 (1)	58	4	26	7	Sens = 0.89 (0.79 to 0.96); spec = 0.87 (0.69 to 0.96)	Insufficient evidence to draw meaningful conclusions
TNF-α (tumour necrosis factor alpha) cut-off threshold < 45.6 pg/ml follicular cycle phase rASRM I-IV[b]	45 (1)	19	11	6	9	Sens = 0.68 (0.48 to 0.84); spec = 0.35 (0.14 to 0.62)	Insufficient evidence to draw meaningful conclusions

TNF-α (tumour necrosis factor alpha) cut-off threshold not reported luteal cycle phase rASRM I-IV	116 (1)	62	10	28	16	Sens = 0.79 (0.69 to 0.88); spec = 0.74 (0.57 to 0.87)	Insufficient evidence to draw meaningful conclusions
Neutrophils cut-off threshold > 4058 cells/ml menstrual cycle phase rASRM I-IV	100 (1)	34	20	30	16	Sens = 0.68 (0.53 to 0.80); spec = 0.60 (0.45 to 0.74)	Insufficient evidence to draw meaningful conclusions
NLR (neutrophil-to-lymphocyte ratio) cut-off threshold > 2.19 menstrual cycle phase rASRM I-IV	100 (1)	38	9	41	12	Sens = 0.76 (0.62 to 0.87); spec = 0.82 (0.69 to 0.91)	Insufficient evidence to draw meaningful conclusions
WBC (white blood cells) cut-off threshold > 6400 cells/ml menstrual cycle phase rASRM I-IV	100 (1)	32	23	27	18	Sens = 0.64 (0.49 to 0.77); spec = 0.54 (0.39 to 0.68)	Insufficient evidence to draw meaningful conclusions
IL-1β (interleukin - 1beta) cut-off threshold < 0.9 pg/ml follicular cycle phase rASRM I-IV[b]	45 (1)	23	11	6	5	Sens = 0.82 (0.63 to 0.94); spec = 0.35 (0.14 to 0.62)	Insufficient evidence to draw meaningful conclusions
IL-4 (interleukin - 4) cut-off threshold ≥ 3 pg/ml follicular cycle phase	50 (1)	21	6	11	12	Sens = 0.64 (0.45 to 0.80); spec = 0.65 (0.38 to 0.86)	Insufficient evidence to draw meaningful conclusions

rASRM I-IV							
IL-6 (interleukin - 6)[a$] cut-off threshold > 1.03 pg/ml follicular or luteal cycle phase rASRM I-IV	138 (1)	55	34	36	13	Sens = 0.81 (0.70 to 0.89); spec = 0.51 (0.39 to 0.64)	Insufficient evidence to draw meaningful conclusions
IL-6 (interleukin - 6)[a$, a^] cut-off threshold > 1.9-2.0 pg/ml cycle phase varied rASRM I-IV	309 (3)	107	43	97	62	Sens = 0.63 (0.52 to 0.75); spec = 0.69 (0.57 to 0.82)	Summary estimates did not meet the predetermined criteria for a triage or replacement test; varying cycle phase across the studies
IL-6 (interleukin - 6)[a$] cut-off threshold > 2.6 pg/ml follicular or luteal cycle phase rASRM I-IV	138 (1)	41	21	49	27	Sens = 0.60 (0.48 to 0.72); spec = 0.70 (0.58 to 0.80)	Insufficient evidence to draw meaningful conclusions
IL-6 (interleukin - 6)[a^] cut-off threshold > 4 pg/ml menstrual cycle phase rASRM I-IV	91 (1)	48	7	28	8	Sens = 0.86 (0.74 to 0.94); spec = 0.80 (0.63 to 0.92)	Insufficient evidence to draw meaningful conclusions
IL-6 (interleukin - 6)[a^] cut-off threshold > 7.5 pg/ml menstrual cycle phase rASRM I-IV	91 (1)	45	5	30	11	Sens = 0.80 (0.68 to 0.90); spec = 0.86 (0.70 to 0.95)	Insufficient evidence to draw meaningful conclusions
IL-6 (interleukin - 6) cut-off threshold < 10 pg/ml follicular cycle phase	45 (1)	20	14	3	8	Sens = 0.71 (0.51 to 0.87); spec = 0.18 (0.04 to 0.43)	Insufficient evidence to draw meaningful conclusions

rASRM I-IV[b]							
IL-6 (interleukin - 6) cut-off threshold > 12.2 pg/ml follicular cycle phase rASRM I-IV	95 (1)	62	5	25	3	Sens = 0.95 (0.87 to 0.99); spec = 0.83 (0.65 to 0.94)	Insufficient evidence to draw meaningful conclusions; meets criteria for a replacement and SnOUT triage test; further diagnostic test accuracy studies recommended
IL-6 (interleukin - 6) cut-off threshold > 15.4 pg/ml follicular cycle phase rASRM I-II	78 (1)	34	7	33	4	Sens = 0.89 (0.75 to 0.97); spec = 0.82 (0.67 to 0.93)	Insufficient evidence to draw meaningful conclusions
IL-6 (interleukin - 6) cut-off threshold > 25.75 pg/ml follicular cycle phase rASRM I-II	84 (1)	8	12	60	3	Sens = 0.73 (0.39 to 0.94); spec = 0.83 (0.73 to 0.91)	Insufficient evidence to draw meaningful conclusions
IL-6 (interleukin - 6) cut-off threshold not specified luteal cycle phase rASRM I-IV	116 (1)	46	9	29	32	Sens = 0.59 (0.47 to 0.70); spec = 0.76 (0.60 to 0.89)	Insufficient evidence to draw meaningful conclusions
IL-8 (interleukin - 8) cut-off threshold > 24 pg/ml menstrual cycle phase rASRM I-IV	101 (1)	31	14	37	19	Sens = 0.76 (0.60 to 0.89); spec = 0.73 (0.58 to 0.84)	Insufficient evidence to draw meaningful conclusions
IL-8 (interleukin - 8) cut-off threshold > 25 pg/ml follicular or luteal cycle phase	91 (1)	50	4	17	20	Sens = 0.71 (0.59 to 0.82); spec = 0.81 (0.58 to 0.95)	Insufficient evidence to draw meaningful conclusions

rASRM III-IV[d]							
IL-8 (interleukin - 8) cut-off threshold not specified luteal cycle phase rASRM I-IV	116 (1)	38	11	27	40	Sens = 0.49 (0.37 to 0.60); spec = 0.71 (0.54 to 0.85)	Insufficient evidence to draw meaningful conclusions
7. Other peptides and proteins shown to influence key events implicated in endometriosis							
Follistatin cut-off threshold > 1433 pg/ml follicular cycle phase rASRM III-IV[d]	104 (1)	48	4	48	4	Sens = 0.92 (0.81 to 0.98); spec = 0.92 (0.81 to 0.98)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a replacement and SnOUT or SpIN triage test; further diagnostic test accuracy studies recommended
STX-5 (syntaxin - 5) cut-off threshold > 55 ng/ml cycle phase not specified rASRM I-IV	80 (1)	47	6	14	13	Sens = 0.78 (0.66 to 0.88); spec = 0.70 (0.46 to 0.88)	Insufficient evidence to draw meaningful conclusions
8. Oxidative stress markers							
Carbonyls cut-off threshold < 14.9 µM cycle phase not specified rASRM stage not reported	108 (1)	63	20	21	4	Sens = 0.94 (0.85 to 0.98); spec = 0.51 (0.35 to 0.67)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SnOUT triage test; further diagnostic test accuracy studies recommended
PON-1 (paraoxonase-1) cut-off threshold < 141.5 U/l follicular cycle phase rASRM I-IV	87 (1)	46	8	32	1	Sens = 0.98 (0.89 to 1.00); spec = 0.80 (0.64 to 0.91)	Insufficient evidence to draw meaningful conclusions; meets criteria for a replacement or SnOUT triage test; further diagnostic test accuracy studies recommended

Thiols cut-off threshold < 396.44 µM cycle phase not specified rASRM stage not reported	108 (1)	49	8	33	18	Sens = 0.73 (0.61 to 0.83); spec = 0.80 (0.65 to 0.91)	Insufficient evidence to draw meaningful conclusions
9. Post-transcriptional regulators of gene expression (microRNAs)							
miR-9* cut-off threshold not specified follicular or luteal cycle phase rASRM I-IV	85 (1)	41	1	24	19	Sens = 0.68 (0.55 to 0.80); spec = 0.96 (0.80 to 1.00)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
miR-17-5 cut-off threshold < 0.9057 follicular or luteal cycle phase rASRM III-IV	40 (1)	14	6	14	6	Sens = 0.70 (0.46 to 0.88); spec = 0.70 (0.46 to 0.88)	Insufficient evidence to draw meaningful conclusions
miR-20a cut-off threshold < 0.6879 follicular or luteal cycle phase rASRM III-IV	40 (1)	12	2	18	8	Sens = 0.60 (0.36 to 0.81); spec = 0.90 (0.68 to 0.99)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
miR-22 cut-off threshold < 0.5647 follicular or luteal cycle phase rASRM III-IV	40 (1)	18	4	16	2	Sens = 0.90 (0.68 to 0.99); spec = 0.80 (0.56 to 0.94)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a replacement or SnOUT triage test; further diagnostic test accuracy studies recommended
miR-122 cut-off threshold not specified follicular or luteal cycle phase rASRM I-IV	85 (1)	48	6	19	12	Sens = 0.80 (0.68 to 0.89); spec = 0.76 (0.55 to 0.91)	Insufficient evidence to draw meaningful conclusions

miR-141* cut-off threshold not specified follicular or luteal cycle phase rASRM I-IV	85 (1)	43	1	24	17	Sens = 0.72 (0.59 to 0.83); spec = 0.96 (0.80 to 1.00)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
miR-145* cut-off threshold not specified follicular or luteal cycle phase rASRM stage not reported	85 (1)	42	1	24	18	Sens = 0.70 (0.57 to 0.81); spec = 0.96 (0.80 to 1.00)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
miR-199a cut-off threshold not specified follicular or luteal cycle phase rASRM I-IV	85 (1)	47	6	19	13	Sens = 0.78 (0.66 to 0.88); spec = 0.76 (0.55 to 0.91)	Insufficient evidence to draw meaningful conclusions
miR-532-3p cut-off threshold not specified follicular or luteal cycle phase rASRM I-IV	85 (1)	48	2	23	12	Sens = 0.80 (0.68 to 0.89); spec = 0.92 (0.74 to 0.99)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
10. Tumour markers							
CA-15.3 (cancer antigen-15.3) cut-off threshold > 15.04 U/ml cycle phase not specified rASRM I-IV	88 (1)	33	14	23	18	Sens = 0.65 (0.50 to 0.78); spec = 0.62 (0.45 to 0.78)	Insufficient evidence to draw meaningful conclusions
CA-15.3 (cancer antigen-15.3) cut-off threshold > 30 U/ml luteal cycle phase	119 (1)	3	3	35	78	Sens = 0.04 (0.01 to 0.10); spec = 0.92 (0.79 to 0.98)	Insufficient evidence to draw meaningful conclusions

rASRM I-IV							
CA-19.9 (cancer antigen-19.9)[a#] cut-off threshold > 7.5 IU/ml luteal cycle phase rASRM I-IV[b]	76 (1)	32	14	18	12	Sens = 0.73 (0.57 to 0.85); spec = 0.56 (0.38 to 0.74)	Insufficient evidence to draw meaningful conclusions
CA-19.9 (cancer antigen-19.9)[a#] cut-off threshold >9.5 IU/ml any cycle phase rASRM I-IV[b]	198 (1)	64	34	47	53	Sens = 0.55 (0.45 to 0.64); spec = 0.58 (0.47 to 0.69)	Insufficient evidence to draw meaningful conclusions
CA-19.9 (cancer antigen-19.9) cut-off threshold > 10.67 IU/ml cycle phase not specified rASRM I-IV	88 (1)	33	14	23	18	Sens = 0.65 (0.50 to 0.78); spec = 0.62 (0.45 to 0.78)	Insufficient evidence to draw meaningful conclusions
CA-19.9 (cancer antigen-19.9) cut-off threshold ≥ 12 IU/ml follicular cycle phase rASRM III-IV[d]	119 (1)	24	24	55	15	Sens = 0.62 (0.45 to 0.77); spec = 0.70 (0.58 to 0.79)	Insufficient evidence to draw meaningful conclusions
42.5. CA-19.9 (cancer antigen-19.9) cut-off threshold > 37 IU/ml cycle phase varied (not specified in 2 studies) rASRM I-IV	330 (3)	88	11	72	159	Summary estimates: Sens = 0.36 (0.26 to 0.45); spec = 0.87 (0.75 to 0.99)	Summary estimates did not meet the predetermined criteria for a triage or replacement test; varying cycle phase across the studies
CA-19.9 (cancer antigen-19.9)[a#] cut-off threshold not specified follicular cycle phase	60 (1) ---------- 116 (1)	21 ---------- 28	2 ---------- 11	18 ---------- 27	19 ---------- 50	Sens = 0.53 (0.36 to 0.68); spec = 0.90 (0.68 to 0.99)	Insufficient evidence to draw meaningful conclusions; varying populations across the studies; unclear thresholds

rASRM stage not reported[d,e] --- luteal cycle phase rASRM I-IV						--- Sens = 0.36 (0.25 to 0.48); spec = 0.71 (0.54 to 0.85)	
CA-72 (TAG-72) (cancer antigen-72 or tumour associated glycoprotein-72) cut-off threshold > 4 U/ml follicular cycle phase rASRM stage not reported	35 (1)	1	4	12	18	Sens = 0.05 (0.00 to 0.26); spec = 0.75 (0.48 to 0.93)	Insufficient evidence to draw meaningful conclusions
CA-72 (TAG-72) (cancer antigen-72 or tumour associated glycoprotein-72) cut-off threshold > 6 U/ml luteal cycle phase rASRM I-IV	119 (1)	7	4	34	74	Sens = 0.09 (0.04 to 0.17); spec = 0.89 (0.75 to 0.97)	Insufficient evidence to draw meaningful conclusions
CA-125 (cancer antigen-125)[a!, a%, a*] cut-off threshold > 10-14.7 U/ml cycle phase varied rASRM stage varied 2 evaluations excluded as overlapping populations (CA-125 cut-off > 11.5 U/ml and cut-off > 13.5 U/ml, Vodolazkaia 2012)	733 (5)	329	174	155	129	Summary estimates: Sens = 0.70 (0.63 to 0.77); spec = 0.64 (0.47 to 0.82)	Summary estimates do not meet the predetermined criteria for a triage or replacement test
CA-125 (cancer antigen-125)[a!] cut-off threshold > 11.5 U/ml follicular cycle phase	45 (1)	24	6	11	4	Sens = 0.86 (0.67 to 0.96); spec = 0.65 (0.38 to 0.86)	Insufficient evidence to draw meaningful conclusions

rASRM I-IV[b] (excluded from the above group as overlapping evaluation)							
CA-125 (cancer antigen-125) [a!] cut-off threshold > 13.5 U/ml luteal cycle phase rASRM I-IV[b] (excluded from the above group as overlapping evaluation)	35 (1)	15	11	5	4	Sens = 0.79 (0.54 to 0.94); spec = 0.31 (0.11 to 0.59)	Insufficient evidence to draw meaningful conclusions
CA-125 (cancer antigen-125)[a#] cut-off value > 16-17.6 U/ml cycle phase varied (not specified in 2 studies) rASRM stage varied (I in 1 study, I-IV in 4 studies)	430 (5)	146	17	154	113	Summary estimates: Sens = 0.56 (0.24 to 0.88); spec = 0.91 (0.75 to 1.00)	Summary estimates approach the criteria for a SpIN triage test; varying populations across the studies
CA-125 (cancer antigen-125)[a@, a^, a&, a*, a!!] cut-off value > 20 IU/ml cycle phase varied rASRM stage varied (1 study[c], 2 studies[d])	1304 (6)	504	200	361	239	Summary estimates: Sens = 0.67 (0.50 to 0.85); spec = 0.69 (0.58 to 0.80)	Summary estimates do not meet the predetermined criteria for a triage or replacement test; varying populations across the studies
CA-125 (cancer antigen-125)[a^, a&] cut-off value > 25-26 U/ml cycle phase varied; not specified in 1 study rASRM stage varied (1 study[d])	963 (3)	373	137	314	139	Summary estimates: Sens = 0.73 (0.67 to 0.79); spec = 0.70 (0.63 to 0.77)	Summary estimates do not meet the predetermined criteria for a triage or replacement test; varying populations across the studies
CA-125 (cancer antigen-125)[a$, a&]	1206 (6)	417	103	411	275	Summary estimates:	Summary estimates do not meet the predetermined criteria for a triage or replacement test; vary-

cut-off value > 30-33 U/ml (1 study > 33 U/ml) cycle phase varied (not specified in 2 studies) rASRM stage varied (2 studies[d])						Sens = 0.62 (0.45 to 0.79); spec = 0.76 (0.53 to 1.00)	ing populations across the studies
CA-125 (cancer antigen-125)[a@, a#, a$, a%, a&, a!!] cut-off value > 35-36 U/ml (1 study > 36 U/ml) cycle phase varied; not specified in 7 studies rASRM stage varied; not reported in 2 studies (1 study[c], 2 studies[d], 1 study[e])	3447 (27)	895	169	1281	1102	Summary estimates: Sens = 0.40 (0.32 to 0.49); spec = 0.91 (0.88 to 0.94)	Summary estimates do not meet the predetermined criteria for a triage or replacement test; varying populations across the studies
CA-125 (cancer antigen-125);[a$] cut-off value > 42 U/ml follicular cycle phase rASRM III-IV[d]	104 (1)	23	5	47	29	Sens = 0.44 (0.30 to 0.59); spec = 0.90 (0.79 to 0.97)	Insufficient evidence to draw meaningful conclusions
CA-125 (cancer antigen-125) cut-off value > 43 U/ml cycle phase not reported rASRM III-IV	63 (1)	42	4	16	0	Sens = 1.00 (0.92 to 1.00); spec = 0.80 (0.56 to 0.94)	Insufficient evidence to draw meaningful conclusions; meets criteria for a replacement and SnOUT triage test; further diagnostic test accuracy studies recommended
CA-125 (cancer antigen-125) cut-off value not specified menstrual cycle phase[a##] rASRM I-IV --- follicular cycle phase[a##]	59 (1) --- 119 (1) --- 60 (1) ---	29 --- 54 --- 33 ---	4 --- 10 --- 2 ---	15 --- 26 --- 18 ---	11 --- 29 --- 7 ---	Sens = 0.72 (0.56 to 0.85); spec = 0.79 (0.54 to 0.94) --- Sens = 0.65 (0.54 to 0.75);	Insufficient evidence to draw meaningful conclusions; 1 study approaches criteria for a SpIN triage test; further diagnostic test accuracy studies recommended with defined cut-off value; varying populations and undefined

rASRM I-IV --- follicular cycle phase rASRM stage not reported[d,e] --- luteal cycle phase[a##] rASRM I-IV	116 (1)	53	11	27	25	spec = 0.72 (0.55 to 0.86) --- Sens = 0.82 (0.67 to 0.93); spec = 0.90 (0.68 to 0.99) --- Sens = 0.68 (0.56 to 0.78); spec = 0.71 (0.54 to 0.85)	cut-off values; not combined in meta-analysis
11. Combined test - 2 blood biomarkers							
CA-125 +/or CA-19.9 U/ml[a#] cut-of threshold CA-125 ≥ 25 U/ml; CA-19.9 ≥ 12 U/ml follicular cycle phase rASRM III-IV[d] combined test by ROC analysis	118 (1)	35	47	32	4	Sens = 0.90 (0.76 to 0.97); spec = 0.41 (0.30 to 0.52)	Insufficient evidence to draw meaningful conclusions
CA-125 + CA-19.9 U/ml[a#] cut-of threshold CA-125 ≥ 25 U/ml; CA-19.9 ≥ 12 U/ml follicular cycle phase rASRM III-IV[d] combined test by ROC analysis	118 (1)	21	8	71	18	Sens = 0.54 (0.37 to 0.70); spec = 0.90 (0.81 to 0.96)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
CA-125 + Prolactin[a$] cut-off threshold CA-125 > 19.8 U/l; Prolactin > 14.8 ng/ml luteal cycle phase	97 (1)	49	4	30	14	Sens = 0.78 (0.66 to 0.87); spec = 0.88 (0.73 to 0.97)	Insufficient evidence to draw meaningful conclusions

rASRM I-IV[c] combined test by ROC analysis							
CA-125 + Prolactin[a$] cut-off threshold CA-125 > 35 U/l; Prolactin > 20 ng/ml luteal cycle phase rASRM I-IV[c] combined test by ROC analysis	97 (1)	28	1	33	35	Sens = 0.44 (0.32 to 0.58); spec = 0.44 (0.32 to 0.58)	Insufficient evidence to draw meaningful conclusions
CA-125 + VEGF cut-off threshold CA-125 > 17.6 U/ml; VEGF > 236 pg/ml follicular cycle phase rASRM I-IV combined test by ROC analysis	95 (1)	50	2	28	15	Sens = 0.77 (0.65 to 0.86); spec = 0.93 (0.78 to 0.99)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
CA-125 + anti-endometrial Abs cut-off threshold CA-125 > 20 U/l; anti-endometrial Abs > 0.3 A-value luteal cycle phase rASRM I-IV selection or classification method not reported	42 (1)	17	3	11	11	Sens= 0.61 (0.41 to 0.78); spec = 0.79 (0.49 to 0.95)	Insufficient evidence to draw meaningful conclusions
CA-125 x NLR cut-off threshold > 43.1 menstrual cycle phase rASRM I-IV combined test ROC analysis	100 (1)	40	7	43	10	Sens = 0.80 (0.66 to 0.90); spec = 0.86 (0.73 to 0.94)	Insufficient evidence to draw meaningful conclusions

CA-125 +/or IL-8 cut-off threshold CA-125 > 30 U/ml; IL-8 ≥ 25 pg/ml follicular or luteal cycle phase rASRM III-IV[d] combined test ROC analysis	83 (1)	56	5	13	9	Sens = 0.86 (0.75 to 0.93); spec = 0.72 (0.47 to 0.90)	Insufficient evidence to draw meaningful conclusions
CA-125 + IL-8 cut-off threshold not specified any cycle phase rASRM I-IV combined test by multivariate analysis using stepwise logistic regression and by ROC analysis	294 (1)	143	27	58	66	Sens = 0.71 (0.64 to 0.77); spec= 0.71 (0.61 to 0.80)	Insufficient evidence to draw meaningful conclusions
IL-6 + TNF-α cut-off threshold IL-6 > 12.2 pg/ml; TNF-α > 12.45 pg/ml follicular cycle phase rASRM I-IV combined test ROC analysis	96 (1)	46	0	30	20	Sens = 0.70 (0.57 to 0.80); spec = 1.00 (0.88 to 1.00)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
IL-6 + CRP cut-off threshold IL-6 >12.2 pg/ml; CRP > 438 µg/ml follicular cycle phase rASRM I-IV combined test by ROC analysis	95 (1)	49	0	30	16	Sens = 0.75 (0.63 to 0.85); spec = 1.00 (0.88 to 1.00)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
TNF-α + CRP cut-off threshold NF-α > 12.45 pg/ml; CRP > 438 µg/ml	95 (1)	48	0	30	17	Sens = 0.74 (0.61 to 0.84);	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test; further

follicular cycle phase rASRM I-IV combined test by ROC analysis						spec = 1.00 (0.88 to 1.00)	diagnostic test accuracy studies recommended
miR-199a + miR-542-3p cut-off threshold not specified follicular or luteal cycle phase rASRM I-IV combined test by discriminant and ROC analysis	85 (1)	58	3	22	2	Sens = 0.97 (0.88 to 1.00); spec = 0.88 (0.69 to 0.97)	Insufficient evidence to draw meaningful conclusions; meets criteria for a replacement and SnOUT triage test; further diagnostic test accuracy studies recommended
miR-199a + miR-122 cut-off threshold not specified follicular or luteal cycle phase rASRM I-IV combined test by discriminant and ROC analysis	85 (1)	48	5	20	12	Sens = 0.80 (0.68 to 0.89); spec = 0.80 (0.59 to 0.93)	Insufficient evidence to draw meaningful conclusions
12. Combined test - 3 blood biomarkers							
CA-125 + CA-19-9 + survivin cut-off threshold not specified follicular cycle phase rASRM stage not reported[e] combined test by logistic regression and ROC analysis	60 (1)	35	2	18	5	Sens = 0.88 (0.73 to 0.96); spec = 0.90 (0.68 to 0.99)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
CA-125 + STX-5 + LN-1 cut-off threshold not specified cycle phase not specified rASRM I-IV	80 (1)	57	6	14	3	Sens = 0.95 (0.86 to 0.99); spec = 0.70 (0.46 to 0.88)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SnOUT triage test and approaches criteria for a replacement test; further diagnostic test accuracy studies recommended

combined test by multivariate logistic regression and ROC analysis							
CA-125 +/or CA-19.9 +/or IL-6 cut-off threshold CA-125 > 35 U/ml; CA-19.9 > 37 U/ml; IL-6 > 2 pg/ml any cycle phase rASRM I-IV combined test by ROC analysis	80 (1)	19	10	25	26	Sens = 0.42 (0.28 to 0.58); spec = 0.71 (0.54 to 0.85)	Insufficient evidence to draw meaningful conclusions
CA-125 +/ or CCR1 +/or MCP-1 CA-125 > 50 U/ml; CCR1 > 1.16; MCP-1 > 140 pg/ml follicular cycle phase rASRM I-IV selection or classification method not reported	151 (1)	94	9	40	8	Sens = 0.92 (0.85 to 0.97); spec = 0.82 (0.68 to 0.91)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a replacement and SnOUT triage test; further diagnostic test accuracy studies recommended
CA-125 + MCP-1 + Leptin cut-off threshold CA-125 > 20 U/ml; MCP-1 > 152.7 pg/ml; Leptin > 3.14 ng/ml any cycle phase rASRM II-IV combined test by a two-tiered algorithm using classification and regression tree (CART)	141 (1)	31	5	73	32	Sens = 0.49 (0.36 to 0.62); spec = 0.94 (0.86 to 0.98)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
CA-125 + IL-8 + TNF-α cut-off threshold not specified luteal cycle phase rASRM I-IV	116 (1)	70	11	27	8	Sens = 0.90 (0.81 to 0.95); spec = 0.71 (0.54 to 0.85)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SnOUT triage test; further diagnostic test accuracy studies recommended

combined test by multivariate analysis using stepwise logistic regression and ROC analysis							
IL-6 + TNF-α + CRP cut-off threshold IL-6 > 12.2 pg/ml; TNF-α > 12.45 pg/ml; CRP > 438 µg/ml follicular cycle phase rASRM I-IV combined test by ROC analysis	95 (1)	41	0	30	24	Sens = 0.63 (0.50 to 0.75); spec = 1.00 (0.88 to 1.00)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test; further diagnostic test accuracy studies recommended
13. Combined test - 4 blood biomarkers							
CA-125 + VEGF + annexin V + glycodelin[a#] cut-off threshold not specified menstrual cycle phase rASRM I-IV[b] combined test by multivariate logistic regression and ROC analysis	19 (1)	9	2	6	2	Sens = 0.82 (0.48 to 0.98); spec = 0.75 (0.35 to 0.97)	Insufficient evidence to draw meaningful conclusions
CA-125 + VEGF + annexin V + glycodelin[a#] cut-off threshold not specified menstrual cycle phase rASRM I-IV[b] combined test by a least squares support vector machines model (LS-SVM) and ROC analysis	19 (1)	9	3	5	2	Sens = 0.82 (0.48 to 0.98); spec = 0.63 (0.24 to 0.91)	Insufficient evidence to draw meaningful conclusions
CA-125 + VEGF + annexin V + sICAM-1 cut-off threshold not specified	19 (1)	9	2	6	2	Sens = 0.82 (0.48 to 0.98);	Insufficient evidence to draw meaningful conclusions

menstrual cycle phase rASRM I-IV[b] combined test by either multivariate logistic regression or a least squares support vector machines model (LS-SVM) and ROC analysis						spec = 0.75 (0.35 to 0.97)	
CA-125 + MCP-1 + Leptin + MIF cut-off threshold CA-125 > 20 U/ml; MCP-1 > 53.5 pg/ml; Leptin > 29.1 ng/ml; MIF > 14.7 ng/ml any cycle phase rASRM II-IV combined test by a two-tiered algorithm using classification and regression tree (CART)	141 (1)	63	51	27	0	Sens = 1.00 (0.94 to 1.00); spec = 0.35 (0.24 to 0.46)	Insufficient evidence to draw meaningful conclusions
miR-199a + miR-122 + miR-145* + miR-542-3p cut-off threshold not specified follicular or luteal cycle phase rASRM I-IV combined test by discriminant and ROC analysis	85 (1)	56	1	24	4	Sens = 0.93 (0.84 to 0.98); spec = 0.96 (0.80 to 1.00)	Insufficient evidence to draw meaningful conclusions; meets criteria for a SpIN triage test and approaches criteria for a replacement test; further diagnostic test accuracy studies recommended
14. Combined test - 6 blood biomarkers							
CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP[a$] cut-off threshold not specified any cycle phase rASRM I-IV	295 (1)	181	44	49	20	Sens = 0.90 (0.85 to 0.94); spec = 0.53 (0.42 to 0.63)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a SnOUT triage test; further diagnostic test accuracy studies recommended

combined test by multivariate analysis using stepwise logistic regression and ROC analysis							
CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP [a$] cut-off threshold not specified menstrual cycle phase rASRM I-IV combined test by multivariate analysis using stepwise logistic regression and ROC analysis	59 (1)	36	5	14	4	Sens = 0.90 (0.76 to 0.97); spec = 0.74 (0.49 to 0.91)	Insufficient evidence to draw meaningful conclusions; approaches criteria for a replacement and SnOUT triage test; further diagnostic test accuracy studies recommended
CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP[a$] cut-off threshold not specified follicular cycle phase rASRM I-IV combined test by multivariate analysis using stepwise logistic regression and ROC analysis	119 (1)	48	10	26	35	Sens = 0.58 (0.46 to 0.69); spec = 0.72 (0.55 to 0.86)	Insufficient evidence to draw meaningful conclusions
CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP[a$] cut-off threshold not specified luteal cycle phase rASRM I-IV combined test by multivariate analysis using stepwise logistic regression and ROC analysis	116 (1)	67	11	27	11	Sens = 0.86 (0.76 to 0.93); spec = 0.71 (0.54 to 0.85)	Insufficient evidence to draw meaningful conclusions

[a] Same biomarker was tested in the same/overlapping cohort; similar symbol designates studies/groups of studies with overlapping cohorts, hence can not be combined in meta-analysis

[b] Only for US-negative endometriosis

[c] Only peritoneal endometriosis

[d] Only ovarian endometriosis versus other benign ovarian cysts

[e] Only deep infiltrating endometriosis or endometrioma + deep infiltrating endometriosis

MW: molecular weight; **rASRM**: revised American Society for Reproductive Medicine; **ROC**: receiver operating characteristic
For a comprehensive list of all biomarkers with their biological annotation, please see Appendix 1.

BACKGROUND

Target condition being diagnosed

Endometriosis

Endometriosis is defined as an inflammatory condition characterised by endometrial-like tissue at sites outside the uterus (Johnson 2013). Endometriotic lesions can occur at different locations, including the pelvic peritoneum and the ovary, or they can penetrate pelvic structures below the surface of the peritoneum (defined as deeply infiltrating endometriosis, or DIE). Current knowledge suggests that each of these types of endometriosis is a separate clinical entity, but they can coexist in the same woman. Rarely, endometriotic implants can be found at more distant sites, including the lung, liver, pancreas and operative scars, with consequent variations in presenting symptoms.

Endometriosis afflicts 10% of reproductive-aged women, causing dysmenorrhoea (painful periods), dyspareunia (painful intercourse), chronic pelvic pain and infertility (Vigano 2004). The clinical presentation can vary from asymptomatic and unexplained infertility to severe dysmenorrhoea and chronic pain. These symptoms can occur with bowel or urinary symptoms, an abnormal pelvic examination or the presence of a pelvic mass; however, no symptom is specific to endometriosis. The estimated prevalence of endometriosis in the symptomatic population is 35% to 50% (Giudice 2004).

Women with endometriosis are also at increased risk of developing several cancers and autoimmune disorders (Sinaii 2002; Somigliana 2006). The presence of disease is associated with changes in the immune response, vascularisation, neural function, the peritoneal environment and the eutopic endometrium (tissue lining the uterine cavity), suggesting that endometriosis is a systemic rather than localised condition (Giudice 2004). Endometriosis has a profound effect on psychological and social well-being and imposes a substantial economic burden on society. Women with endometriosis may incur significant direct medical expenses from diagnostic and therapeutic surgeries, hospital admissions and fertility treatments, while indirect costs, including absenteeism and loss of productivity, compound the economic impact (Gao 2006; Simoens 2012). In the United States, the financial burden of endometriosis is about USD 12,419 per woman (Simoens 2012).

Although research has not been able to fully elucidate the pathogenesis of endometriosis, specialists commonly believe that it occurs when endometrial tissue contained within the menstrual fluid implants at an ectopic site within the pelvic cavity through retrograde flow (Sampson 1927). However, this theory does not explain the fact that only 10% of women develop endometriosis, while retrograde menstruation occurs in up to 90% of women (Halme 1984). There is evidence that a variety of environmental, immunological and hormonal factors are associated with endometriosis and genetic loci that confer a risk of endometriosis, but the relative contribution of these and other causal factors is still unclear (Nyholt 2012; Vigano 2004).

Although it is impossible to time the onset of disease, on average, women have a 6- to 12-year history of symptoms before obtaining a surgical diagnosis, indicative of considerable diagnostic delay (Matsuzaki 2006). Untreated endometriosis is associated with reduced quality of life and contributes to outcomes such as depression, inability to work, sexual dysfunction and missed opportunity for motherhood (Gao 2006). Since endometriosis is a progressive disease in up to 50% of women, early diagnosis has the potential to offer early treatment and prevent progression (D'Hooghe 2002).

Treatment of endometriosis

There is no cure for endometriosis. Treatment options include expectant management, pharmacological (hormonal) therapy and surgery (Johnson 2013). Treatment is individualised, taking into consideration a therapeutic goal (pain relief or conception) and the location of the disease. Current pharmacological therapies such as the combined oral contraceptive pill, progestogens, weak androgens and GnRH agonists and antagonists act to reduce the effect of oestrogen on endometrial tissues and suppress menstruation. These drugs can ameliorate the symptoms of dysmenorrhoea and chronic pelvic pain but are associated with side effects such as breast discomfort, irritability, androgenic symptoms and bone loss. Surgical excision of endometriotic lesions can reduce pain symptoms, but it is associated with high recurrence rates of 40% to 50% at five years postsurgery (Guo 2009). Early treatment of endometriosis improves pain levels as well as physical and psychological functioning. Furthermore, improvements in menstrual management (the use of the intrauterine system (hormonal coil) and the continuous use of the combined contraceptive pill) and fertility preservation (oocyte vitrification) raise the possibility of suppressing the progression of endometriosis and prospectively managing subfertility in endometriosis sufferers. The potential success of these preventive strategies depends on an accurate and early diagnosis. A major impediment to earlier and more efficacious treatment of this disease is diagnostic delay, due to the invasive nature of standard diagnostic tests (Dmowski 1997).

Diagnosis of endometriosis

Clinical history and pelvic examination can raise the possibility of a diagnosis of endometriosis, but the heterogeneity in clinical presentation, the high prevalence of asymptomatic endometriosis (2% to 50%) and the poor association between presenting symptoms and severity of the disease contribute to the difficulty in obtaining a reliable diagnosis based solely on presenting symptoms (Ballard 2008; Fauconnier 2005; Spaczynski 2003). Although an abnormal pelvic examination correlates with the presence of endometriosis on laparoscopy in 70% to 90% of cases (Ling 1999), there is a wide differential diagnosis for most positive physical findings. Furthermore, a normal clinical examination does not exclude endometriosis, as laparoscopically proven disease has been diagnosed in more than 50% of women with a clinically normal pelvic examination (Eskenazi 2001). A variety of tests utilising pelvic imaging, blood markers, eutopic endometrium characteristics, urinary markers or peritoneal fluid components have been suggested as diagnostic measures for endometriosis. Although large numbers of the reported markers distinguish women with and without endometriosis in small pilot studies, many do not show convincing potential as a diagnostic test when they are evaluated in larger studies by different research groups. The diagnostic value of these tests has not previously been fully systematically evaluated and summarised using Cochrane methods. Currently, there is no simple non-invasive test for the

diagnosis of endometriosis that is routinely implemented in clinical practice.

Surgical diagnostic procedures for endometriosis include laparoscopy (minimal access, or keyhole surgery) or laparotomy (open surgery via an abdominal incision). In the last several decades, laparoscopy has become an increasingly common procedure and has largely replaced traditional open surgery in patients suspected of having endometriosis (Yeung 2009). Laparoscopy has significant advantages over laparotomy, including fewer complications and shorter recovery times. Furthermore, a magnified view at laparoscopy allows better visualisation of the peritoneal cavity. Despite continuing controversy in the literature with regard to the superiority of one surgical modality over another in treating pelvic pathology, laparoscopy is the preferred technique to evaluate the pelvis and abdomen and to treat benign conditions such as ovarian endometriomas (Medeiros 2009). Surgery is currently also the only acceptable method of determining the extent and severity of endometriosis. There are several different classification systems for endometriosis (Adamson 2008; Batt 2003; Chapron 2003a; Martin 2006), but most researchers and clinicians use the revised American Society for Reproductive Medicine (rASRM) classification, which is internationally accepted as a respected tool for the objective assessment of the disease (ASRM 1997). The rASRM classification system considers the appearance, size and depth of peritoneal or ovarian implants and adhesions that are visualised during laparoscopy and allows uniform documentation of the extent of disease (Table 1). Unfortunately, this classification system has little value in clinical practice due to the lack of correlation between laparoscopic staging, the severity of symptoms and response to treatment (Chapron 2003b; Guzick 1997; Vercellini 1996). The World Endometriosis Society has recently undertaken an endeavour to attain consensus around the optimal classification for endometriosis (Johnson 2015).

The European Society of Human Reproduction and Embryology (ESHRE) Special Interest Group for Endometriosis stated in their diagnostic and treatment guidelines that for most forms of endometriosis, women presenting with symptoms cannot obtain a definitive diagnosis without visual inspection of the pelvis at laparoscopy as the gold standard investigation (Kennedy 2005). Currently the visual or histological identification of endometriotic tissue in the pelvic cavity during surgery is not just the best available but the only diagnostic test for endometriosis in clinical practice.

The disadvantages of laparoscopic surgery include (but are not limited to) the high cost, the need for general anaesthesia and the potential for adhesion formation postprocedure. Laparoscopy has been associated with a 2% risk of injury to pelvic organs, a 0.001% risk of damaging a major blood vessel and a mortality rate of 0.0001% (Chapron 2003c). Even though the major complications of laparoscopy are rare, it is difficult to determine the exact incidence of complications, and delayed recognition adds to surgical morbidity and mortality. Only a third of women who undertake a laparoscopic procedure will receive a diagnosis of endometriosis; therefore many disease-free women are unnecessarily exposed to surgical risk (Frishman 2006).

The validity of laparoscopy as a reference test for endometriosis has is highly dependent on the skills of the surgeon. The diagnostic accuracy of laparoscopic visualisation has been compared with histological confirmation in a sole systematic review, and it was estimated as having a sensitivity of 0.94 and specificity of 0.79 (Wykes 2004). Subsequent studies suggested that incorporating histological verification in the diagnosis of endometriosis may improve diagnostic accuracy (Almeida Filho 2008; Marchino 2005; Stegmann 2008), but these papers have not been systematically reviewed. The clinical significance of histological verification remains debatable, and a diagnosis based on visual findings is generally reliable as long as properly trained and experienced surgeons perform an appropriate inspection of the abdominal cavity (Redwine 2003). Furthermore, excised potential endometriotic tissues are rarely serially sectioned in clinical practice, and pathologists can miss small lesions in mild disease. Thus, sampling inconsistencies are also likely to influence the accuracy of histological reporting.

Summary

A diagnostic test without the need for surgery would reduce the associated surgical risks, increase accessibility to a diagnostic test and improve treatment outcomes. The need for an accurate non-invasive diagnostic test for endometriosis continues to encourage extensive research in the field and was endorsed at the international consensus workshop at the 10th World Congress of Endometriosis in 2008 (Rogers 2009). Although multiple markers and imaging techniques have been explored as diagnostic tests for endometriosis, none of them have been implemented routinely in clinical practice, and many have not been subject to a systematic review.

Index test(s)

This review assesses blood-based biomarkers that have been proposed as non-invasive tests for the diagnosis of endometriosis as part of the review series on non-invasive diagnostic tests for endometriosis (Table 2). The other reviews from this series include 'Imaging modalities for the non-invasive diagnosis of endometriosis', 'Endometrial biomarkers for the non-invasive diagnosis of endometriosis', 'Urinary biomarkers for the non-invasive diagnosis of endometriosis', and 'Combination of the non-invasive tests for the diagnosis of endometriosis', which also summarises all the reviews from the series.

The definition of 'non-invasive' varies between medical dictionaries but refers to a procedure that does not involve penetration of skin or physical entrance to the body (McGraw-Hill Dictionary of Medicine 2006; The Gale Encyclopedia of Medicine 2011). Although venipuncture for blood collection is invasive by this definition, blood tests are generally considered to be non-invasive or minimally invasive when compared to diagnostic surgery. For the purpose of these reviews, we will define all tests that do not involve anaesthesia and surgery as non-invasive.

The advantages of using a blood test for the diagnosis of endometriosis are that it is minimally invasive, readily available, acceptable to women, provides a rapid result and is more cost-effective when compared to surgery. However, blood testing is dependent on the reliability of laboratory techniques and quality control protocols. Blood biomarker levels may also be susceptible to variation during the menstrual cycle.

Research has identified cellular and molecular processes that characterise ectopic endometrium and peritoneal fluid in human and animal models (D'Hooghe 2001; Hull 2008; Kao 2003). Different studies have evaluated markers of these

pathophysiological processes in blood samples as a single test or a combination of several biomarkers. Categories of blood markers include: angiogenic and growth factors; markers of apoptosis; cell adhesion molecules and other matrix-related proteins; cytoskeleton molecules; DNA-repair/telomere maintenance molecules; hormonal markers; high-throughput molecular markers; hormonal markers; immune system and inflammatory markers; nerve growth markers; oxidative stress markers; post-transcriptional regulators of gene expression (circulating nuclear DNAs, microRNAs); tumour markers; and other peptides/proteins shown to influence key events implicated in endometriosis. Most blood-based tests have only been evaluated in a limited number of small studies with varying methods, laboratory techniques and types of assay.The most extensively studied biomarker for endometriosis is cancer antigen-125 (CA-125), a glycoprotein expressed on coelomic epithelial tissues such as the peritoneum. An older meta-analysis concluded that CA-125 had a limited ability to diagnose endometriosis (Mol 1998). However, the review did not describe the selection process to include studies. Since then, further studies evaluating CA-125 have been published, and the methodologies of diagnostic test reviews have improved, so an updated review of CA-125 is warranted (Brosens 2003; Bedaiwy 2004; Matalliotakis 2008; Yang 2004).

A large systematic review of all proposed biomarkers for endometriosis in serum, plasma and urine identified over 100 putative biomarkers, but the authors were unable to identify any biomarker (single or in a panel) that they could recommend for use in clinical practice (May 2010). A more recent narrative review concurred with this conclusion (Fassbender 2015). There is a current need to re-evaluate the diagnotic test accuracy of blood tests for endometriosis using Cochrane methods.

Clinical pathway

Women presenting with symptoms of endometriosis (dysmenorrhoea, dyspareunia, chronic pelvic pain or difficulty conceiving) are generally investigated with a pelvic ultrasound scan to exclude other pathologies, which is in line with international guidelines (ACOG 2010; Dunselman 2014; SOGC 2010). There are no other standard investigative tests, and although evidence suggests that MRI is superior to ultrasound, it is used conservatively because of its cost. If patients seek pain management rather than conception, physicians generally initiate empirical treatment with progestogens or the combined oral contraceptive pill. Diagnostic laparoscopy is considered if empirical treatment fails or if women decline or do not tolerate empirical treatment. In women who have difficulty conceiving, laparoscopy can be undertaken before fertility treatment (particularly if severe pelvic pain or endometrioma are present) or after failed assisted reproductive technology (ART) treatments. Physicians may also diagnosis endometriosis during fertility investigations in women who have minimal or no pain symptomatology.

On average there is a delay of 6 to 12 years from onset of symptoms to definitive diagnosis at surgery. Early referral to a gynaecologist with the capability to perform diagnostic surgery is associated with a shorter time to diagnosis. Collectively, young women, women in remote and rural locations and women of lower socioeconomic status have reduced access to surgery and are less likely to obtain a prompt diagnosis of endometriosis.

Prior test(s)

Most women presenting with symptoms suggestive of endometriosis have a full history and examination and a routine gynaecological ultrasound before physicians recommend they undergo diagnostic surgery. However, there is no consensus on whether any other test should be routinely used as part of a standardised approach.

Role of index test(s)

A new diagnostic test can fulfil one of three roles.

1. Replacement: replacing an existing test due to better accuracy or a similar accuracy with other advantages.
2. Triage: used as an initial step in a diagnostic pathway to identify the group of patients who need further testing with an existing test. Although ideally a triage test has a high sensitivity and specificity, it may have a lower sensitivity but higher specificity than the current test or vice versa. The triage test does not aim to improve the diagnostic accuracy of the existing test but rather to reduce the number of individuals having an unnecessary diagnostic test.
3. Add-on: used in addition to existing testing to improve diagnostic performance (Bossuyt 2008).

Ideally a diagnostic test is expected to correctly identify all patients with a disease and to exclude all patients without that disease; in other words, it should have a sensitivity and specificity of 1.00. A high sensitivity indicates that there are a low number of patients who have a negative test and do have the disease (i.e. a low number of false negative results). High specificity corresponds to a low number of patients who have a positive test but do not have the disease (i.e. low false positive results). In practice, however, it is extremely rare to find a test with equally high sensitivity and specificity. An acceptable replacement test would need to have a similar or higher sensitivity and specificity than the current gold standard. In the case of laparoscopy for diagnosis of endometriosis, the only systematic review reported a sensitivity of 0.94 and a specificity of 0.79, and we have taken this as a cut-off for a replacement test (Wykes 2004).

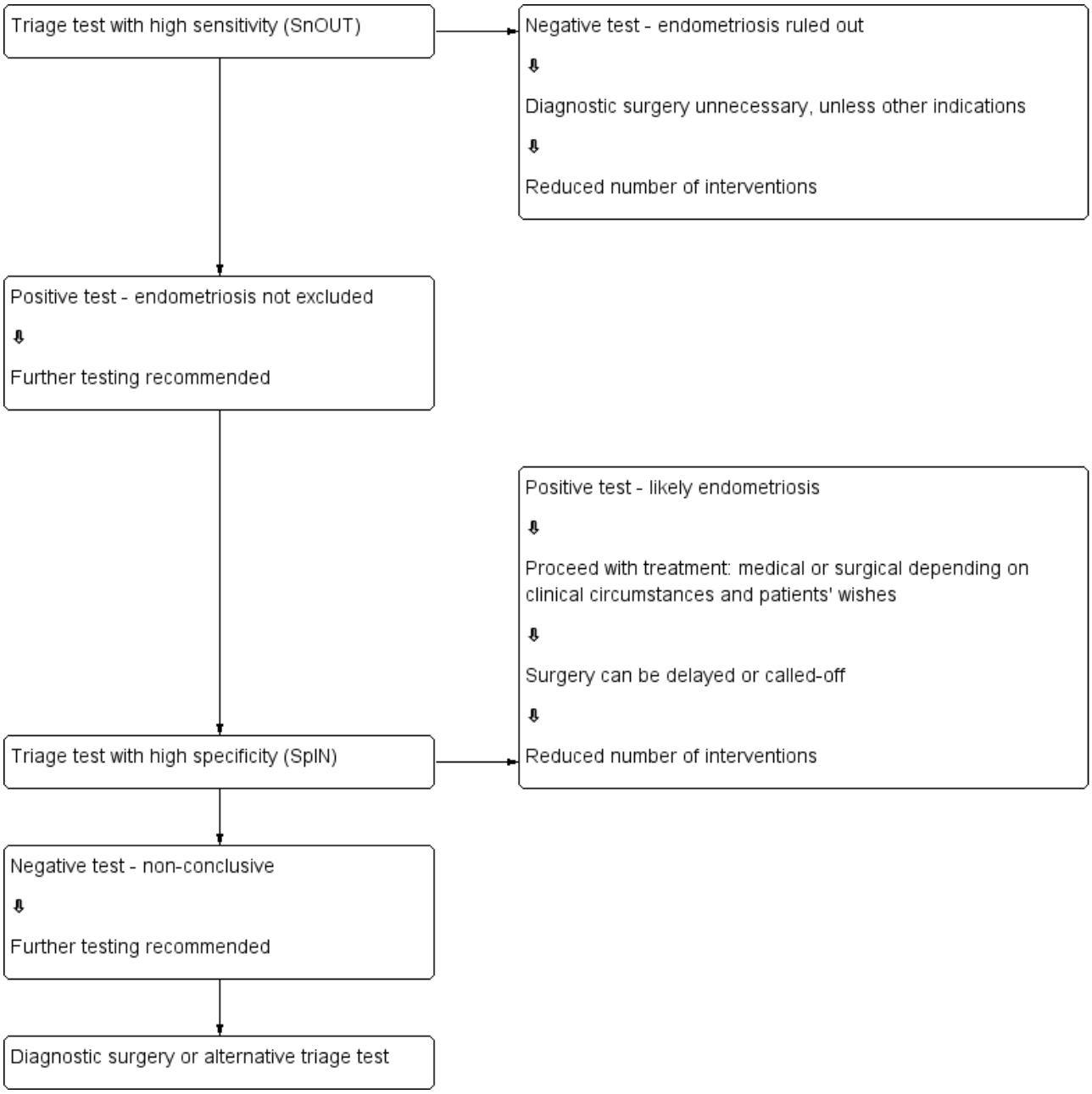
The purpose of triage tests can vary depending on the clinical context and patients' priorities. One reasonable approach is to exclude the diagnosis to avoid further unnecessary and expensive diagnostic investigations. High sensitivity tests have few false negative results and act to rule conditions out (SnOUT). A negative result from a test with high sensitivity will exclude the disease with high certainty independent of the specificity. As women without disease would be assured of having a negative test, unnecessary invasive interventions can be avoided. However, a positive result has less diagnostic value, particularly when the specificity is low. We predetermined that a clinically useful SnOUT triage test should have a sensitivity of 0.95 or more and a specificity of 0.50 and above. We set the sensitivity cut-off for a SnOUT triage test at 0.95 and above, assuming that a 0.05 false negative rate is statistically and clinically acceptable. We set the specificity cut-off at 0.50 and above, to avoid diagnostic uncertainty in more than 50% of the population with a positive result.

An alternative approach would be to avoid a missed diagnosis. High specificity tests have few false positive results and act to rule conditions 'in' (SpIN). A positive result for a highly specific

triage test indicates a high likelihood of having endometriosis. This information could be used to prioritise these patients for surgical treatment. A positive SpIN test could also provide a clinical rationale to start targeted disease-specific medical management in a patient without a surgical diagnosis, under the assumption that disease is present. Surgical management could then be reserved for cases when conservative treatment fails. This is particularly relevant in some populations where the therapeutic benefits of surgery for endometriosis have to be carefully balanced with the disadvantages (e.g. young women, women with medical conditions or pain-free patients with a history of infertility). In this scenario we considered a sensitivity of 0.50 and above and a specificity of 0.95 and higher as suitable cut-offs for a SpIN triage test.

We evaluated blood tests for their potential to replace surgery (replacement test) or to improve the selection of women for surgery (triage test to rule out (SnOUT) or rule in (SpIN) the disease). Both types of triage tests are clinically useful, minimising the number of unnecessary interventions. Sequential implementation of SnOUT and SpIN tests can also optimise a diagnostic algorithm (Figure 1). We did not assess any test as an add-on test, as we sought tests that reduce the need for surgery and not tests that improve the accuracy of the currently available surgical diagnosis.

Figure 1. Sequential approach to non-invasive testing of endometriosis.



Alternative test(s)

There are no routine alternative tests for the diagnosis of endometriosis in clinical practice.

Rationale

Many women with endometriosis suffer longstanding pelvic pain and infertility prior to a diagnosis. Surgery is the only current method of diagnosing endometriosis, but it is associated with high costs and surgical risks. A simple and reliable non-invasive test for endometriosis, with the potential to either replace laparoscopy or to triage women in order to reduce surgery, would minimise surgical risk and reduce diagnostic delay. Physicians could then detect endometriosis at less advanced stages and institute earlier interventions. Early diagnosis would provide the opportunity for a preventive approach for this debilitating disease, potentially reducing healthcare-related costs and favouring more cost-effective and efficient treatments. Furthermore, identifying blood biomarkers that do not pertain to endometriotic disease would help clinicians and researchers focus on clinically relevant biomarker detection.

OBJECTIVES

Primary objectives

To evaluate blood biomarkers as replacement tests for diagnostic surgery and as triage tests to inform decisions to undertake surgery for endometriosis. Specific objectives include the following.

1. To provide summary estimates of the diagnostic accuracy of blood biomarkers for the diagnosis of peritoneal, ovarian and deep infiltrating pelvic endometriosis, compared to surgical diagnosis as a reference standard.
2. To assess the diagnostic utility of biomarkers that could differentiate ovarian endometrioma from other ovarian masses.

Secondary objectives

1. To investigate the influence of heterogeneity on the diagnostic accuracy of blood biomarkers for endometriosis. Potential sources of heterogeneity include:
   a. participant characteristics: age (adolescents versus later reproductive years), clinical presentation (subfertility, pelvic pain, ovarian mass, asymptomatic women), stage of disease (rASRM classification system), geographic location of study;
   b. histological confirmation in conjunction with laparoscopic visualisation compared to laparoscopic visualisation alone;
   c. changes in technology over time: year of publication, modifications applied to conventional laboratory techniques;
   d. methodological quality: differences in the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) evaluation (Table 3), including low versus unclear or high risk; consecutive versus non-consecutive enrolment; and blinding of surgeons to the results of index tests;
   e. study design (single-gate design versus two-gate design studies).
2. To assess biomarkers that were not affected by endometriosis and hence were unlikely to discriminate between women with and without the disease.

METHODS

Criteria for considering studies for this review

Types of studies

Published peer-reviewed studies that compared the results of one or several types of blood biomarker tests with the results obtained from a surgical diagnosis of endometriosis.

We included the following types of studies.

1. Randomised controlled trials (RCTs).
2. Observational studies with the following designs.
   a. Single-gate design (studies with a single set of inclusion criteria defined by clinical presentation). All participants had clinically suspected endometriosis.
   b. Two-gate design (studies where participants are sampled from distinct populations with respect to clinical presentation). The same study includes participants with a clinical suspicion of having the target condition (e.g. women with pelvic pain) and also participants in whom the target condition is not suspected (e.g. women admitted for tubal ligation). Two-gate studies were eligible only where all cases and controls belonged to the same population with respect to the reference standard (i.e. all the participants were scheduled for laparoscopy) (Rutjes 2005).
3. Studies performed on prospectively collected samples, irrespective of the actual time of the test assay. The timing of sample collection relative to surgery is important because the surgical excision of endometriotic lesions could influence blood biomarker expression and hence bias the results. Therefore, we only included studies that drew blood before the surgical procedure, i.e. 'prospectively collected'. We considered to be eligible the studies performed on tissue bank samples collected from prospectively recruited, well-defined populations, which prevented the omission of valuable data from adequately designed studies. The time interval between sample collection and laboratory testing may influence test outcomes, which could be dependent on sample storage conditions and the stability of each individual biomarker during storage and freeze-thawing. This information was not readily available for most molecules, and we did not address it in this review, but we will consider it in future updates if more evidence emerges.

We did not impose limits on eligibility related to the healthcare settings where the study took place, the language of publication, the number of participants in the included studies or the number of studies that evaluated each index test.

We excluded the following types of studies.

1. Narrative or systematic reviews.
2. Studies of retrospective design where investigators collected samples after execution of the reference test.
3. Studies of retrospective design where investigators selected participants from retrospective review of the case notes/ archived samples and where information on recruitment methods or study population was not available.
4. Case reports or case series.
5. Studies reported only in abstract form or in conference proceedings where the full text was not available. We applied

this limitation after facing substantial difficulty in obtaining the information from the abstracts, which precluded a reliable assessment of eligibility and methodological quality.

Participants

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.

Participants came from populations of women undergoing abdominal surgery for the following indications.

- Clinically suspected endometriosis (pelvic pain, infertility, abnormal pelvic examination, or a combination of the above).
- Ovarian mass, regardless of symptoms.
- A mixed group consisting of women with suspected endometriosis/ovarian mass or women with other benign gynaecological conditions (e.g. surgical sterilisation, fibroid uterus, etc).
- Asymptomatic women who have an incidental finding of endometriosis at surgery performed for another indication.

Studies that included participants of postmenopausal age were eligible when the data for the reproductive age group was available in isolation. We excluded studies with participants that clearly would not undergo the index test in the relevant clinical situation or would not benefit from the test (e.g. women with ectopic pregnancies or acute pelvic inflammatory disease). We also excluded publications that only analysed participants with a positive index test or reference standard and did not provide data for the whole cohort.

Index tests

We assessed any type of blood-based biomarker for endometriosis either separately or in combination with other blood tests. We included index tests performed on whole blood, plasma or serum. We present the assessed index tests in Table 2 (classified by biological subgroups) and in Appendix 1 (alphabetical order with annotation for biological subgroups). We included the tests performed in one or several phases of menstrual cycle.

The combined evaluations of blood biomarkers with other methods for diagnosing endometriosis (e.g. pelvic examination, imaging, urine or endometrial tests) are beyond the scope of this review and are presented separately in another review, 'Combined tests for the non-invasive diagnosis of endometriosis'. We excluded studies that solely assessed specific technical aspects, presented qualitative descriptions of lesion appearance or reported interobserver variability of the index tests, without reporting the data on diagnostic performance. When the evaluated biomarker(s) showed differential expression between the groups of women with and without endometriosis, we only considered the study if it reported data with sufficient detail for the construction of 2 x 2 contingency tables. However, when the contingency tables were not available because the expression level of index test did not significantly differ between the groups and the inclusion criteria were otherwise met, we made a critical appraisal and presented the study in the descriptive part of the review. Thus, we evaluated the adequately designed studies that identified biomarkers without diagnostic value, as they provide information that is likely to focus future research on other more clinically useful biomarkers.This methodology also identified biomarkers that were associated with endometriosis in some but not other studies. We did not include evaluations of screening or predictive accuracy tests in this review.

We considered the diagnostic performance of an index test to be high when the test reached the criteria for a replacement test (sensitivity of equal or greater than 0.94 with specificity of equal or greater than 0.79) or triage test (sensitivity of equal or greater than 0.95 with specificity of equal or greater than 0.50 or vice versa) or approached these criteria (diagnostic estimates within 0.05 of the set thresholds). We considered all other diagnostic estimates to be low.

Target conditions

Pelvic endometriosis, defined as endometrial tissue located in the pelvic cavity: involving any of the following: pelvic organs, peritoneum and pouch of Douglas.

We assessed three types of pelvic endometriosis.

1. Peritoneal endometriosis, defined as endometrial deposits detected on peritoneum covering pelvic organs, pelvic side walls or pouch of Douglas.
2. Ovarian endometriosis (endometrioma), defined as an ovarian cyst lined by endometrial tissue, appearing as an ovarian mass of varying size.
3. Deep infiltrating endometriosis (DIE), defined as subperitoneal infiltration of endometrial implants, i.e. when the endometriotic implants penetrate the retroperitoneal space at a distance of 5 mm or more (Koninckx 1991). DIE may be present in multiple locations, involving either the anterior or posterior pelvic compartments, or both.

We did not include certain rare types of endometriosis such as extrapelvic, bladder and ureteric endometriosis because the majority were reported in case reports or case series, and laparoscopy or laparotomy are not reliable reference standards for these conditions.

We excluded the studies where diagnosis of endometriosis was not the primary outcome (e.g. malignant versus benign masses or normal versus abnormal pelvis) and separate data for endometriosis was not available.

We also excluded the studies where the findings of the index test formed the basis of selection for the reference standard, because this was likely to distort an assessment of the diagnostic value of the index test.

We did include studies that recruited selected populations of women with endometriosis (i.e. those with specific rASRM stages), because there is a poor correlation between the rASRM classification and infertility or pain symptoms. Exclusion of these studies could result in the loss of potentially important diagnostic information from otherwise eligible publications. Where possible, we addressed the impact of these studies in the assessment of heterogeneity. When a study analysed a large population with a wide spectrum of endometriosis and additionally reported a subgroup analysis of the different stages of disease severity, we only considered estimates for the entire population. This is because a subgroup analysis would not directly address the review question regarding the clinical utility of the biomarker in disease detection.

Reference standards

The reference standard was visualisation of endometriosis at surgery (laparoscopy or laparotomy) with or without histological confirmation, as this is currently the best available test for endometriosis. If reported, we reviewed information regarding the inter- and intraobserver correlation of the reference standard.

We only included studies in which the reference test was performed within 12 months of the blood sample collection, on the assumption that disease status could change within a period of one year or longer, either naturally or as a result of treatment. We excluded studies in which the participants did not undergo the reference standard or where the findings of the index test formed the basis of selection for undertaking the reference standard, as this was likely to distort an assessment of the diagnostic value of the index test.

Summary of inclusion and exclusion criteria

Inclusion criteria

1. Types of studies
   a. Published and peer-reviewed
   b. RCTs
   c. Observational designs, including:
      i. single-gate design (single set of inclusion criteria defined by clinical presentation): all the participants had clinically suspected endometriosis;
      ii. two-gate design (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.
   d. Published in any language
   e. Performed in any healthcare setting
   f. Any sample size
2. Participants
   a. Reproductive-aged women
   b. Clinically suspected endometriosis, including:
      i. women who underwent abdominal surgery for other benign gynaecological conditions and had a surgical assessment for presence/absence of endometriosis;
      ii. asymptomatic women who have an incidental finding of endometriosis at surgery performed for another indication.
   c. Undertook both the index test and reference standard
3. Index tests
   a. One or several types of blood biomarkers
   b. Data reported in sufficient detail for the construction of 2 x 2 tables for the tests that showed differential expression between the groups
   c. Biomarkers where a 2 x 2 tables could not be constructed because the results did not differ between women with and without endometriosis, but all other inclusion criteria were met.
4. Target condition
   a. Pelvic endometriosis
      i. Peritoneal endometriosis
      ii. Ovarian endometrioma
      iii. DIE
      iv. Combinations of the above
5. Reference standard
   a. Surgical visualisation of lesions for the diagnosis of endometriosis (laparoscopy or laparotomy) with or without histological verification
   b. Performed within 12 months of the endometrial sample collection

Exclusion criteria

1. Types of studies
   a. Narrative or systematic reviews
   b. Retrospective design where the execution of reference test preceded the collection of the blood sample
   c. Prospectively collected samples that were selected from the archived material, but where information on the study population or the selection process was unclear
   d. Case reports or case series
   e. Conference proceedings
2. Participants
   a. Included cohort was not representative of the target population that would benefit from the test (e.g. women with known genital tract malignancy, ectopic pregnancies or acute pelvic inflammatory disease)
   b. Study included participants of postmenopausal age, and the data for the reproductive age group were not available in isolation
   c. Analysis only included participants with positive index test or positive reference standard
3. Index tests
   a. Blood biomarkers presented in combination with other diagnostic tests for endometriosis, and separate information for blood biomarkers was not available
   b. Study presented only specific technical aspects of an index test or focused on the biological events, rather than diagnostic performance of the test
   c. Study assessed screening or predictive test accuracy
4. Target condition
   a. Endometriosis was not the primary outcome of the trial (e.g. malignant versus benign masses or normal versus abnormal pelvis)
   b. Atypical, rare sites of endometriosis
5. Reference standard
   a. Reference standard performed only in a subset of study/control group
   b. Findings of the index test formed the basis of selection for the reference standard
   c. Other than specified in inclusion criteria

Search methods for identification of studies

We developed the search strategy in collaboration with the Trials Search Coordinator of the Gynaecology and Fertility Review Group, following recommendations of the *Cochrane Handbook for*

Systematic Reviews of Diagnostic Test Accuracy (De Vet 2008). We did not limit the searches to particular types of study design or impose language or publication date restrictions. The search strategy incorporated words in the title, abstract, text words across the record and the medical subject headings (MeSH). We initially created the search for one broad review looking at all diagnostic markers for endometriosis, but due to complexity, the review team split the originally planned review into five separate reviews. We designed two separate search strategies: one for all the biomarkers-based tests, and another for the imaging tests; we used the former in this review. We performed all searches from database inception to April - July 2015. We present the search strategies for each database and the number of hits per search in Appendix 2; Appendix 3; Appendix 4; Appendix 5; Appendix 6. The summary of the results is presented in Results of the search.

Electronic searches

We searched the following databases to identify the published studies that assessed the diagnostic value of blood biomarkers for endometriosis.

- CENTRAL (2015, July).
- MEDLINE (inception to May 2015).
- EMBASE (inception to May 2015).
- CINAHL (inception to April 2015).
- PsycINFO (inception to April 2015).
- Web of Science (inception to April 2015).
- LILACS (inception to April 2015).
- OAIster (inception to April 2015).
- TRIP (inception to April 2015).
- Databases of the trial registers.
  - ClinicalTrials.gov (inception to April 2015).
  - World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (inception to April 2015).
- Databases to identify reviews and guidelines as sources of references to potentially relevant studies.
  - MEDION (inception to January 2014, the last available date).
  - DARE (inception to April 2015).
  - PubMed, a 'Systematic Review' search under the 'Clinical Queries' link (inception to April 2015).
- Searches for papers recently published and not yet indexed in the major databases:
  - PubMed (simple search for the 6 months to April 2015).

Searching other resources

We handsearched the reference list of all relevant publications (retrieved full texts of the key articles and identified reviews).

We abandoned an initial attempt to locate the grey literature (unpublished studies and conference proceedings), as we faced substantial difficulty in obtaining full text publications or further details of studies reported in an abstract form.

Data collection and analysis

Selection of studies

Three authors of this review (RS, DA, VN) and three authors from the other reviews in this series (Emily Liu, Devashana Gupta and Lucy Prentice) scanned the titles of studies identified by our search to remove any clearly irrelevant articles. We reviewed the titles and abstracts of the remaining studies to select potentially relevant publications, and we divided the relevant articles into four categories of endometriosis biomarkers: blood, endometrial, urinary and combined tests. Two out of four review authors (of VN, LH, RS and DA) independently reviewed each of the full text versions of the articles that we had selected by title and abstract, assessing them for eligibility based on the criteria listed in 'Criteria for considering studies for this review'. A single failed eligibility criterion was sufficient for a study to be excluded from the review.

The review authors who assessed the relevance of the studies and eligibility for inclusion were not blind to the information about each article, including the publishing journal, the names of authors, the institution and the results. We resolved any disagreements by discussion and, if necessary, in consultation with a third review author (VJ) who is an expert in methodological aspects of Cochrane systematic reviews.

When papers updated previous publications and were performed on the same study population at different recruitment points, we used the most complete data set that superseded previous publications to avoid double counting participants or studies. We retrieved missing data directly by contacting authors to clarify study eligibility. When we found potentially relevant studies in languages other than English, we had them translated. For excluded studies, we documented the reasons for exclusion and details of which criteria were not met. We present the characteristics of included studies, excluded studies and studies awaiting classification in 'Characteristics of included studies', 'Characteristics of excluded studies' and 'Characteristics of studies awaiting classification', respectively.

Data extraction and management

Two out of five review authors (of VN, LH, RS, DA and CS) extracted data from each eligible study, resolving any disagreements by adjudication from the third review author (VJ). If required, we contacted study investigators to resolve any questions regarding the data.

To collect details from included studies, we used a purpose-designed data extraction form, designed specifically for this review and pilot tested on three studies of diagnostic accuracy tests for endometriosis. The following information was recorded for each study.

- General information and study design: first author, year of publication, country, language, setting, objectives, inclusion/exclusion criteria, type of enrolment.
- Characteristics of the study participants: age, symptoms/history/previous tests, type of target condition and its prevalence in the study population, number of participants enrolled and available for analysis, reasons for withdrawal.
- Features of the index test and reference standard: type, diagnostic criteria, number and experience of the operators, blinding of the operators to other tests or clinical data, interobserver variability, time interval between index test and reference standard.
- The reported number of true positives (TP), false negatives (FN), true negatives (TN) and false positives (FP), which we used to construct a 2 x 2 table for each index test. If studies did not report

these values, we attempted to reconstruct the 2 x 2 tables from the summary estimates presented in the study.

We extracted data into Review Manager 5 software (RevMan 2014), which we used to graphically display the quality assessment, the diagnostic estimates data and the descriptive analyses.

Assessment of methodological quality

To assess the quality of each included study, we used QUADAS-2, a modified version of the QUADAS tool for systematic reviews of diagnostic accuracy studies (Whiting 2011).

We present the review-specific QUADAS-2 tool and explanatory document in Table 3. We judged each study to be at 'low', 'high' or 'unclear' risk for each of four domains, and we assessed concerns about applicability in three domains. We considered studies as having low methodological quality when they were at high or unclear risk of bias or when we had a high concern regarding applicability at least in one domain. Two out of the four reviewers (of RS, DA, VN and LH) independently performed the assessment of each included study, settling disagreements with a third author (VJ) or by consensus. Two review authors (VN, RS) independently piloted the topic-specific tool to rate four of the included studies with a high level of agreement. We made modifications specific to the blood biomarkers review to the signalling questions of the original QUADAS-2 tool as follows.

Domain 1: We rephrased an original signalling question, 'Was a case-control design avoided?' as 'Was a two-gate design avoided?'. The diagnostic studies are cross-sectional in nature, aiming to compare the result of an index test with the result of the reference standard in the same group of participants. Study investigators measure the parameters at a single point in time and classify the groups by the outcome of the reference standard, albeit they perform the analysis retrospectively. Therefore, unlike epidemiological studies, the terminology 'cohort' and 'case-control' is less informative for diagnostic test trials, so we substituted them for 'single-gate' and 'two-gate' designs. We included this question because a two-gate design has more potential to introduce selection bias.

Domain 2: We introduced an additional signalling question, 'Was the phase of the menstrual cycle considered in interpreting the index test?' to assess bias in the interpretation of the test results. Some biochemical markers are sensitive to fluctuation in steroid sex hormone levels across a menstrual cycle, which could result in the differential expression of endometriosis biomarkers at different cycle phases.

We undertook the assessment of methodological quality for each domain, but we did not calculate a summary score to estimate the overall quality of studies (Whiting 2005).

Statistical analysis and data synthesis

We generated the estimates of sensitivity and specificity in forest plots and plotted them in the receiver operating characteristic (ROC) space for each index test using Review Manager 5 software (RevMan 2014). We investigated the diagnostic performance of each test and visually explored interstudy variation in the performance of each index test in relation to patient characteristics, study design and study quality considerations. When there were two or more tests evaluated in the same cohort, we included them as separate data sets, since the unit of analysis was the test result, not the patient.

For studies that reported subgroup analyses per phase of the menstrual cycle, we presented the data in a clinically relevant way. For instance, we presented pooled estimates when there was no statistically significant difference in biomarker expression between cycle phases. Alternatively, where putative biomarkers demonstrated cycle-dependent expression or were noted to be modulated by ovarian hormones, we reported the test performance either at several time points across the menstrual cycle or in the phase that demonstrated the most distinct difference between groups.

We estimated the expected operating point (mean sensitivity and specificity) and corresponding 95% confidence region using the bivariate logit normal random-effects model for all meta-analyses with four studies or more. When the number of studies was fewer than four, we did not attempt to estimate the covariance and reported a zero. To estimate the performance of the other tests in small meta-analyses (two or three data sets), we performed fixed-effect meta-analysis of sensitivity and specificity, in the absence of substantial heterogeneity. We performed the meta-analyses using SAS NLMIXED software (Cary, NC: SAS Institute Inc). We entered results from SAS into Review Manager 5 to provide plots of the mean or summary point(s) and confidence region(s), superimposed on the study specific estimates of sensitivity and specificity (RevMan 2014).

We assessed the comparative accuracy of index tests in two ways. In direct, fully paired comparisons where all the study participants received more than one index test as well as the reference standard, we plotted the estimates in Review Manager 5 (RevMan 2014). If meta-analysis was possible, we used test-level covariates in the bivariate logit normal model to identify statistically significant differences. Otherwise we reported the available comparative data in a narrative way and illustrated it using forest and ROC plots.

When judging test performance against the predetermined diagnostic criteria, we considered the point estimates of sensitivity and specificity as the most informative presentation of test performance. We acknowledge that tests with point estimates that did not reach the predetermined criteria, but with confidence intervals (CIs) that contained values above the threshold, could have diagnostic value. Furthermore, tests with point estimates that reached the criteria but with CIs containing values below the threshold could have an overestimated diagnostic value. If we use the range of the CIs rather than the point estimates of the data, the predetermined cut-off becomes meaningless. Therefore we did not consider CIs in qualifying the test performance but used this information in interpreting the reliability of the obtained data.

Dealing with missing data

We defined missing data as any information on the study population, index tests or reference standard that were not available from the publication and that were required to determine the eligibility of the study for inclusion, assess the methodological quality, or construct the results table. If we identified missing data, we contacted the authors in an attempt to obtain them. If missing data prevented a clear judgment regarding applicability for inclusion or the construction of accurate 2 x 2 tables and the data were unavailable from the primary investigators (for example

we were unable to locate the contact details of the authors, there was no reply from the authors or the authors replied that the requested information was unavailable), we excluded the study from the review.

Investigations of heterogeneity

We initially assessed heterogeneity by visually examining the forest plots of sensitivities and specificities and the ROC plots for each index test. We describe the potential sources of heterogeneity in the Secondary objectives. For diagnostic tests where there were more than 10 eligible studies, we initially planned to formally explore heterogeneity by using study level covariates, and to assess the sensitivity of results to the inclusion and exclusion of outlier studies in all analyses. However, we refrained from taking these steps because of the small numbers of studies in most analyses. It is important to use caution when interpreting small meta-analyses (few studies) with a limited total sample size.

Sensitivity analyses

We planned to conduct sensitivity analyses to assess the impact of the methodological quality of included studies on the results of any meta-analyses if sufficient data were available. We defined low quality studies as those for which we identified a high risk of bias for one or more QUADAS-2 domains. We also planned to use the 'leave-one-out' procedure (Higgins 2008) to assess the impact of each study on the meta-analysis results (leading study effect). However, we could not undertake this action due to the paucity of studies evaluating each biomarker, except CA-125.

Assessment of reporting bias

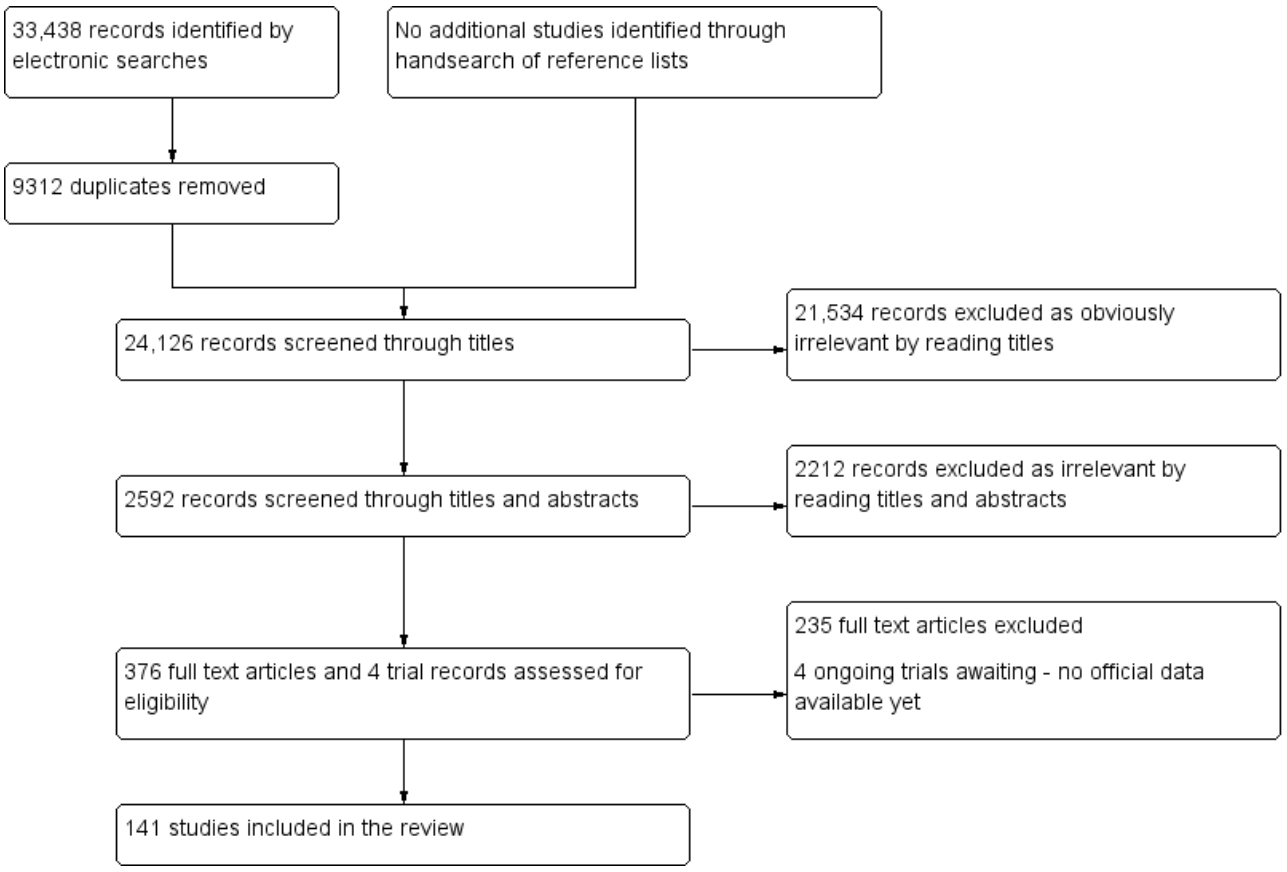
A comprehensive search of multiple sources for eligible studies, a search of trial registers and no language restrictions minimised the risk of reporting bias. However, publication bias generally arises when studies have a higher chance of being published if their results are positive. Therefore we initially searched and evaluated unpublished and published study databases and conference proceedings. During the process of qualifying the studies for inclusion in this review, we faced substantial difficulty in obtaining full text publications or further details of studies published in an abstract form. This precluded a reliable assessment of eligibility and methodological quality, and we decided not to include these publication sources in this review.

RESULTS

Results of the search

The literature search identified 33,438 references in the following databases: CENTRAL (N = 226), MEDLINE (N = 10,328), EMBASE (N = 10,313), CINAHL (N = 1131), PsycINFO (N = 174), Web of Science (N = 7425), LILACS (N = 420), OAIster (N = 446), Trip (N = 1648), trial registers for ongoing and registered trials (N = 523), MEDION (N = 2), DARE (N = 99), PubMed, a 'systematic review' search (N = 418) and simple search (N = 267). We present the flow of the selection process in Figure 2. We screened titles to exclude duplicates (N = 9312) and clearly irrelevant studies (N = 21,534). We eliminated a further 2212 references after reviewing the abstracts because they either did not address the research question or clearly did not meet the inclusion criteria. We retrieved the full texts of the remaining 376 records and assessed them for eligibility. Data from 86 studies required additional clarification from the authors, and we had 42 non-English publications translated. Ultimately, 141 studies were eligible and provided data for the review, while we excluded 235 studies. We identified four ongoing trials through clinical trials registries (Characteristics of ongoing studies), but as trial outcomes were not available, we will address the progress of these studies in future updates.

Figure 2. Flow of the studies identified in literature search for systematic review on imaging modalities for a non-invasive diagnosis of endometriosis.



Basic features of the included studies

We present a list of the details of the included studies in 'Characteristics of included studies'. The 141 eligible studies included 15,141 participants, with a median of 88 women per study (range 17 to 834). Of these studies, 70 estimated the diagnostic accuracy of blood biomarkers, 82 reported negative findings and 11 were in both groups. Seventy studies included enough data to estimate the diagnostic performance of an investigated test (N = 8716 participants, median 97, range 35 to 775 women). Each study evaluated one or several biomarkers, and some authors reported several estimates for the same biomarker at different menstrual cycle phases, different cut-off thresholds or both. When this occurred, we considered every estimation to be a separate test; however, we did not combine the diagnostic data sets for the biomarker of interest in one meta-analysis if obtained from the same or an overlapping cohort. Most studies reported diagnostic estimates for biomarkers that demonstrated differential expression between women with and without endometriosis, although in eight studies this assessment was undertaken for biomarkers that demonstrated no differential expression (Ferreira 1994; Gurgan 1990; Molo 1994; Muscatello 1992; Somigliana 2004; Tokmak 2011; Vigil 1999; Yang 1994). Eighty-two studies did not show any difference in the expression between the women with and without endometriosis, and they did not evaluate the diagnostic test accuracy of the blood biomarker (N = 7482 participants, median 73, range 17 to 834 women). This set of studies were methodologically eligible, but the biomarkers identified are unlikely to be of diagnostic utility and hence may not be worthy of further investigation.

Seventy of the included studies took place in Europe, 31 in Asia, 17 in North America, 14 in South America, 5 in the Middle East, 2 in Australia and 2 in unspecified locations. Ninety-five per cent (130/137) of the studies took place in university hospitals, of which at least 14 were referral centres for endometriosis. The earliest study was published in 1986, 107 studies were published after 2000, and 44 studies were published after 2010. All the included studies assessed women of reproductive age, and two focused exclusively on adolescent girls after menarche. All the studies were observational, mainly of cross-sectional design. Seventy-eight studies had a single-gate design, where both cases and controls were from the same patient population. Of these, 57 studies included women with suspected endometriosis based on clinical presentation (women presenting with pelvic pain, infertility, ovarian mass or a combination of these), 10 studies included only women undergoing an infertility work-up, eight studies included only a population with a persistent ovarian mass, two studies reported pelvic pain as a sole presenting symptom and one study evaluated asymptomatic women. Sixty-one studies had a two-gate design and included a wider group of participants who were undergoing surgery for various indications. Two studies presented insufficient information to determine whether they used a single- or two-gate design. Laparoscopy was the predominant surgical modality in the included studies; surgeons used either laparoscopy or laparotomy in 29 studies, and three studies did

not report information on the type of surgery. Seventy-five of the included studies used histopathology to confirm the surgical diagnosis.

Most of the studies (N = 123) evaluated pelvic endometriosis, 13 studies addressed only ovarian endometriosis, two studies focused on a combination of ovarian endometriosis and DIE, two studies looked only at peritoneal endometriosis, and one study considered only ultrasound-negative endometriosis.The reported prevalence of endometriosis varied from 16% to 84%. Eleven studies included only participants with minimal-mild endometriosis (rASRM stage I-II), 15 studies included only participants with moderate-severe endometriosis (rASRM stage III to IV), and eight studies did not report information regarding the severity of the disease. Fifty-one studies received financial support, of which 8 reported funding by biotech or pharmaceutical companies. In six of the eight commercially supported studies, there was no statement regarding a conflict of interest. For the remaining two studies, one group of authors reported that most of the authors worked in the biotechnology industry, and one group had nothing to declare. Overall, the authors of 33 studies declared no conflict of interest, with five reporting that there was no financial support from any external source. Three groups reported conflicts of interest (employee of a biotech company, lecturing honorarium from pharmaceutical companies and not specified), and no information was available from the remaining studies.

Basic features of the excluded studies

We present the list and descriptions of the excluded studies in 'Characteristics of excluded studies'. Based on a full text assessment, we excluded 235 studies, of which 23 were retrospective with the blood samples being collected after the surgical procedure. A further 88 studies reported biomarker levels that were statistically significant when the study and control groups were compared, but they did not provide enough information for the construction of 2 x 2 contingency tables. Forty-six of the excluded studies used a reference standard other than abdominal surgery and did not provide information regarding the surgical diagnosis. We excluded an additional 20 studies because they did not provide enough detail on the research methods, the study population or both to assess eligibility, and this information was not available from the authors. In 24 studies, the index test was outside the inclusion criteria, including comparisons between different types or stages of endometriosis without including a disease-free group (N = 13); reports on biological events or technical aspects of the test without direct comparison of biomarker levels between the groups (N = 6); evaluations of a screening or predictive rather than a diagnostic test (N = 3); or use of male or umbilical cord samples as control group (N = 2). In nine studies, the population was outside the inclusion criteria because they enrolled postmenopausal women, pregnant women or women with genital tract malignancies, and an independent assessment of reproductive-aged women without these conditions was not possible. We excluded a further nine studies as their population cohort overlapped with another updated, included study. In five of the excluded studies, the target condition was outside the inclusion criteria, comparing a benign versus malignant mass or normal versus abnormal pelvis without any independent data for endometriosis. We excluded three studies because they were review articles, and we were unable to locate the full text for another three studies.

Methodological quality of included studies

We illustrate the quality of the included studies in the QUADAS-2 results summary (Figure 3; Figure 4). Overall, the studies were of poor methodological quality, and all studies had an unclear or high risk of bias in at least one domain.

Figure 3. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies

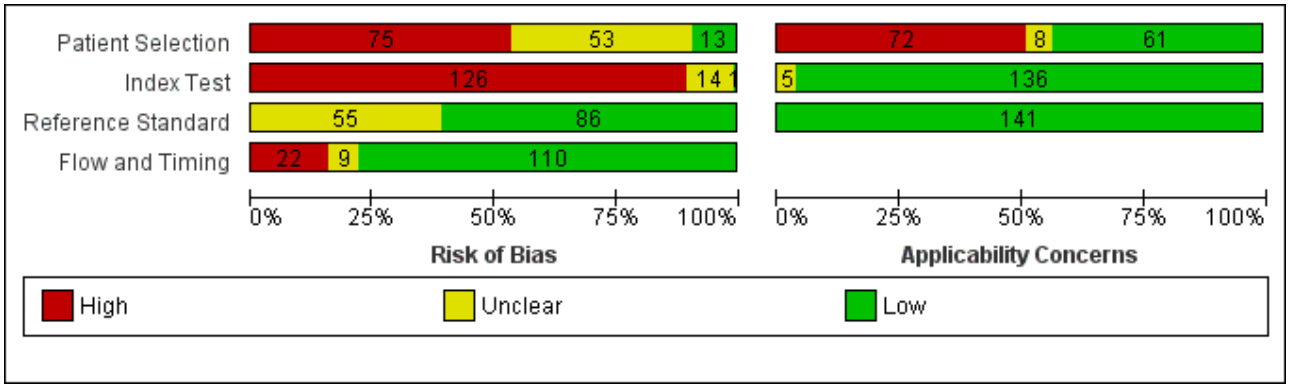


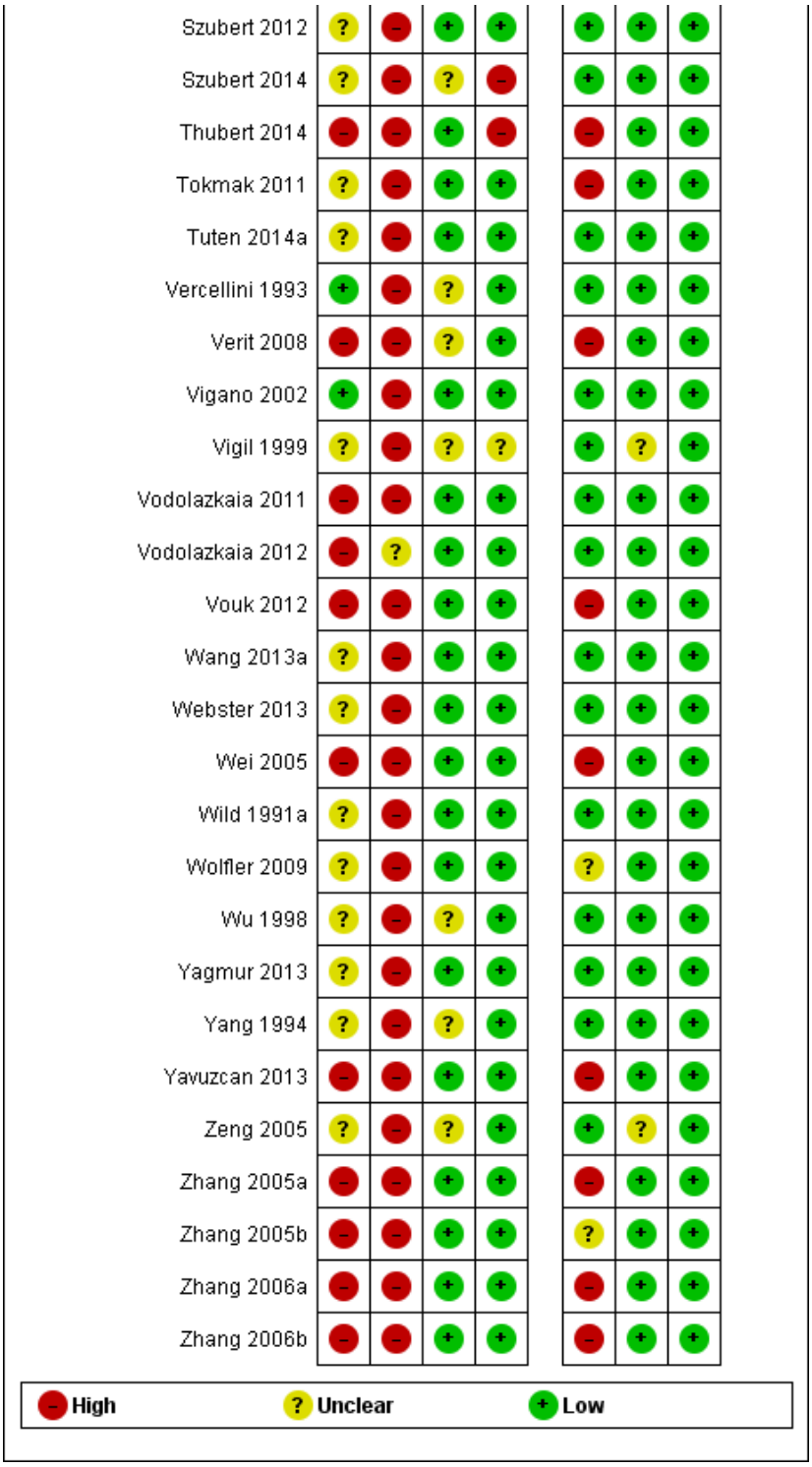
Figure 4. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study

Figure 4. (Continued)

Figure 4. (Continued)

Figure 4. (Continued)

Figure 4. (Continued)



Thirteen studies presented a low risk of patient selection bias (Chen 1998; Fairbanks 2009; Guerriero 1996a; Guerriero 1996b; Koninckx 1996; Molo 1994; Podgaec 2007; Ramos 2012; Rosa E Silva 2007; Somigliana 2002; Somigliana 2004; Vercellini 1993; Vigano 2002), 53 studies demonstrated an unclear risk, and 75 studies were assessed at high risk for this domain. Non-consecutive or non-random patient selection, utilisation of a two-gate design for patient selection, the absence of a clear definition of inclusion/exclusion criteria and use of a highly selected group of women were the main reasons for a high risk assessment of bias.

One study demonstrated a low risk of index test interpretation bias (Pittaway 1989), 14 studies demonstrated an unclear risk and 126 studies carried a high risk. A lack of clear pre-specified criteria for a positive diagnosis and index test operators not being blind to the results of reference standard were the main reasons for a high risk assessment. We also assigned a high risk of bias for this domain to studies where the phase of menstrual cycle was not considered when interpreting the index test. This was considered an important criterion, since varying ovarian hormones across the cycle could influence biomarker expression and undermine the reliability of the results. Studies rarely reported the skill level of a test operator or the interobserver variability, both of which directly relate to test performance. As the positive index test criteria were variable between the studies and the index test protocols were not standardised, quality judgements for the index test were complex.

Eighty-six studies were at low risk of bias in the reference standard domain (Agic 2008; Barbosa 2009; Barcz 2002; Bilibio 2014; Borkowski 2008; Calienno 2008; Chen 1998; Cho 2007; Dayangan Sayan 2013; De Placido 1998; Drosdzol-Cop 2012a; Drosdzol-Cop 2012b; Elgafor el Sharkwy 2013; Fairbanks 2009; Fassbender 2009; Fassbender 2012; Fedele 1989; Ferreira 1994; Ferrero 2005a; Florio 2007; Florio 2009; Gagne 2003a; Gagne 2003b; Gazvani 1998; Glitz 2009; Gogacz 2014; Guerriero 1996a; Guerriero 1996b; Gurgan 1999; Hallamaa 2012; Hassa 2009; Jee 2008; Jia 2013; Kalu 2007; Khan 2006; Khan 2012; Khan 2013; Khanaki 2012; Kim 2008; Kitawaki 2005; Kocbek 2013; Kocbek 2014a; Kocbek 2014b; Kubatova 2013; Kuessel 2014; Kurdoglu 2009; Lambrinoudaki 2009; Li 2005; Lin 2005; Mabrouk 2012; Mihalyi 2010; Mohamed 2013; Odukoya 1996; Ohata 2008; Olkowska-Truchanowicz 2013; Othman 2008; Paiva 2014; Patton 1986; Philippoussis 2004; Pittaway 1989; Podgaec 2007; Ramos 2012; Rosa E Silva 2014; Salehpour 2009; Somigliana 2002; Somigliana 2004; Szczepanska 2001b; Szubert 2012; Thubert 2014; Tokmak 2011; Tuten 2014a; Vigano 2002; Vodolazkaia 2011; Vodolazkaia 2012; Vouk 2012; Wang 2013a; Webster 2013; Wei 2005; Wild 1991a; Wolfler 2009; Yagmur 2013; Yavuzcan 2013; Zhang 2005a; Zhang 2005b; Zhang 2006a; Zhang 2006b), while the rest (N = 55) were at unclear risk. No studies demonstrated a high risk. We assigned an unclear risk of bias if there was not enough information to assess how likely the reference standard was to have correctly classified the target condition. This could occur when authors did not adequately describe surgical procedures, state the positive reference standard criteria, clarify whether they used histology to confirm the surgical diagnosis or provide information regarding the expertise of the surgeons and pathologists involved.

One hundred and ten studies presented a low risk of bias in the flow and timing domain (Acien 1989; Agic 2008; Akoum 1996; Andreoli 2011; Barbati 1994; Barbosa 2009; Bilibio 2014; Borkowski 2008; Braun 1996; Calienno 2008; Chen 1998; Cho 2007; Colacurci 1996a; De Placido 1998; Drosdzol-Cop 2012a; Drosdzol-Cop 2012b; Fairbanks 2009; Fassbender 2009; Fassbender 2012; Ferreira 1994; Ferrero 2005a; Florio 2007; Florio 2009; Foda 2012; Gagne 2003a; Gagne 2003b; Glitz 2009; Gogacz 2014; Goluda 1998; Gorai 1993; Guerriero 1996a; Guerriero 1996b; Gurgan 1990; Gurgan 1999; Hallamaa 2012; Harada 2002; Hornstein 1995; Iwasaki 1993; Jee 2008; Jia 2013; Khanaki 2012; Kianpour 2012; Kianpour 2013; Kitawaki 2005; Kocbek 2013; Kubatova 2013; Kuessel 2014; Lambrinoudaki 2009; Li 2005; Lima 2006; Lin 2005; Liu 2009; Mabrouk 2012; Maeda 2002a; Maeda 2002b; Maiorana 2007; Markham 1997a; Martinez 2007; Matalliotakis 2003a; Matalliotakis 2004; Matveeva 1990; Mier-Cabrera 2011; Mihalyi 2010; Mohamed 2013; Molo 1994; Morin 2005; Muscatello 1992; Odukoya 1996; Oku 2004; Olkowska-Truchanowicz 2013; Othman 2008; Ozhan 2014; Paiva 2014; Patton 1986; Philippoussis 2004; Pittaway 1989; Podgaec 2007; Ramos 2012; Riley 2007; Rosa E Silva 2007; Salehpour 2009; Somigliana 2002; Somigliana 2004; Steff 2004a; Suen 2014; Szczepanska 2001a; Szczepanska 2001b; Szubert 2012; Tokmak 2011; Tuten 2014a; Vercellini 1993; Verit 2008; Vigano 2002; Vodolazkaia 2011; Vodolazkaia 2012; Vouk 2012; Wang 2013a; Webster 2013; Wild 1991a; Wolfler 2009; Wu 1998; Yagmur 2013; Yang 1994; Yavuzcan 2013; Zeng 2005; Zhang 2005a; Zhang 2005b; Zhang 2006a; Zhang 2006b), nine studies demonstrated an unclear risk and 22 studies carried a high risk. All participants received the same reference standard. The time interval between the index test and the reference standard was 12 months or less, and the most commonly reported time interval was immediately before surgery. We assigned an unclear risk if authors did not clearly state the time interval, but if their descriptions suggested that the interval was reasonably short. We assigned a high risk of bias if there were unexplained withdrawals that exceeded 5% of the enrolled population or if the reason for withdrawal could introduce selection bias regarding the samples analysed.

Sixty-one studies presented a low concern for patient selection applicability (Barbati 1994; Borkowski 2008; Chen 1998; Colacurci 1996a; Drosdzol-Cop 2012a; Drosdzol-Cop 2012b; Fairbanks 2009; Fassbender 2009; Fassbender 2012; Fedele 1989; Ferreira 1994; Foda 2012; Franchi 1993; Gogacz 2014; Gurgan 1999; Harada 2002; Hassa 2009; Hornstein 1995; Inagaki 2003; Iwasaki 1993; Khan 2006; Khan 2012; Kim 2008; Lamp 2012; Lanzone 1991; Lin 2005; Liu 2009; Mabrouk 2012; Matalliotakis 2003a; Matalliotakis 2004; Mihalyi 2010; Muscatello 1992; Odukoya 1996; Oku 2004; Othman 2008; Ozhan 2014; Paiva 2014; Philippoussis 2004; Pittaway 1989; Podgaec 2007; Ramos 2012; Rosa E Silva 2007; Salehpour 2009; Somigliana 2002; Somigliana 2004; Szczepanska 2001b; Szubert 2012; Szubert 2014; Tuten 2014a; Vercellini 1993; Vigano 2002; Vigil 1999; Vodolazkaia 2011; Vodolazkaia 2012; Wang 2013a; Webster 2013; Wild 1991a; Wu 1998; Yagmur 2013; Yang 1994; Zeng 2005), eight demonstrated an unclear concern and 72 were of high concern. We assigned high concern in patient selection applicability if the study utilised two-gate selection for cases and controls or if it only evaluated a limited spectrum of disease. In our view, any sampling deviation from a representative group of the entire clinically relevant population could skew the estimates of diagnostic accuracy in either direction. We reported unclear concern if this information was unclear, for example if the severity of endometriosis was not reported.

In 136 studies there was a low concern in index test applicability, whereas in five studies the concern was unclear (Calienno 2008; Kurdoglu 2009; Rosa E Silva 2007; Vigil 1999; Zeng 2005), and none of the studies presented a high concern. We assigned an unclear concern when the study did not present sufficient information regarding the conduct of the tests, such as the laboratory methods, reagents used or the level of expertise of the test operators.

All 141 studies were of low concern for applicability with regard to the reference standard, and none had a high or unclear concern. All the included studies implemented pelvic surgery (laparoscopy or laparotomy) as a reference standard, which could be relied upon to match the review question.

Findings

We evaluated a total of 122 blood biomarkers in 141 included studies; 47 biomarkers had a diagnostic evaluation in 70 studies. Studies assessed 44 biomarkers as a single blood test, along with 29 combinations of two to six biomarkers (Summary of findings 1). The presence of endometriosis did not alter 97 biomarkers evaluated in 79 studies (Appendix 7). Twenty-two biomarkers demonstrated altered levels in endometriosis in some studies and showed no difference in expression in other studies. We report the findings for two separate groups: blood biomarkers that were evaluated for the diagnosis of pelvic endometriosis, when any type of endometriosis was assessed against disease-free controls; and blood biomarkers that could differentiate ovarian endometrioma from other benign ovarian cysts in women of reproductive age, when assessing ovarian endometriosis against other types of ovarian masses. We have biologically subcategorised biomarkers and presented them under these categories in alphabetical order. To assist readers in the search for a specific biomarker, we present an index of the biomarkers with biological annotation in Appendix 1 .

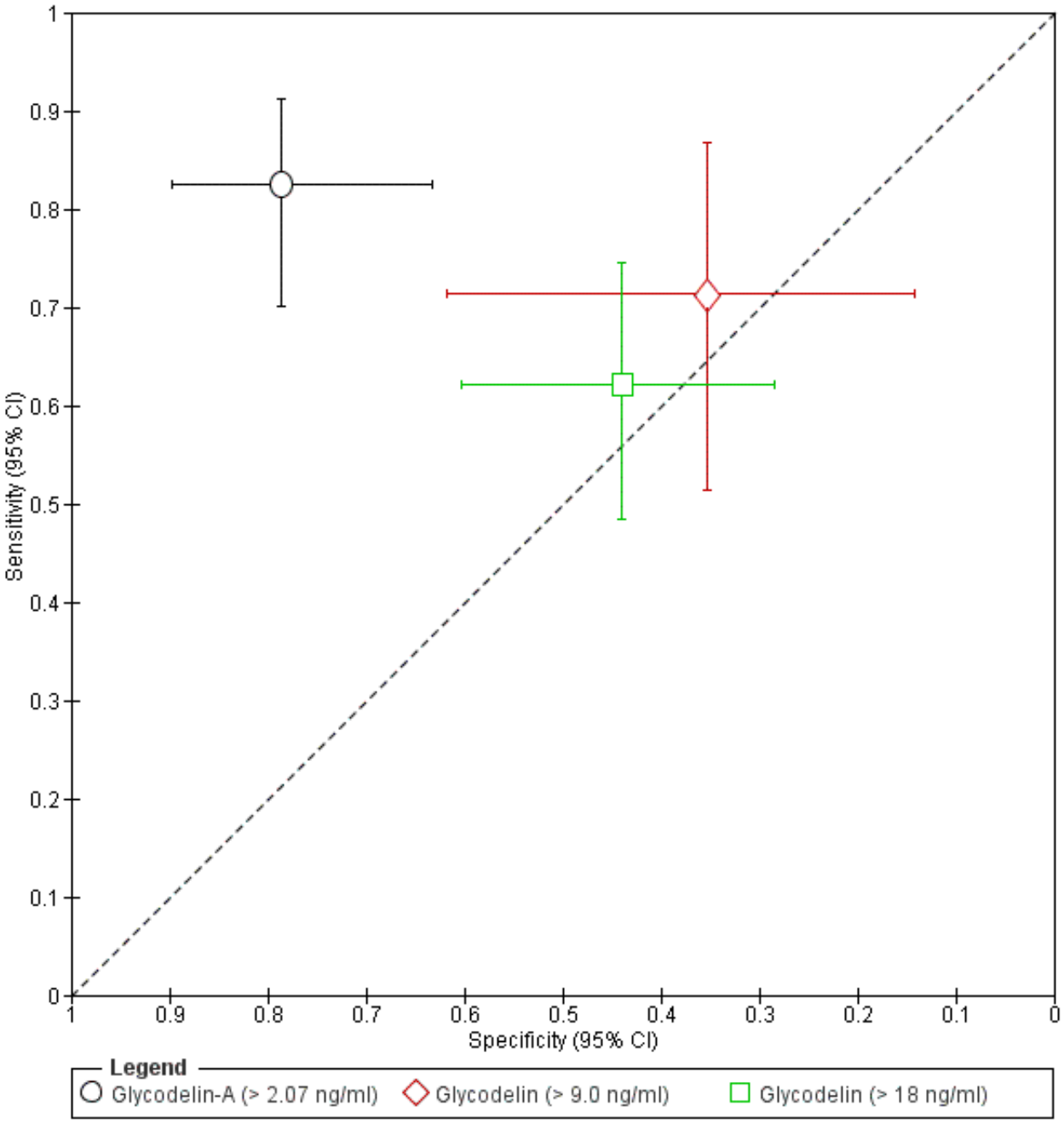
Blood biomarkers evaluated for the diagnosis of pelvic endometriosis (peritoneal, ovarian and deep infiltrating)

1. Angiogenesis/growth factors and their receptors

1.1. Glycodelin-A (PP14 or PAEP) (or placental protein 14 or progestogen-associated endometrial protein)

Two studies, including three data sets with a total of 198 participants, assessed the value of glycodelin in detecting pelvic endometriosis (Figure 5). Investigators assigned three different cut-off thresholds in each data set.

Figure 5. Summary ROC plot of Glycodelin for detection of endometriosis. Each point represents the pair of sensitivity and specificity for each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The bars correspond to 95% CIs of each individual evaluation. Two evaluations (> 9 ng/ml and > 18 ng/ml) were performed on overlapping populations. The data were not assessed by meta-analysis.



1. > 2.07 ng/ml (1 study, 99 participants, follicular or luteal cycle phase, rASRM I to IV), showing a sensitivity of 0.82 (95% CI 0.70 to 0.91) and a specificity of 0.79 (95% CI 0.63 to 0.90); Kocbek 2013.
2. > 18 ng/ml (1 study, 99 participants, all cycle phases, rASRM I to IV), showing a sensitivity of 0.62 (95% CI 0.48 to 0.74) and a specificity of 0.44 (95% CI 0.28 to 0.60); Kocbek 2013.
3. > 9.0 ng/ml (1 study, 45 participants, follicular cycle phase, rASRM I to IV), showing a sensitivity of 0.71 (95% CI 0.51 to 0.87) and a specificity of 0.35 (95% CI 0.14 to 0.62); Vodolazkaia 2012.

The same study (Kocbek 2013) performed two tests on an overlapping population of women, and other studies used varying

thresholds, so it was not possible to combine studies in a meta-analysis. In three contrasting studies (206 participants, rASRM I to IV), glycodelin concentrations did not change in women with endometriosis in the follicular phase (Drosdzol-Cop 2012a), follicular or luteal phase (Joshi 1986), or when the cycle phase was not specified (Paiva 2014). It appears that there is little clinical value in using glycodelin-A to diagnose endometriosis.

1.2. IGFBP-3 (insulin-like growth factor-binding protein-3)

One study evaluated the accuracy of IGFBP-3 in detecting pelvic endometriosis in 99 women with ultrasound negative, rASRM I to IV endometriosis (Vodolazkaia 2012). This study included two evaluations: all the participants in all phases of menstrual cycle (cut-off threshold > 210 ng/ml), demonstrating a sensitivity of 0.55 (95% CI 0.42 to 0.68) and a specificity of 0.44 (95% CI 0.28 to 0.60); and only in participants in the follicular cycle phase (45 women, cut-off threshold > 200 ng/ml), with a sensitivity of 0.71 (95% CI 0.51 to 0.87) and a specificity of 0.29 (95% CI 0.10 to 0.56). There were no significant differences in for IGFBP-3 levels between women with and without endometriosis in two additional studies (116 participants, follicular and luteal or only luteal cycle phase, rASRM I to IV) (Gurgan 1999; Philippoussis 2004). These data suggest that IGFBP-3 is not sensitive or specific enough to be clinically useful in diagnosing endometriosis.

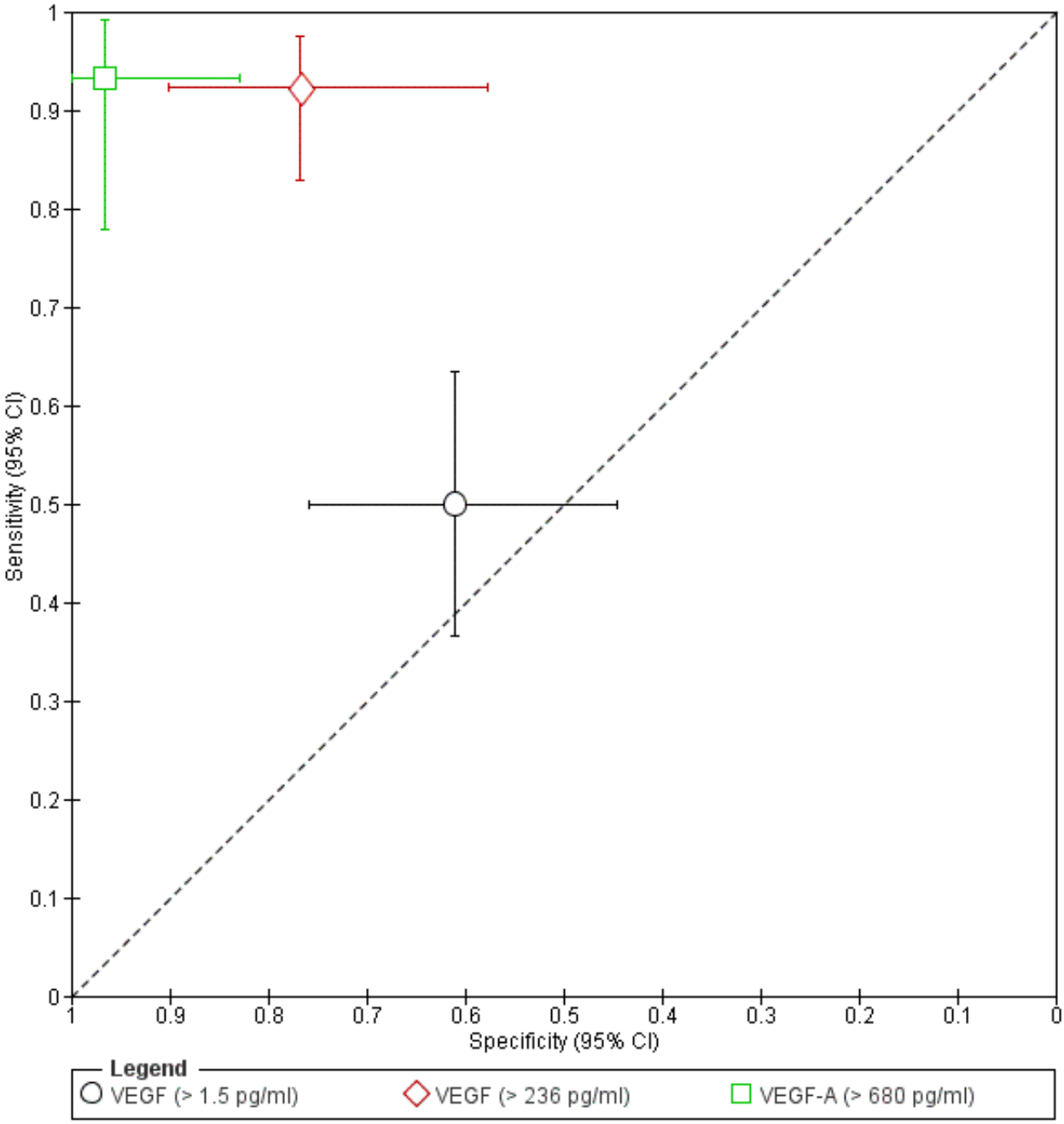
1.3. Leptin

The diagnostic performance of leptin was assessed as a component of a combination of blood biomarkers (see below under 'Combined tests'). Four other studies (311 participants, rASRM I to IV) demonstrated that leptin levels alone did not differ between the groups of women with and without endometriosis when tested in all phases of menstrual cycle or when the cycle phase was not specified (Ozhan 2014; Paiva 2014; Vigano 2002; Wei 2005). Overall, leptin did not appear to be reliable as a marker for endometriosis.

1.4. VEGF (vascular endothelial growth factor)

Three studies with a total of 254 participants evaluated VEGF for the diagnosis of pelvic endometriosis (Vodolazkaia 2012; Foda 2012; Mohamed 2013)(Figure 6). Each study differed with regard to the population studied, the cycle phase and cut-off thresholds.

Figure 6. Summary ROC plot of VEGF for detection of endometriosis. Each point represents the pair of sensitivity and specificity for each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The bars correspond to 95% CIs of each individual evaluation. The data were not assessed by meta-analysis.



- VEGF with a cut-off of > 1.5 ng/ml (1 study, 99 participants in all phases of menstrual cycle, ultrasound negative endometriosis, rASRM I to IV) had a sensitivity of 0.50 (95% CI 0.37 to 0.63) and a specificity of 0.61 (95% CI 0.45 to 0.76); Vodolazkaia 2012.
- VEGF with a cut-off of > 236.00 pg/ml (1 study, 95 participants in follicular cycle phase, rASRM I to IV) showed a sensitivity of 0.92 (95% CI 0.83 to 0.97) and a specificity of 0.77 (95% CI 0.58 to 0.90); Foda 2012.
- VEGF with a cut-off of > 680.00 pg/ml (1 study, 60 participants in follicular cycle phase, rASRM III to IV) demonstrated a sensitivity of 0.93 (95% CI 0.78 to 0.99) and a specificity of 0.97 (95% CI 0.83 to 1.00); Mohamed 2013.

The last test had the highest diagnostic accuracy, but investigators only evaluated it for moderate-severe disease. Substantial variations in the methodology and the populations studied precluded combining this data in a meta-analysis. Another seven studies (842 women, rASRM I to IV) demonstrated that VEGF levels were not influenced by endometriosis in the follicular phase (Da Silva 2014; Mabrouk 2012), luteal phase (Gagne 2003b), follicular or luteal phase (Cho 2007; Kianpour 2013; Othman 2008), or when the cycle phase was not specified (Paiva 2014). There is considerable inconsistency in the VEGF-A data, although follicular phase VEGF-A testing appears to have some potential in diagnosing endometriosis. Further work to confirm or refute this observation and determine the value of VEGF-A blood testing to diagnose endometriosis is warranted.

1.5. Urocortin

Two studies reported on urocortin, both of which assessed the accuracy of this biomarker in discriminating ovarian endometriosis from other benign ovarian masses. These studies are presented separately under 'Blood biomarkers that could differentiate ovarian endometrioma from other benign ovarian cysts in women of reproductive age'.

1.6. Angiogenesis/growth factors that exhibited no differential expression in endometriosis

We present a detailed summary of other angiogenesis and growth factors that did not display significant differences in expression levels in women with endometriosis in Appendix 7. The list includes:

1. angiogenic activity of serum (Barcz 2002; 84 participants);
2. CAC (Webster 2013; 64 participants);
3. EGF (Philippoussis 2004; 72 participants);
4. sEGF-R (Matalliotakis 2003a; 48 participants);
5. sFlt-1 (sVEGFR-1) (Cho 2007; 70 participants);
6. HGF (Khan 2006; 58 participants);
7. IGF-1 (Matalliotakis 2003a; Steff 2004a; 196 participants);
8. IGF-2 (Gurgan 1999; 44 participants);
9. PDGF (Kalu 2007; 40 participants).

Collectively, the results were discouraging, but not sufficient to draw conclusions regarding the role of the biomarkers in detecting endometriosis.

2. Apoptosis markers

2.1. Annexin-V

Investigators evaluated the diagnostic performance of annexin-V in conjunction with other blood biomarkers as a part of combined test, and we report the findings below under 'Combined tests'. One additional study (101 participants, cycle phase not reported, rASRM I to IV) demonstrated that annexin-V was not differentially expressed in women with and without endometriosis (Paiva 2014). Further work in a well-characterised population is needed to support this observation.

2.2. Survivin

One study (60 participants, follicular cycle phase) evaluated survivin for the diagnosis of DIE and ovarian endometriosis and found a very low sensitivity of 0.07 (95% CI 0.02 to 0.20) and a specificity of 0.90 (95% CI 0.68 to 0.99). The authors did not report a cut-off threshold. There were no other eligible studies that assessed this biomarker.

2.3. Apoptosis markers that exhibited no differential expression in endometriosis

We present additional markers of apoptosis that were not altered in endometriosis in Appendix 7, including: anti-survivin antibody (Lamp 2012; 145 participants); apoptotic cells (Mier-Cabrera 2011; 62 participants); and sFas (Kalu 2007; 40 participants). Overall, there was insufficient data to make recommendations regarding these biomarkers.

3. Cell adhesion molecules and other matrix-related proteins

3.1. sICAM-1 (soluble form of intercellular-adhesion molecule-1)

Two studies evaluated the accuracy of sICAM-1 in detecting pelvic endometriosis. One study included women with ultrasound negative pelvic endometriosis, rASRM I to IV, and presented two overlapping data sets, which we therefore did not combine in a meta-analysis (Vodolazkaia 2012). One data set from this study included 99 participants at all phases of menstrual cycle and demonstrated a sensitivity of 0.55 (95% CI 0.42 to 0.68) and a specificity of 0.50 (95% CI 0.34 to 0.66) for a cut-off threshold of < 243 ng/ml. The second data set comprised 28 participants in the menstrual cycle phase and showed a sensitivity of 0.73 (95% CI 0.39 to 0.94) and a specificity of 0.29 (95% CI 0.10 to 0.56) for a cut-off of < 254.6 ng/ml. Another study (60 participants, rASRM I to IV, cycle phase not reported) demonstrated an opposite direction of differential expression of sICAM-1 in endometriosis (higher sICAM-1 levels in endometriosis as opposed to the former study where expression in endometriosis was lower than in controls (Zhang 2006b). Utilising a cut-off threshold of > 241.46 µg/ml, the sensitivity was 0.6 (95% CI 0.41 to 0.77) and the specificity was 0.87 (95% CI 0.69 to 0.96). Four studies reported negative findings for the same test (271 participants, various phases of menstrual cycle); three of those studies assessed a wide spectrum of pelvic endometriosis, rASRM I to IV (De Placido 1998; Paiva 2014; Somigliana 2002), and one study assessed only minimal-mild disease, rASRM I-II (Goluda 1998). This evidence suggests that sICAM-1 molecule is not reliable as a diagnostic test for endometriosis.

3.2. LN-1 (laminin-1)

One study evaluated the value of LN-1 in detecting pelvic endometriosis (73 participants, cycle phase not specified, rASRM II to IV), demonstrating a sensitivity of 0.72 (95% CI 0.58 to 0.83) and a specificity of 0.70 (95% CI 0.46 to 0.88). There is insufficient evidence to comment on the diagnostic performance of this biomarker.

3.3. Cell adhesion molecules that exhibited no differential expression in endometriosis

Three studies reported negative findings for three additional biomarkers from this group (Appendix 7): biglycan (Kocbek 2014b; 96 participants); MMP-9 (Mabrouk 2012; 60 participants); and sE-selectin (Goluda 1998; 20 participants). In view of the paucity of data, the diagnostic role of these biomarkers in endometriosis requires further investigation.

4. Cytoskeleton molecules

4.1. CK19 (Cytokeratin-19) exhibited no differential expression in endometriosis

One study (79 participants, follicular or luteal cycle phase, severity not reported) evaluated expression of CK19 in pelvic endometriosis and demonstrated no significant differences in CK19 expression between the groups (Kuessel 2014). This observation provides too few data to draw conclusions regarding the diagnostic role of this blood biomarker in endometriosis.

5. DNA-repair/telomere maintenance molecules

5.1. Telomere length exhibited no differential expression in endometriosis

One study evaluated telomere length of peripheral blood mononuclear cells (50 participants, luteal cycle phase, rASRM I to IV) and demonstrated no significant difference between the women diagnosed with pelvic endometriosis and the disease-free group (Hapangama 2008). Further studies are required before the diagnostic role of telomere length in peripheral blood cells in the diagnosis of endometriosis can be determined.

6. High-throughput molecular markers

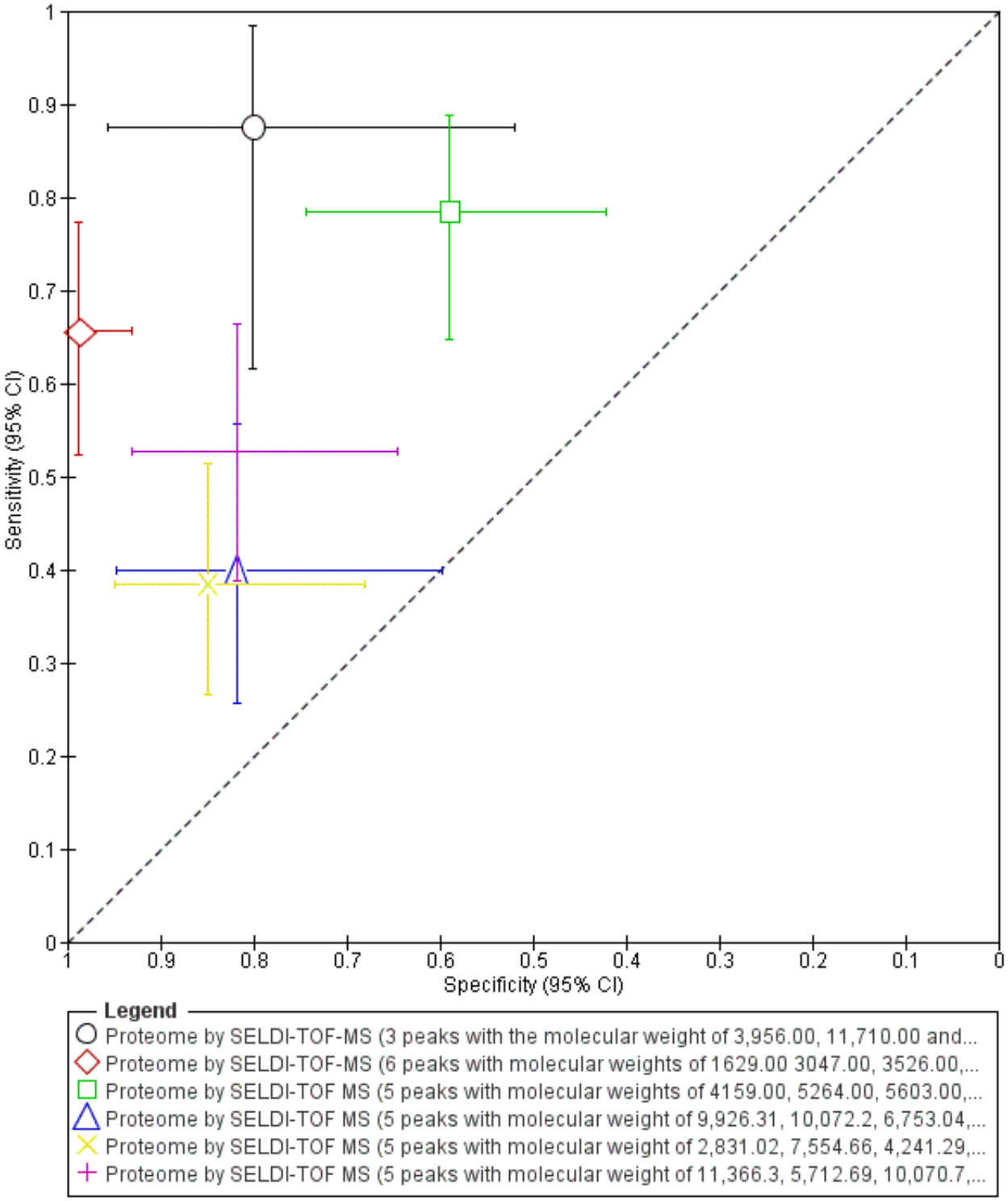
6.1. Metabolome

One study assessed the accuracy of the metabolome in detecting endometriosis (92 participants, all phases of menstrual cycle, ovarian endometriosis, rASRM III to IV) using electrospray ionisation mass spectrometry (ESI-MS/MS). A diagnostic model including hydroxy sphingomyelin SMOH C16:1 and the ratio of phosphatidylcholine PCaa C36:2 to ether-phospholipid PCae C34:2 was selected using stepwise regression. When adjusted for age and BMI, it showed a sensitivity of 0.90 (95% CI 0.76 to 0.97) and a specificity of 0.85 (95% CI 0.72 to 0.93). These estimates approach the criteria for a SnOUT triage test; however, the trial assessed a limited spectrum of disease and did not provide the cut-off thresholds. Although promising, these findings require further confirmation in a broader group of women with endometriosis.

6.2. Proteome

Four studies included six data sets with a total of 425 participants, assessing the accuracy of the proteome in detecting endometriosis. All the included studies evaluated rASRM I to IV pelvic endometriosis and performed matrix-assisted Surface-Enhanced Laser Desorption Ionization Time-of-Flight Mass Spectrometry (SELDI-TOF-MS) (Figure 7). The different groups took varying approaches to the data analysis and the construction of a diagnostic model. They described distinct sets of proteins as discriminating between women with and without endometriosis, precluding a meta-analysis. One study (31 participants, cycle phase not reported) identified three protein peaks with molecular weights of 3,956.00 Da, 11,710.00 Da and 6,986.00 Da, and it reported a sensitivity of 0.88 (95% CI 0.62 to 0.98) and a specificity of 0.80 (95% CI 0.52 to 0.96) (Liu 2009). Another group (139 participants, all phases of menstrual cycle) showed six protein peaks with molecular weights of 1629.00 Da, 3047.00 Da, 3526.00 Da, 3774.00 Da, 5046.00 Da and 5068.00 Da. This test demonstrated a sensitivity of 0.66 (95% CI 0.52 to 0.77) and a specificity of 0.99 (95% CI 0.93 to 1.00) (Seeber 2010) which meets the criteria for a SpIN triage test. A further study (90 participants, follicular or luteal cycle phase), demonstrated that five protein peaks with molecular weights of 4159.00 Da, 5264.00 Da, 5603.00 Da, 9861.00 Da and 10,533.00 Da had a sensitivity of 0.78 (95% CI 0.65 to 0.89) and a specificity of 0.59 (95% CI 0.42 to 0.74) in detecting endometriosis (Wolfler 2009). The most recent study reported three separate evaluations for each menstrual cycle phase with varying sets of proteins for each cycle phase (Fassbender 2012). Specifically, testing 67 participants in the menstrual cycle phase revealed five peaks with molecular weights of 9,926.31 Da, 10,072.20 Da, 6753.04 Da, 4302.67 Da and 9328.49 Da, with a sensitivity of 0.40 (95% CI 0.26 to 0.56) and a specificity of 0.82 (95% CI 0.60 to 0.95). Evaluation of 98 women in the follicular cycle phase showed that five peaks with molecular weight of 2831.02 Da, 7554.66 Da, 4241.29 Da, 2953.25 Da and 9927.73 Da had a sensitivity of 0.38 (95% CI 0.27 to 0.51) and a specificity of 0.85 (95% CI 0.68 to 0.95). In the same study, five protein peaks in 88 women in the luteal cycle phase had molecular weights of 11,366.30, 5712.69, 10,070.70, 3017.68, 3824.44 Da had a sensitivity of 0.53 (95% CI 0.39 to 0.66) and a specificity of 0.82 (95% CI 0.65 to 0.93) to detect endometriosis. None of the studies reported diagnostic cut-off thresholds. Further evaluations of this diagnostic approach through using standardised analytical processes with similar sets of markers and defined cut-off thresholds is required for a comprehensive assessment of this diagnostic tool.

Figure 7. Summary ROC plot of proteome by SELDI-TOF-MS for detection of endometriosis. Each point represents the pair of sensitivity and specificity for each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different sets of proteins determined by molecular weight (MW) in daltons (Da). The bars correspond to 95% CIs of each individual evaluation. The data were not assessed by meta-analysis.



7. Hormonal markers

7.1. Prolactin

One study that included two data sets with a total of 97 participants in the luteal cycle phase explored the diagnostic accuracy of prolactin for pelvic endometriosis, rASRM I to IV. The study evaluated two different cut-off thresholds: > 14.8 ng/ml, which demonstrated a sensitivity of 0.44 (95% CI 0.32 to 0.58) and a specificity of 0.94 (95% CI 0.80 to 0.99) and > 20 ng/ml, with a sensitivity of 0.21 (95% CI 0.11 to 0.33) and a specificity of 1.00 (95% CI 0.90 to 1.00). Despite the high specificity, the sensitivity for the both thresholds remains unacceptably low even for a triage test. These data are not sufficient to draw meaningful conclusions.

7.2. Hormonal biomarkers that exhibited no differential expression in endometriosis

Blood levels of the following hormonal markers showed no statistically significant difference in women with and without endometriosis (Appendix 7): E2 and progesterone (Hapangama 2008; 50 participants); FSH and LH (Lima 2006; 49 participants). Even though we only identified one study for each of these markers, the findings are consistent with other studies in the literature addressing hormonal alterations in endometriosis. We do not therefore recommend further research on the diagnostic accuracy of these biomarkers for endometriosis.

8. Immune system and inflammatory markers

8.1. Autoantibodies

8.1.a. Anti-endometrial autoantibodies (anti-endometrial Abs)

Five studies comprising 795 participants assessed the value of anti-endometrial Abs in detecting pelvic endometriosis. Of these, four studies (759 participants, varying phases of menstrual cycle, rASRM I to IV (3 studies) or unclear severity (1 study)) evaluated IgG anti-endometrial Abs using various immunofluorescence methods and different definitions of a positive test.The estimates for sensitivity ranged from 0.56 to 0.87 and for specificity from 0.57 to 0.93. The mean sensitivity and specificity of all these evaluations were 0.81 (95% CI 0.76 to 0.87) and 0.75 (95% CI 0.46 to 1.00), which did not meet the criteria for either a replacement or triage test. Forest plots (Figure 8) and the ROC plot (Figure 9) showed a high degree of heterogeneity for estimates of both sensitivity and specificity. An additional study (36 participants, cycle phase not reported, rASRM I to IV) demonstrated that anti-endometrial Abs of a specific molecular weight (MW) were differentially expressed in endometriosis, and the expression of at least one of the antibodies with MWs of 26 kDa, 34 kDa or 42 kDa had a sensitivity of 1.00 (95% CI 0.81 to 1.00) and a specificity of 0.39 (95% CI 0.17 to 0.64) (Gorai 1993) (Figure 9; Figure 8). This study could not be added to the meta-analysis as the definition of the index test was different, and we considered it separately. The same study assessed an alternate set of antibodies with MWs of 28 kDa, 38 kDa and 64 kDa, the expression of which was not altered in presence of endometriosis. A further study (80 participants, cycle phase not reported, rASRM I to IV) also demonstrated that the serum levels of anti-endometrial Abs were comparable between control and endometriosis groups (Ozhan 2014).

Figure 8. Forest plot of anti-endometrial Abs for detection of endometriosis. Plot shows study-specific estimates of sensitivity and specificity (squares) with 95% CI (black line), country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

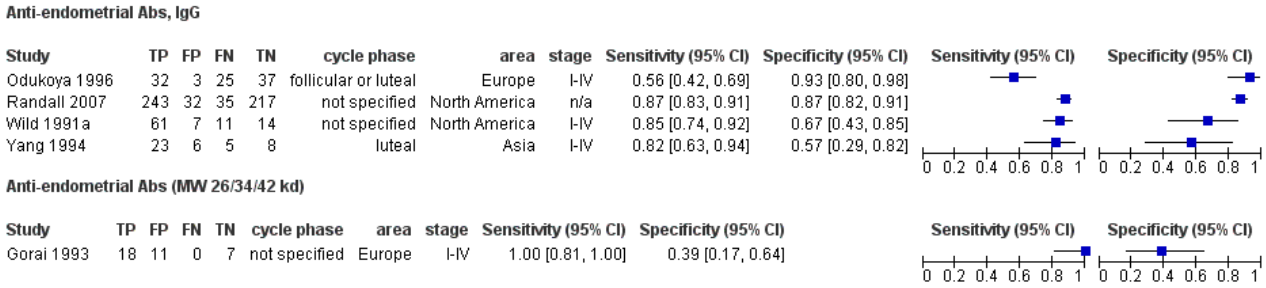
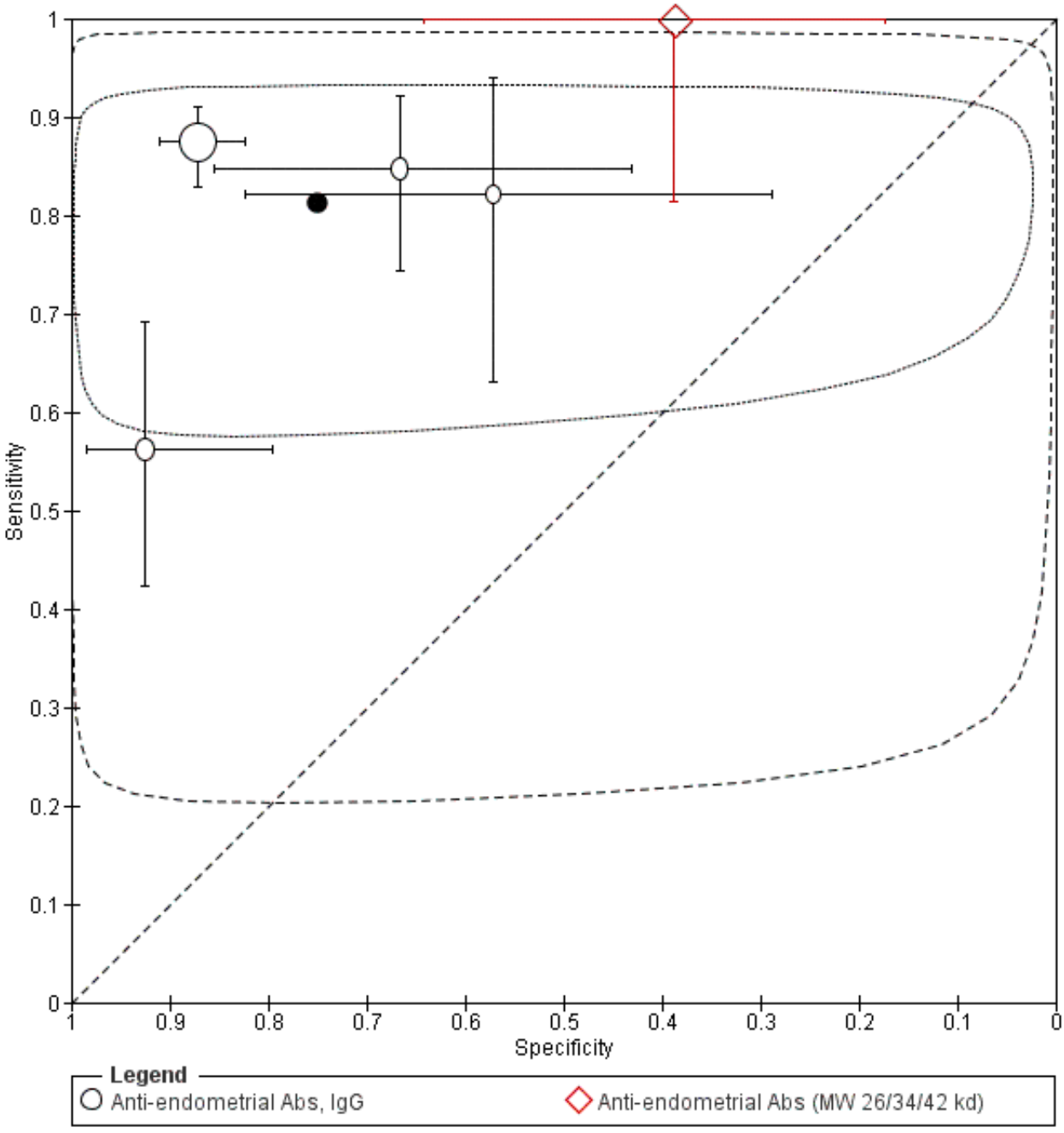


Figure 9. Summary ROC plot of anti-endometrial Abs, IgG for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different sets of antibodies tested. The bars correspond to 95% CIs of each individual evaluation. The solid black circle represents the mean sensitivity and specificity, which is surrounded by a 95% confidence region (dotted line) and by 95% prediction region (dashed line). Meta-analysis was performed for 4 studies (the data for Anti-endometrial Abs (MW 26/34/42 kd) were not included, considering it as a separate test).



8.1.b. Anti-laminin autoantibodies (anti-laminin-1 Abs)

One study (68 participants, cycle phase not reported, rASRM I to IV) evaluated an accuracy of anti-laminin-1 Abs in detecting pelvic endometriosis. Using a cut-off threshold of > 1 U/ml, the test had a sensitivity of 0.40 (95% CI 0.26 to 0.57) and a specificity of 0.88 (95% CI 0.70 to 0.98). Although there is insufficient evidence to have certainty regarding the role of anti-laminin-1 Abs as a marker for endometriosis, these data suggest it is of limited value.

8.1.c. Autoantibodies that exhibited no differential expression in endometriosis

Two additional types of autoantibodies, anti-sperm and anti-zona pellucida Abs, were evaluated in association with minimal endometriosis in one study (98 participants, luteal cycle phase, rASRM I) (Szczepanska 2001a). The levels of these antibodies did not significantly differ in women with and without endometriosis; however, further data from additional studies for broader spectrum of disease is required to draw meaningful conclusions.

8.2. Chemokines

8.2.a. CCR1 (C-C motif receptor 1)

None of the eligible studies assessed the performance of CCR1 as a single test for detecting endometriosis. This biomarker was a part of a panel that constitutes a combined blood test for endometriosis, as presented below under 'Combined tests'.

8.2.b. MCP-1 (monocyte chemotactic protein-1)

One study assessed the diagnostic accuracy of MCP-1 in pelvic endometriosis (101 participants, menstrual cycle phase, rASRM I to IV) and demonstrated a sensitivity of 0.65 (95% CI 0.51 to 0.77) and a specificity of 0.61 (95% CI 0.45 to 0.76). Four other studies (361 participants, various phases of menstrual cycle) revealed that MCP-1 levels were not altered by a wide spectrum of pelvic endometriosis, rASRM I to IV (Drosdzol-Cop 2012b; Kim 2008; Paiva 2014) or by only minimal-mild disease, rASRM I-II (Kalu 2007). Based on the available evidence, MCP-1 in blood appears to have little value as a diagnostic test for endometriosis.

8.3. Other Cytokines

8.3.a. IFN-γ (interferon-gamma)

One study (45 participants, follicular cycle phase, rASRM I to IV) evaluated IFN-γ and demonstrated a sensitivity of 0.68 (95% CI 0.48 to 0.84) and a specificity of 0.65 (95% CI 0.38 to 0.86) for the diagnosis of pelvic ultrasound negative endometriosis using a cut-off value of < 76.00 pg/ml. Another five studies (455 participants, rASRM I to IV) demonstrated no difference in IFN-γ levels in women with and without pelvic endometriosis in the follicular phase (Hassa 2009), follicular or luteal phase (Podgaec 2007; Seeber 2008), or when the cycle phase was not specified (Matalliotakis 2003a; Wu 1998). In view of the data available, IFN-γ appears to be unreliable as a test for pelvic endometriosis.

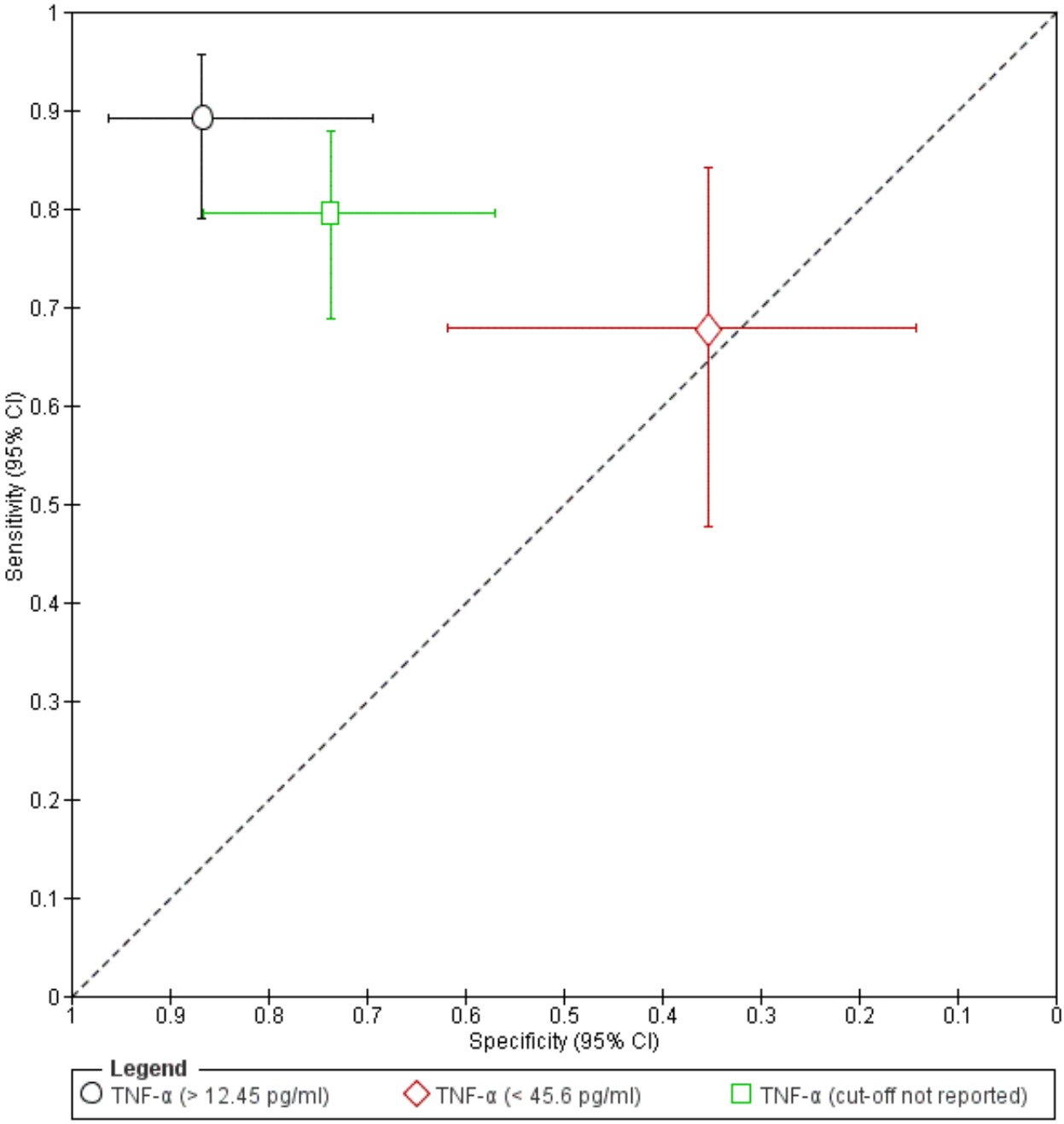
8.3.b. MIF (macrophage migration inhibitory factor)

One study evaluated the value of MIF in detecting pelvic endometriosis (93 participants, follicular or luteal cycle phase, rASRM I to IV), and showed a sensitivity of 0.65 (95% CI 0.51 to 0.78) and a specificity of 0.66 (95% CI 0.49 to 0.80) at a cut-off threshold of > 0.57 ng/ml. Three studies (322 participants, menstrual cycle phase not reported, rASRM I to IV) reported that MF levels were not altered in pelvic endometriosis in follicular or luteal cycle phase (Seeber 2008) or when the cycle phase was not specified (Ozhan 2014; Paiva 2014), suggesting that MIF has little value in diagnosing endometriosis.

8.3.c. TNF-α (tumour necrosis factor alpha)

Three studies evaluated the accuracy of TNF-α in detecting pelvic endometriosis (256 participants, rASRM I to IV), (Figure 10). Two studies evaluated diagnostic test performance in the follicular phase, using contradictory cut-off values of above 12.45 pg/ml (Foda 2012) and below 45.60 pg/ml (Vodolazkaia 2012), and another study assessed the test in luteal cycle phase with no reported cut-off value (Mihalyi 2010). The estimates of sensitivity ranged from 0.68 to 0.89 and the estimates of specificity ranged from 0.35 to 0.87 (Summary of findings 1). We did not perform a meta-analysis because of the diverse definitions of a positive test. Alternatively, eight studies (633 participants, various phases of menstrual cycle) showed unchanged levels of TNF-α in blood in a wide spectrum of pelvic endometriosis, rASRM I to IV (Da Silva 2014; Drosdzol-Cop 2012a; Othman 2008; Podgaec 2007; Seeber 2008; Vercellini 1993; Yagmur 2013) or in only minimal-mild disease, rASRM I-II (Kalu 2007). These conflicting results indicate that TNF-α for the detection of endometriosis is unlikely to be clinically useful.

Figure 10. Summary ROC plot of TNF-α for detection of endometriosis. Each point represents the pair of sensitivity and specificity for each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The bars correspond to 95% CIs of each individual evaluation. The data were not assessed by meta-analysis.



8.3.d. Other cytokines that exhibited no differential expression in endometriosis

Additional cytokines evaluated in the included studies (Appendix 7) were Epo (Yagmur 2013; 55 participants) and sGM-CSF (Matalliotakis 2003a; Othman 2008; Paiva 2014; 287 participants). While the data is scarce for Epo, it is sufficient to suggest that GM-CSF is an inadequate marker for the diagnosis of pelvic endometriosis.

8.4. Immune cells

8.4.a. Neutrophils, NLR (neutrophil-to-lymphocyte ratio), WBC (white blood cells)

One study (100 participants, menstrual phase, rASRM I to IV) evaluated the accuracy of immune cells in diagnosing pelvic endometriosis and reported unsatisfactory estimates for neutrophils at a cut-off of > 4058 cells/ml (sensitivity 0.68, 95% CI 0.53 to 0.80, specificity 0.60, 95% CI 0.45 to 0.74); for NLR (neutrophil lymphocyte ratio), at a cut-off of > 2.19 (sensitivity 0.76, 95% CI 0.62 to 0.87, specificity 0.82, 95% CI 0.69 to 0.91) and for WBC at a cut-off of > 6400 cells/ml (sensitivity 0.64, 95% CI 0.49 to 0.77, specificity 0.54, 95% CI 0.39 to 0.68). Other studies reported negative findings for these biomarkers, specifically for neutrophils and NLR (Yavuzcan 2013); 94 participants, cycle phase not reported, rASRM III to IV) and for WBC (3 studies, 222 participants, follicular or undetermined cycle phase, rASRM I to IV in Gogacz 2014 and Tuten 2014a), or rASRM III to IV in Yavuzcan 2013). These data indicate that WBC levels are not reliable as a diagnostic test for endometriosis, whilst the data for neutrophils and NLR are discouraging but scant.

8.4.b. Immune cells that exhibited no differential expression in endometriosis

We investigated additional peripheral immune cells and found them to be similar in women with and without endometriosis, as presented in Appendix 7. The tested markers from this subgroup included:

1. lymphocytes (Gogacz 2014; Hassa 2009; Matveeva 1990; Yavuzcan 2013; 352 participants);
2. B-lymphocytes (Iwasaki 1993; Maeda 2002a; Zhang 2006a; 223 participants);
3. monocytes/macrophages (Maeda 2002a; 54 participants);
4. NK cells (Hassa 2009; Iwasaki 1993; Maeda 2002a; Zhang 2006a; 320 participants);
5. NKR CD158b+ (KIR2DL2+NK) and NKR CD94+ (Maeda 2002b; Zhang 2006a; 206 participants);
6. T-lymphocytes and specific T-cell populations:
   a. T-cells (Iwasaki 1993; Maeda 2002a; Matveeva 1990; Zhang 2006a; 6 data sets, 342 participants);
   b. T-inducers (Iwasaki 1993; 45 participants);
   c. T-helpers (Hassa 2009; Iwasaki 1993; Maeda 2002a; Matveeva 1990; Mier-Cabrera 2011; Zhang 2006a; 501 participants);
   d. T-suppressors (Hassa 2009; Maeda 2002a; Matveeva 1990; Mier-Cabrera 2011; Zhang 2006a; 6 data sets, 456 participants);
7. Treg cells (regulatory T cells) (Gogacz 2014; Olkowska-Truchanowicz 2013; 3 data sets, 74 participants);
8. haemoglobin (Yavuzcan 2013; 94 participants);
9. MPV (Yavuzcan 2013; 94 participants);
10. platelet count (Yavuzcan 2013; 94 participants);
11. PLR (Yavuzcan 2013; 94 participants).

This evidence clearly indicates that most of the evaluated peripheral blood mononuclear cells have no role as a diagnostic marker for endometriosis. The finding is consistent with the general theme of literature addressing other components of full blood count (haemoglobin, platelets, MPV). Therefore, except for the unexplored phenotypes of Treg cells, we do not recommend further research on the diagnostic accuracy of these biomarkers for endometriosis.

8.5. Interleukins

8.5.a. IL-1β (interleukin-1β)

One study (45 participants, follicular cycle phase, rASRM I to IV) evaluated the diagnostic role of IL-1β in ultrasound negative pelvic endometriosis, showing a sensitivity of 0.82 (95% CI 0.63 to 0.94) and a specificity of 0.35 (95% CI 0.14 to 0.62) for the cut-off value of < 0.90 pg/ml. Four additional studies (248 participants, various cycle phases) showed that IL-1β remained unchanged in a wide spectrum of pelvic endometriosis, rASRM I to IV (Bedaiwy 2002; Oku 2004; Szubert 2014), or in only minimal-mild disease, rASRM I-II (Kalu 2007). Taken together, these results demonstrate that IL-1β has a limited value in detecting pelvic endometriosis.

8.5.b. IL-4 (interleukin - 4)

One study reported the diagnostic accuracy of IL-4 (50 women, follicular cycle phase, rASRM I to IV), showing inadequate estimates for both sensitivity and specificity for a cut-off value ≥ 3.00 pg/ml (0.64, 95% CI 0.45 to 0.80 and 0.65, 95% CI 0.38 to 0.86, respectively). Two other studies reported negative data for this biomarker (195 participants, rASRM I to IV) in either the follicular cycle phase or follicular and luteal cycle phase (Hassa 2009; Podgaec 2007), indicating that IL-4 is unlikely to be an accurate diagnostic test for endometriosis.

8.5.c. IL-6 (interleukin-6)

Eight studies including 12 data sets with a total of 726 participants assessed the diagnostic accuracy of IL-6 for endometriosis. All the studies evaluated pelvic endometriosis (rASRM I to IV in 6 studies and rASRM I-II in 2 studies), but were performed at various phases of the menstrual cycle and utilised different cut-off values (Figure 11). The cut-offs varied from > 1.03 pg/ml to > 25.75 pg/ml, whilst one study used a cut-off of < 10.00 pg/ml. We only included three studies (309 participants, of varying cycle phase, rASRM I to IV) in a meta-analysis, which revealed the summary sensitivity and specificity of 0.63 (95% CI 0.52 to 0.75) and 0.69 (95% CI 0.57 to 0.82) for the cut-off threshold > 1.90 to 2.00 pg/ml. The test did not satisfy the criteria for either a replacement or triage test. Forest plots (Figure 12) and the ROC plot (Figure 13) showed a high degree of heterogeneity for diagnostic estimates, ranging from 0.20 to 0.89 for sensitivity and from 0.66 to 0.80 for specificity. Individual studies evaluated other cut-off thresholds, as presented in Summary of findings 1. Studies reported the highest diagnostic estimates for cut-off value > 12.20 pg/ml (95 participants, follicular cycle phase, rASRM I to IV) with a sensitivity of 0.95 (95% CI 0.87 to 0.99) and a specificity of 0.83 (95% CI 0.65 to 0.94), which met the criteria for a replacement test; however, wide confidence intervals, especially for specificity, advises caution in interpreting these results (Foda 2012). Two studies compared different cut-off values, specifically > 1.03 pg/ml versus > 1.90 pg/ml versus > 2.60 pg/ml (Othman 2008) and > 2.00 pg/ml versus > 4.00 pg/ml versus > 7.50 pg/ml (Bedaiwy 2002); however, all had wide overlapping confidence intervals and presented inconclusive results (Figure 14). In contrast, six other studies (473 participants, various phases of menstrual cycle) demonstrated that IL-6 levels were not affected by the presence of endometriosis when considering different spectra of disease: rASRM I to IV (Drosdzol-Cop 2012a; Seeber 2008; Somigliana 2004), rASRM I-II (Kalu 2007) or rASRM III to IV (Jee 2008; Suen 2014). Although the reports are conflicting, further testing of IL-6 in the

follicular cycle phase at a cutoff value of > 12.20 pg/ml could reveal some diagnostic benefit.

Figure 11. Forest plot of IL-6 (all the included evaluations) for detection of endometriosis. Plot shows estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

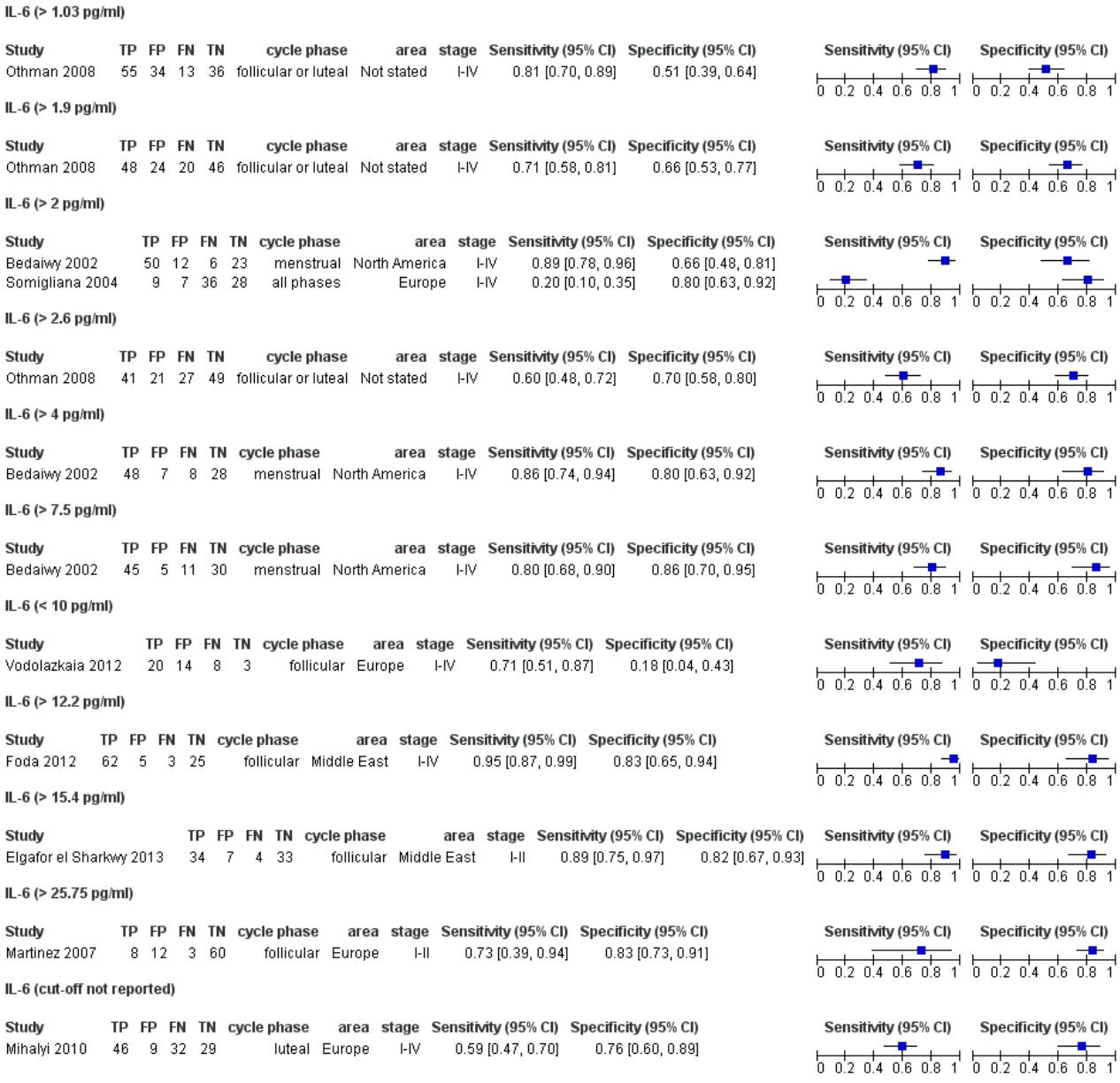


Figure 12. Forest plot of IL-6 with cut-off values above 1.9-2 pg/ml for detection of endometriosis. Plot shows estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

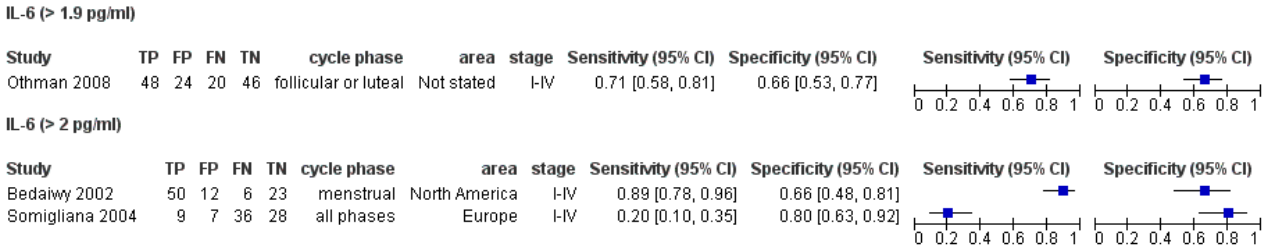


Figure 13. Summary ROC plot of IL-6 with cut-off values ranging > 1.9-2 pg/ml for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size the shape designates the tests with different sets of antibodies tested. The bars correspond to 95% CIs of each individual evaluation. The solid black circle represents the summary sensitivity and specificity.

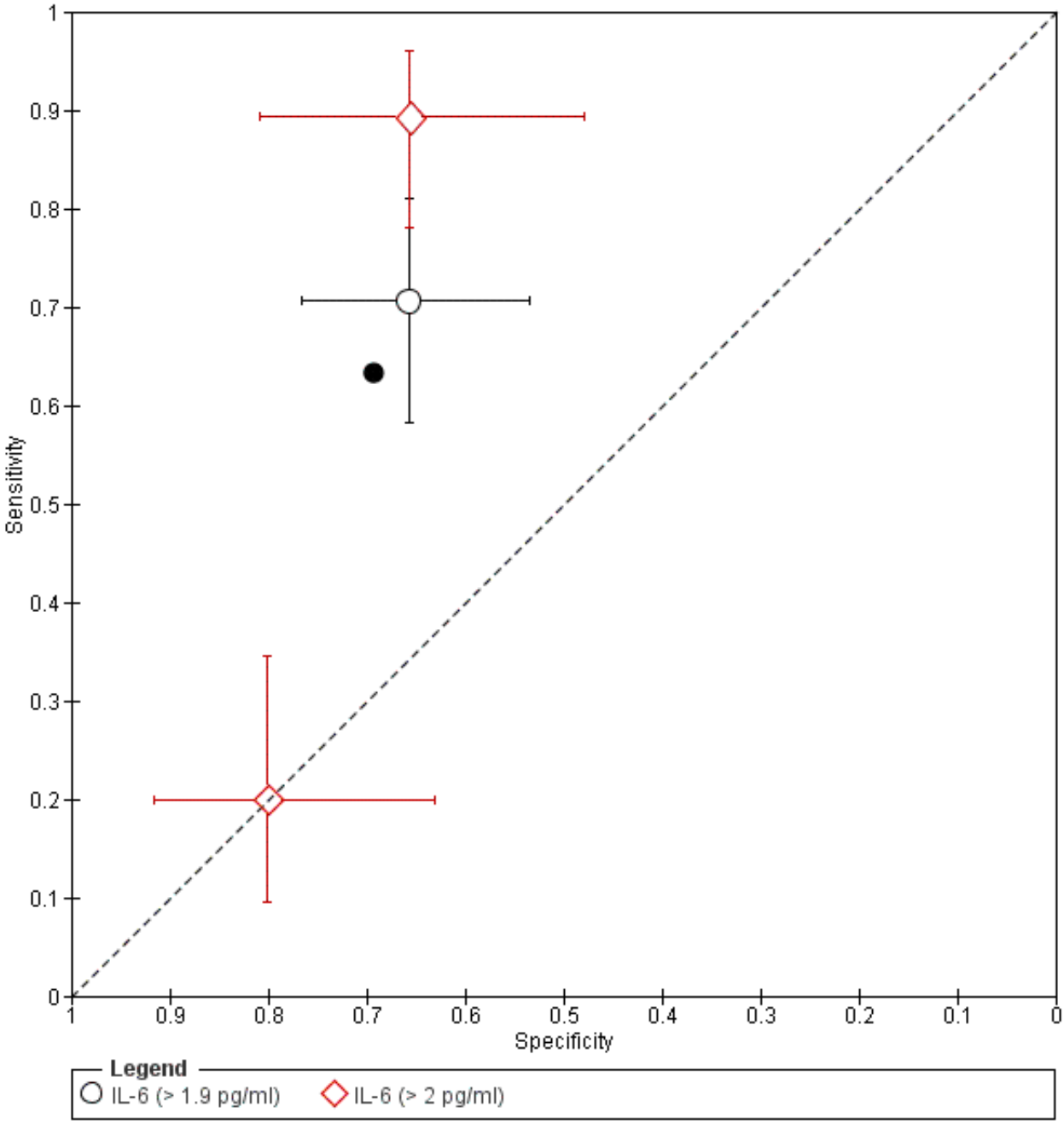
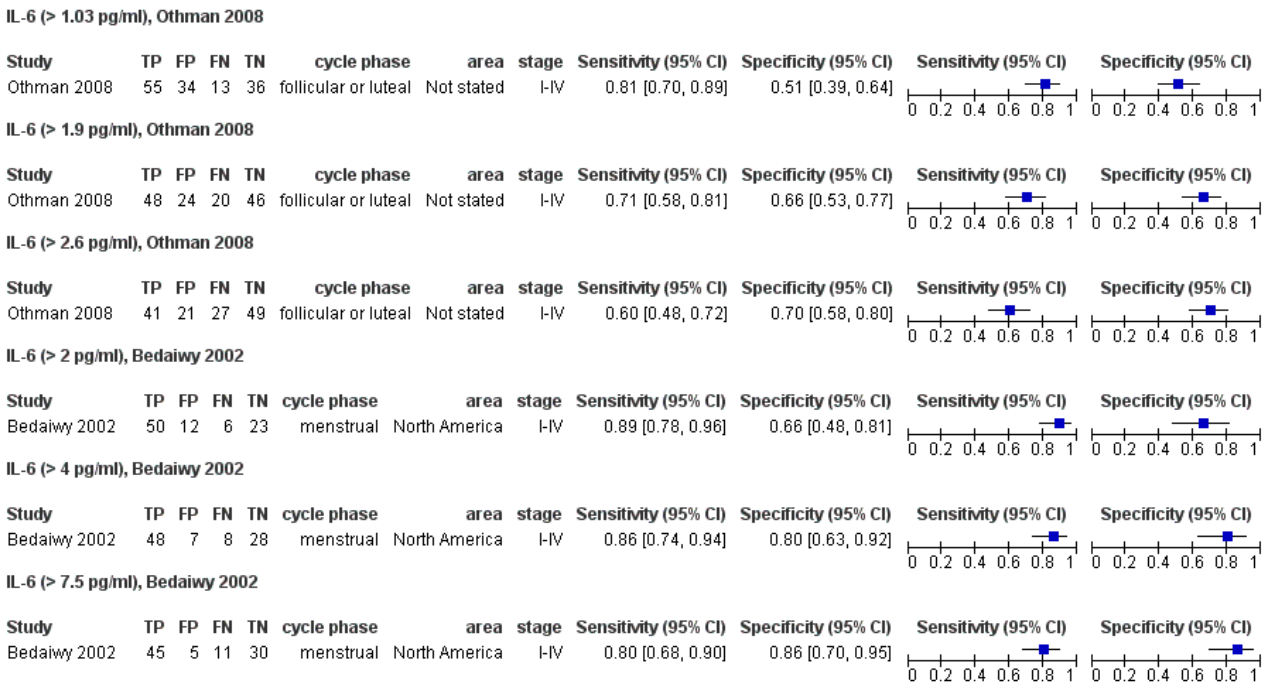


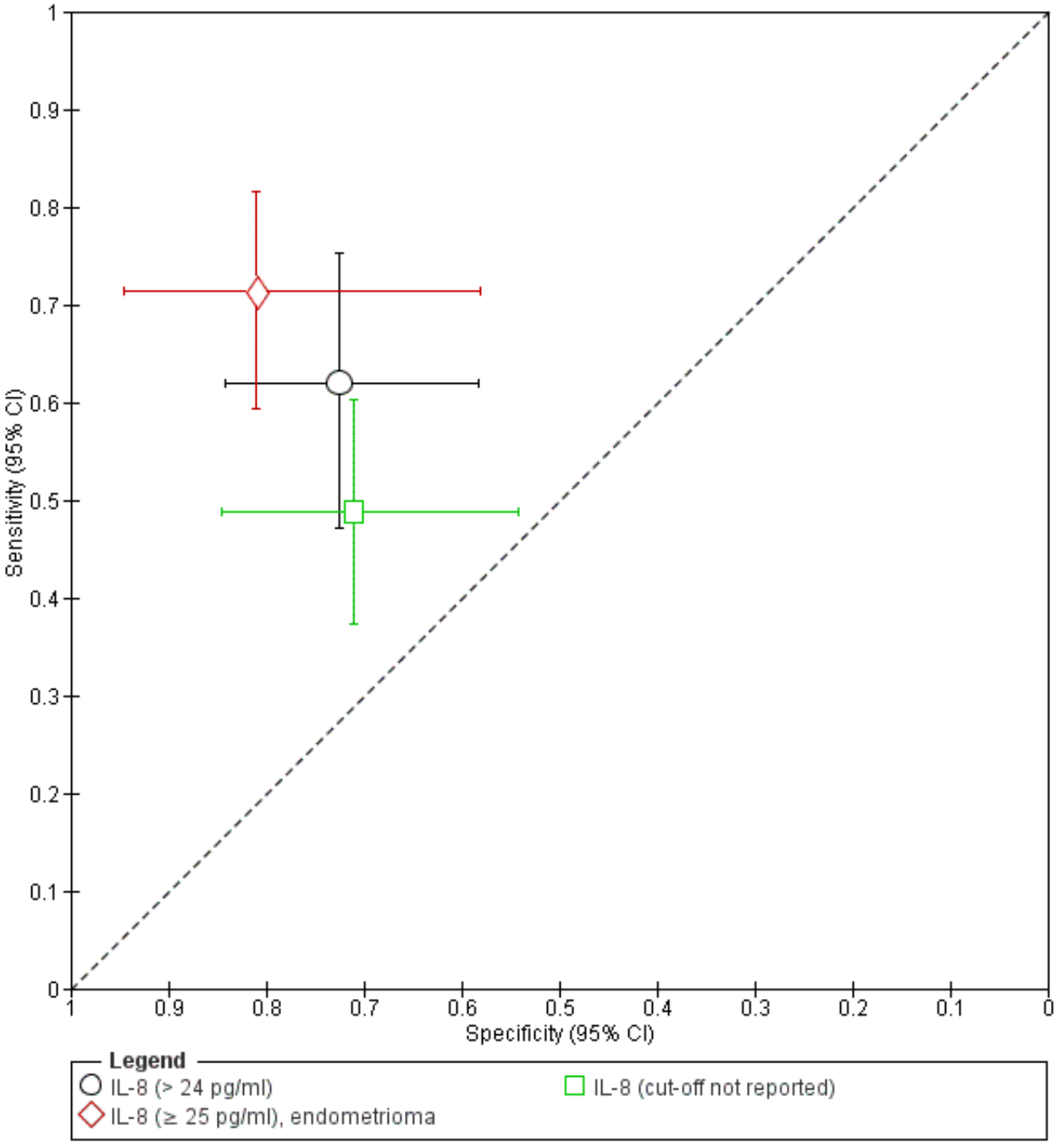
Figure 14. Forest plot of direct comparisons of IL-6 for detection of endometriosis performed between different cut-off values in 2 separate studies. Plot shows the estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.



8.5.d. IL-8 (interleukin-8)

Two studies explored the accuracy of IL-8 in diagnosing pelvic endometriosis (217 participants, various cycle phases, rASRM I to IV), of which one utilised a cut-off value of > 24.00 pg/ml and one did not report a diagnostic threshold. Due to the heterogeneity of the methodology, we could not perform a meaningful meta-analysis. The estimates of sensitivity ranged between 0.49 and 0.62 and of specificity between 0.71 and 0.73 (Summary of findings 1; Figure 15). An additional study (91 participants, cut-off value > 25.00 pg/ml) specifically addressed ovarian endometriosis versus other benign ovarian cysts; we present its findings below under 'Blood biomarkers that could differentiate ovarian endometrioma from other benign ovarian cysts in women of reproductive age'. Five other studies (389 participants,various cycle phases) reported negative data for IL-8 in pelvic endometriosis rASRM I to IV (Barcz 2002; Gazvani 1998; Othman 2008; Ozhan 2014), rASRM I-II (Kalu 2007), or rASRM III to IV (Calienno 2008). These conflicting results suggest that IL-8 has questionable value as a diagnostic test for endometriosis.

Figure 15. Summary ROC plot of IL-8 for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The bars correspond to 95% CIs of each individual evaluation. The data were not assessed by meta-analysis.



8.5.e. Interleukins that exhibited no differential expression in endometriosis

The included studies reported negative findings for endometriosis with additional types of interleukins, as presented in Appendix 7.

1. IL-2 Drosdzol-Cop 2012b; Hassa 2009; Li 2005; Othman 2008; Podgaec 2007; 433 participants).
2. IL-10 (Andreoli 2011; Braun 1996; Hassa 2009; Podgaec 2007; 305 participants).
3. IL-12 (Andreoli 2011; Bedaiwy 2002; Fairbanks 2009; Kubatova 2013; Suen 2014; Szczepanska 2001b; 433 participants).

4. IL-13 (Bedaiwy 2002; 53 participants).
5. IL-15 (Othman 2008; 138 participants, biomarker below detection limit in both groups).
6. IL-16 (Lin 2005; Zhang 2005a; 88 participants).
7. IL-17 (Andreoli 2011; Paiva 2014; 181 participants).
8. IL-18 (Fairbanks 2009; Glitz 2009; Oku 2004; Zhang 2005b; 301 participants).
9. IL-23 (Andreoli 2011: 80 participants).

Many of these interleukins were evaluated by more than one study and are unlikely to be worthy of further investigation as diagnostic biomarkers for endometriosis.

8.6. Other immune/inflammatory markers

8.6.a. sCD23 (soluble CD23)

One study evaluated the diagnostic performance of sCD23 for pelvic endometriosis (97 participants, follicular or luteal cycle phase, rASRM I to IV), demonstrating a sensitivity of 0.25 (95% CI 0.14 to 0.38) and a specificity of 0.93 (95% CI 0.80 to 0.98). Another study (102 participants, menstrual or follicular cycle phase, rASRM I to IV) demonstrated no significant difference in sCD23 levels in women with and without endometriosis, indicating that sCD23 is likely to have limited diagnostic value, albeit further studies are needed to support this statement (Ramos 2012).

8.6.b. Copeptin, vasopressin surrogate

One study evaluated the accuracy of copeptin in detecting pelvic endometriosis (87 participants, cycle phase not reported, rASRM I to IV), showing a sensitivity of 0.65 (95% CI 0.50 to 0.78) and a specificity of 0.58 (95% CI 0.41 to 0.74). There is insufficient data to draw meaningful conclusions on the findings from this single study.

8.6.c. hs-CRP (high sensitive C-reactive protein)

Three studies including six data sets (506 participants, various menstrual cycle phases, rASRM I to IV) explored the diagnostic accuracy of hs-CRP for pelvic endometriosis, using various cut-off thresholds, ranging from 0.60 mg/l to 438 mg/l. Five data sets included overlapping populations. We did not perform a meta-analysis because of the methodological heterogeneity. Diagnostic estimates from the included studies varied, with sensitivities ranging from 0.41 to 0.83 and specificities ranging from 0.47 to 0.87 (Summary of findings 1; Figure 16; Figure 17). Studies reported the highest estimates for hs-CRP with a cut-off of > 438 mg/l (1 study, 95 participants in follicular cycle phase) with a sensitivity of 0.83 (95% CI 0.72 to 0.91) and a specificity of 0.87 (95% CI 0.69 to 0.96) (Foda 2012). One group compared hs-CRP diagnostic estimates in the menstrual, follicular, luteal or combination of all phases of the cycle in a total of 295 participants (Vodolazkaia 2011). The authors established the best cut-off values in a ROC analysis, which varied depending on the cycle phase. The diagnostic estimates were low for all evaluations, ranging from 0.54 to 0.68 for sensitivity and from 0.47 to 0.64 for specificity (Figure 16). Six additional studies (1333 participants, various cycle phases) demonstrated no difference in expression levels of CRP or hs-CRP in a wide spectrum of pelvic endometriosis, rASRM I to IV (Dayangan Sayan 2013; Riley 2007; Szubert 2014; Thubert 2014; Tuten 2014a) or when the severity of the disease was not reported (Kianpour 2012). The methods included the hs-CRP assay in Thubert 2014 (834 participants) and the CRP assay in the other studies (499 participants). A comparison between the two assay methods concluded that hs-CRP assay had higher diagnostic accuracy than the traditional CRP assay (Vodolazkaia 2011). Collectively, the available evidence suggests that CRP evaluated by either method is not a reliable biomarker for detecting endometriosis.

Figure 16. Forest plot of hs-CRP for detection of endometriosis. Plot shows the estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

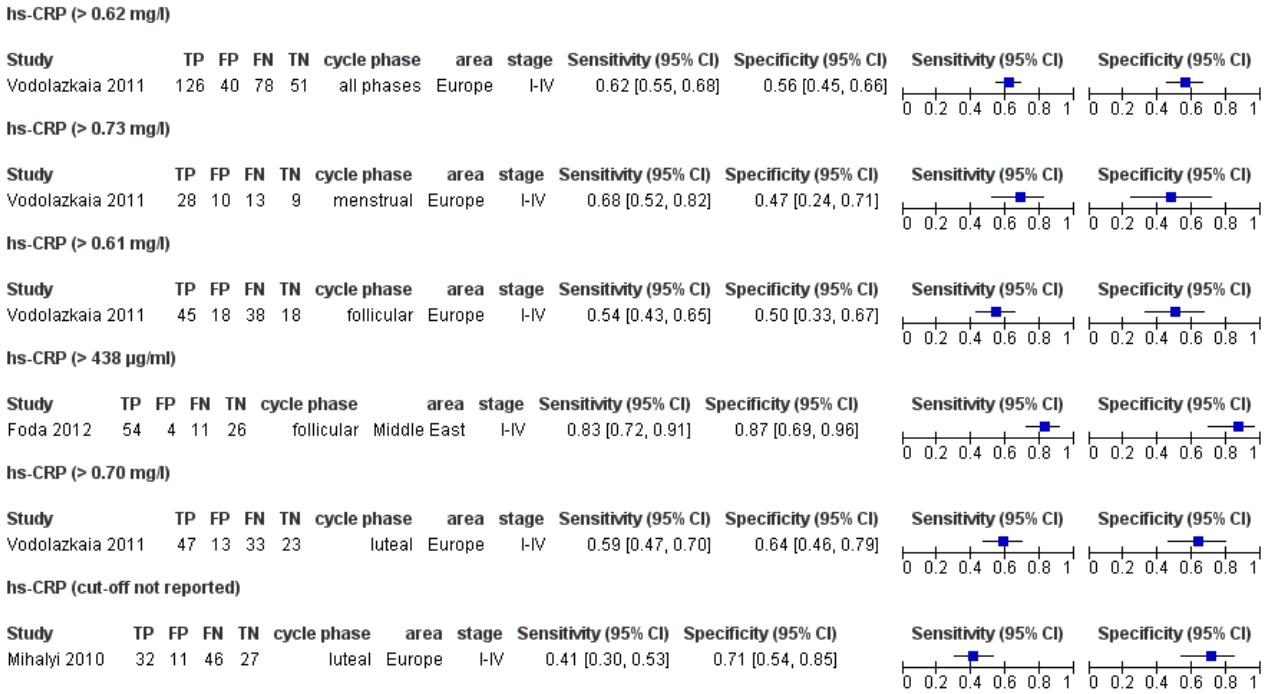
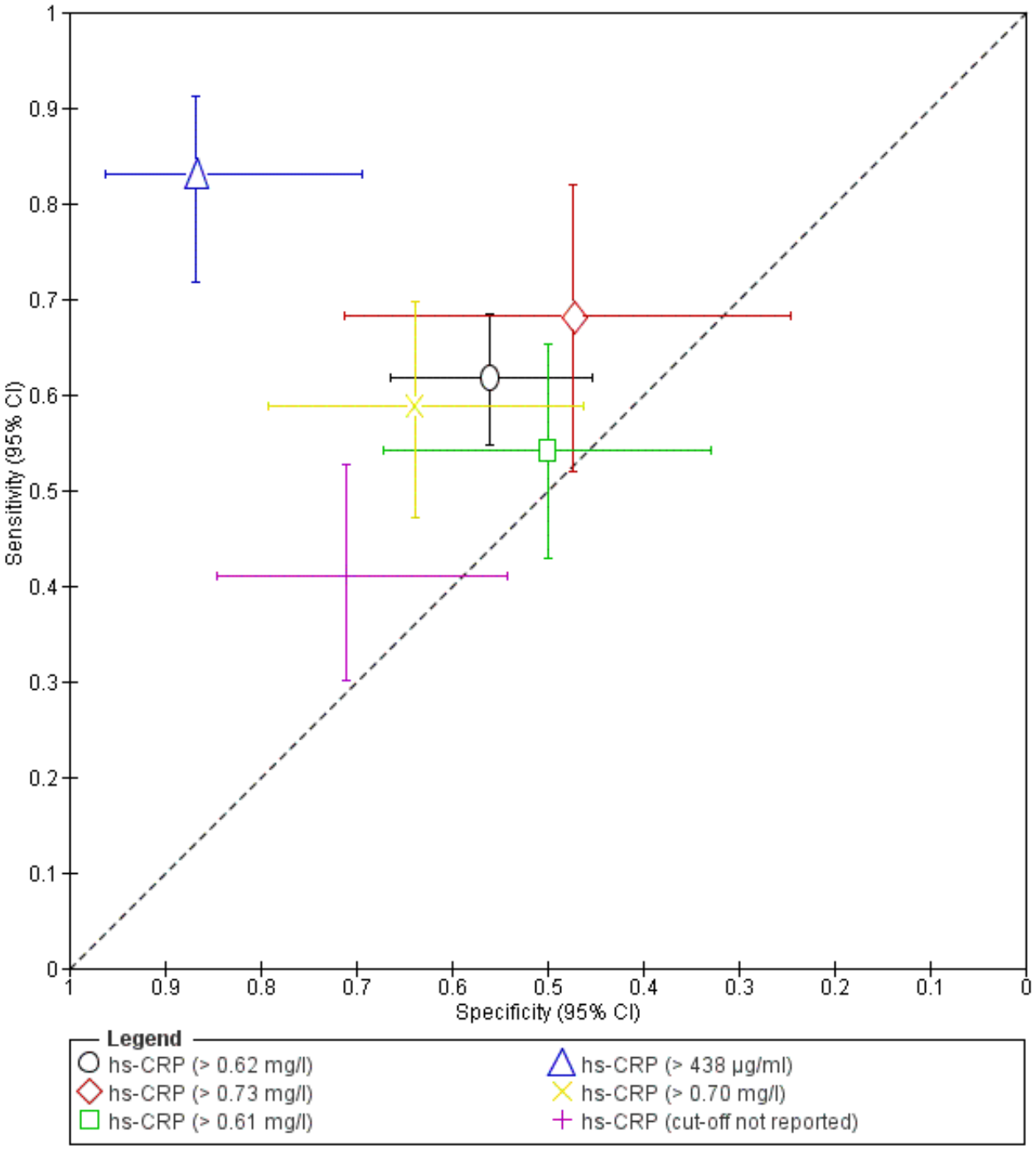


Figure 17. Summary ROC plot of hs-CRP for detection of endometriosis. Each point represents the pair of sensitivity and specificity for each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The bars correspond to 95% CIs of each individual evaluation. Five evaluations (excluding 1 with a cut-off > 438 µg/ml) were performed on overlapping populations. The data were not assessed by meta-analysis.



8.6.d. Other immune system and inflammatory markers that exhibited no differential expression in endometriosis

The other immune system and inflammatory biomarkers for which only negative data were reported included (Appendix 7):

1. C3a (Fassbender 2009; 160 participants);
2. sHLA-I (De Placido 1998; 30 participants);
3. immunoglobulins: IgA, IgG (Matveeva 1990; 119 participants);

4. MPO (Da Silva 2014; 17 participants);
5. NAG (Da Silva 2014; 17 participants);
6. PGE2 (Khan 2012; 86 participants);
7. phospholipid fatty acids (Khanaki 2012; 138 participants, 16 fatty acids);
8. PLA2G2A (Kocbek 2014a; 91 participants);
9. RANTES (Kalu 2007; Markham 1997a; 72 participants).

Except for RANTES, all other biomarkers from this group were assessed in a single study, and their association with endometriosis remains unclear.

9. Nerve growth markers

9.1. Nerve growth markers that exhibited no differential expression in endometriosis

One study (101 participants, cycle phase not reported, rASRM I to IV) evaluated four nerve growth markers (CNTF, GDNF, NGF, NT4), showing no association between any of these tests and endometriosis (Paiva 2014), as presented in Appendix 7. Future research needs to confirm the expression of these biomarkers in endometriosis and their value in the diagnosis of the disease.

10. Other peptides/proteins shown to influence key events implicated in endometriosis

10.1. Follistatin

Follistatin was only evaluated in the context of ovarian endometrioma and is presented below under 'Blood biomarkers that could differentiate ovarian endometrioma from other benign ovarian cysts in women of reproductive age'.

10.2. STX-5 (syntaxin-5)

One study reported on the diagnostic performance of STX-5 in endometriosis (80 participants, cycle phase not reported, rASRM I to IV), using a cut-off of > 55 ng/ml, with a sensitivity of 0.78 (95% CI 0.66 to 0.88) and a specificity of 0.70 (95% CI 0.46 to 0.88); Ozhan 2014. The diagnostic estimates did not meet the criteria for an adequate diagnostic test (replacement or triage), but additional studies need to support this observation.

10.3. Other peptides/proteins that exhibited no differential expression in endometriosis

Three additional proteins (DBP (Borkowski 2008; Ferrero 2005a; 171 participants), enolase and PDPK1 (Ozhan 2014; 80 participants) were evaluated for their association with endometriosis. Their serum levels did not distinguish women with endometriosis from controls (Appendix 7).

11. Oxidative stress markers

11.1. Carbonyls

One study (Rosa E Silva 2014) assessed the diagnostic role of carbonyls in endometriosis (108 participants, cycle phase and spectrum of the disease not reported), demonstrating a sensitivity of 0.94 (95% CI 0.85 to 0.98) and a specificity of 0.51 (95% CI 0.35 to 0.67) at a cut-off value of < 14.9 µm. This approaches the criteria for a SnOUT triage test, but large high quality studies need to confirm this finding.

11.2. PON-1 (paraoxonase-1)

One study (Verit 2008) reported on the ability of PON-1 to diagnose pelvic endometriosis (87 participants, follicular cycle phase, rASRM I to IV). The diagnostic estimates were high enough to fulfil the criteria for a replacement test (sensitivity 0.98, 95% CI 0.89 to 1.00 and specificity 0.80, 95% CI 0.64 to 0.91), using a cut-off threshold of < 141.5 U/ml. Further studies are required to confirm this finding.

11.3. Thiols

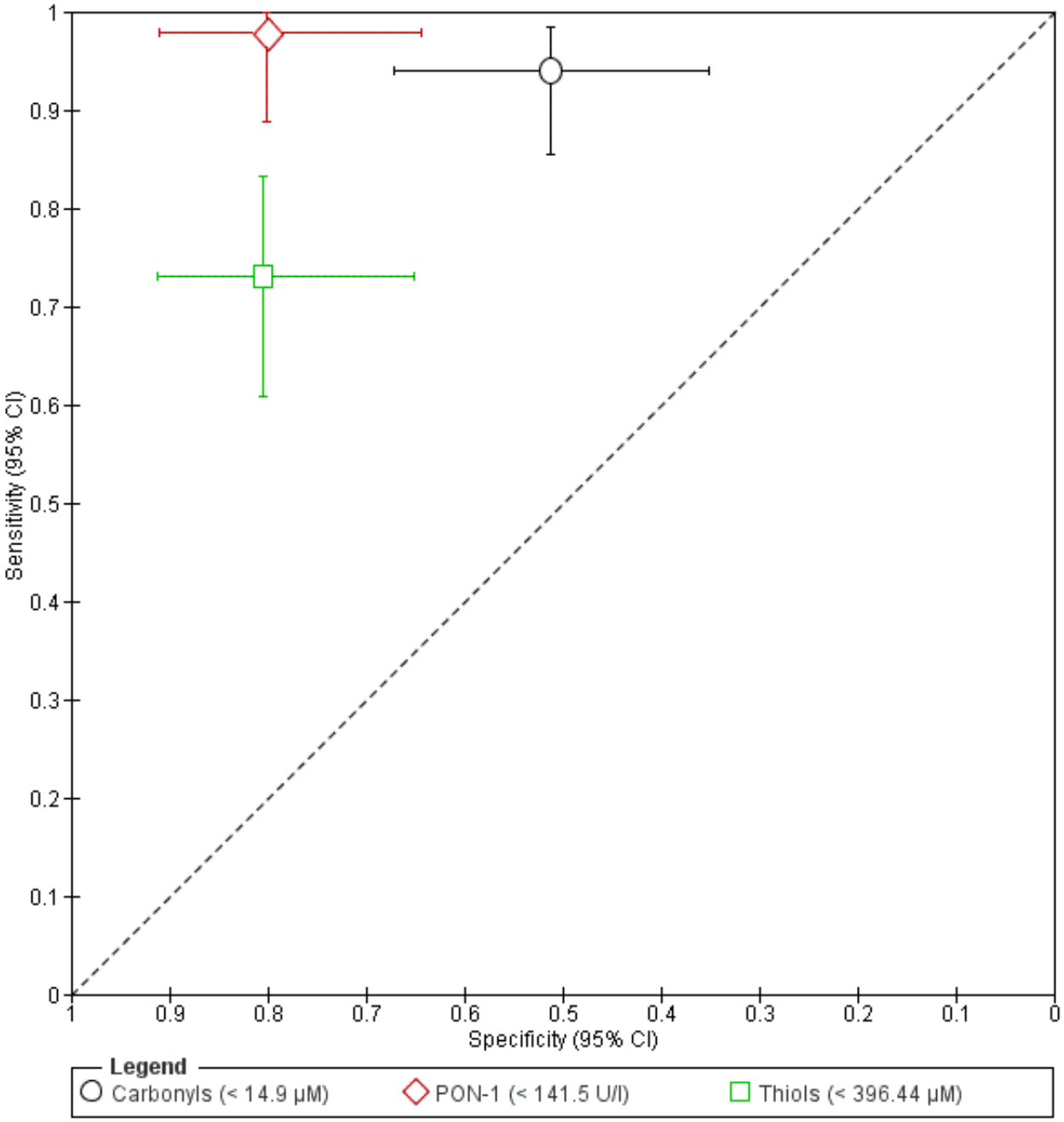
One study (Rosa E Silva 2014) tested the accuracy of thiols in detecting pelvic endometriosis (108 participants, cycle phase and spectrum of the disease not reported), showing a sensitivity of 0.73 (95% CI 0.61 to 0.83) and a specificity of 0.80 (95% CI 0.65 to 0.91) at a cut-off value of < 396.44 µm. Further data is required before a comment can be made on its diagnostic role.

11.4. Oxidative stress markers that exhibited no differential expression in endometriosis

Additional oxidative stress markers that appeared to have comparable levels in women with and without endometriosis (Appendix 7) included ascorbic acid and malondialdehyde (Mier-Cabrera 2011; 62 participants); GSH, nitrotyrosine, SOD3 and vitamin E (Paiva 2014; 101 participants); HSP70 (Khan 2013; Lambrinoudaki 2009; 116 participants); and IMA and TRX (Lambrinoudaki 2009; 66 participants).

Although the diagnostic studies for these biomarkers are encouraging (Figure 18), there is insufficient evidence to draw meaningful conclusions regarding any biomarker from this group, and further research is recommended to confirm the positive and negative findings presented above.

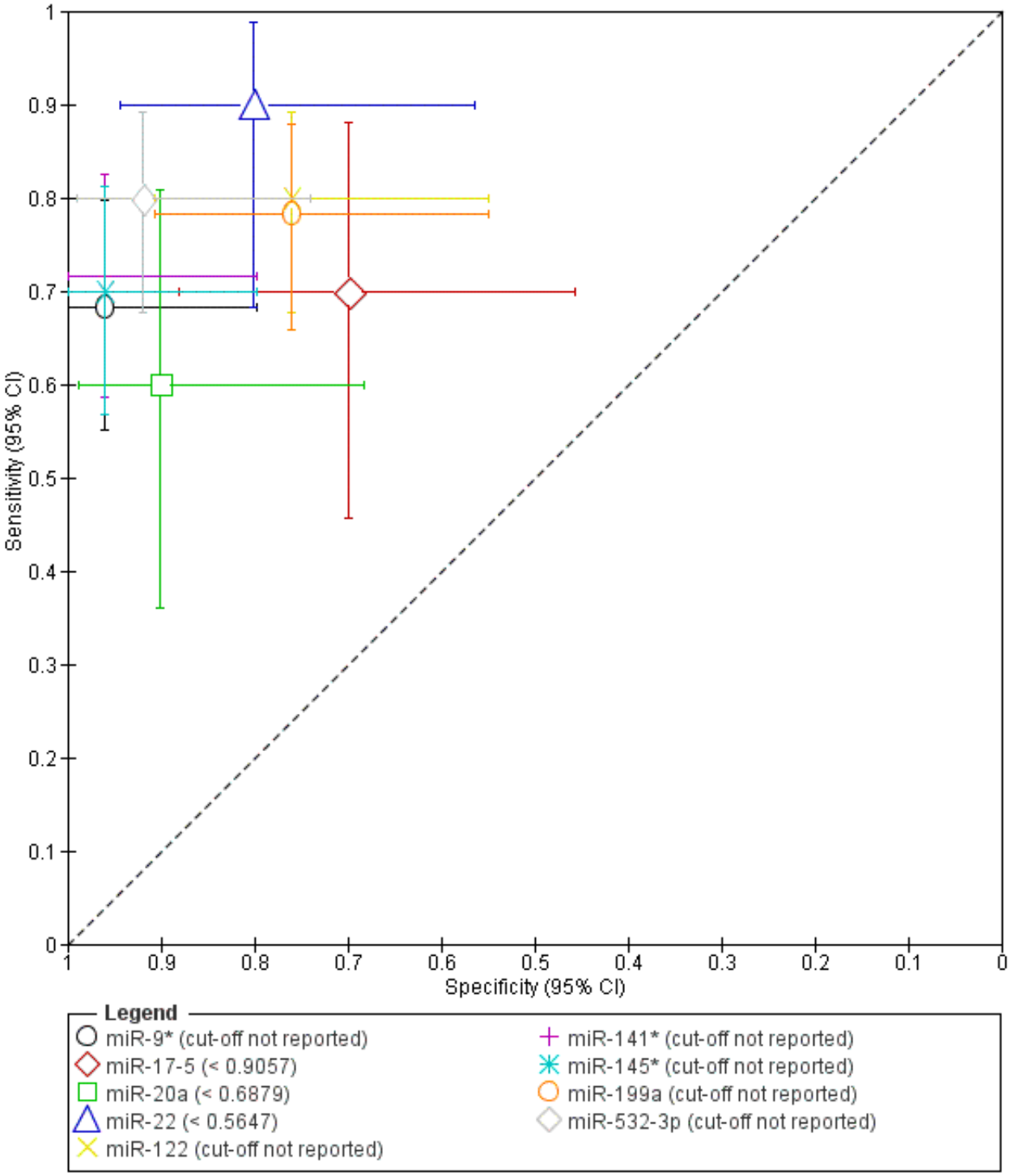
Figure 18. Summary ROC plot of oxidative stress biomarkers for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates different biomarkers from this group, each assessed in a single study. The bars correspond to 95% CIs of each individual evaluation. The data were not assessed by meta-analysis.



12. Post-transcriptional regulators of gene expression (microRNAs)

There were two eligible studies that evaluated the role of microRNAs (miRs) in detecting endometriosis (Figure 19). One study (85 participants, follicular or luteal cycle phase) assessed diagnostic accuracy of six microRNAs in pelvic endometriosis, rASRM I to IV (Wang 2013a).

Figure 19. Summary ROC plot of microRNAs for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates different biomarkers from this group, each assessed in a single study. The bars correspond to 95% CIs of each individual evaluation. The data were not assessed by meta-analysis.



1. miR-9* (sensitivity 0.68, 95% CI 0.55 to 0.80 and specificity 0.96, 95% CI 0.80 to 1.00).
2. miR-122 (sensitivity 0.80, 95% CI 0.68 to 0.89 and specificity 0.76, 95% CI 0.55 to 0.91).

3. miR-141* (sensitivity 0.72, 95% CI 0.59 to 0.83 and specificity 0.96, 95% CI 0.80 to 1.00).
4. miR-145* (sensitivity 0.70, 95% CI 0.57 to 0.81 and specificity 0.96, 95% CI 0.80 to 1.00).
5. miR-199a (sensitivity 0.78, 95% CI 0.66 to 0.88 and specificity 0.76, 95% CI 0.55 to 0.91).
6. miR-532-3p (sensitivity 0.80, 95% CI 0.68 to 0.89 and specificity 0.92, 95% CI 0.74 to 0.99).

The authors did not report the cut-off values for any of the tested biomarkers.

Another group published data on diagnostic performance of three microRNAs (40 participants, follicular or luteal cycle phase) in moderate-severe pelvic endometriosis, rASRM III to IV (Jia 2013): miR-17-5 (sensitivity 0.70, 95% CI 0.46 to 0.88 and specificity 0.70, 95% CI 0.46 to 0.88 for the cut-off of < 0.9057); miR-20a (sensitivity 0.60, 95% CI 0.36 to 0.81 and specificity 0.90, 95% CI 0.68 to 0.99, for the cut-off of < 0.6879) and miR-22 (sensitivity 0.90, 95% CI 0.68 to 0.99 and specificity 0.80, 95% CI 0.56 to 0.94 for the cut-off of < 0.5647). Both Jia 2013 and Wang 2013a varied in laboratory methodology and approach to quantifying and analysing the data. MiR-9*, miR-141* and miR-145* met the criteria of a SpIN triage test, and miR-532-30, miR-20a and miR-22 approached these criteria. While several microRNAs show some promise as diagnostic markers for endometriosis, the two published studies identified completely independent microRNA biomarkers. These results require further validation in a large, well-defined population with a wide spectrum of disease, using a standardised reproducible methodology.

13. Tumour markers

13.1. CA-15.3 (cancer antigen-15.3)

Two studies (Tuten 2014a; Muscatello 1992) (207 participants, various phases of menstrual cycle, rASRM I to IV) assessed the diagnostic performance of CA-15.3 in endometriosis with substantially heterogeneous estimates. Each study used different cut-off thresholds, so we did not include them in a meta-analysis. In both studies the levels of CA-15.3 were not significantly different in women with and without endometriosis, although the diagnostic test estimates were calculated. None of the included studies exhibited high diagnostic accuracy, with sensitivities ranging from 0.65 to 0.04 and specificities from ranging 0.62 to 0.92.

13.2. CA-19.9 (cancer antigen-19.9)

Seven studies (8 data sets, 793 participants, various phases of menstrual cycle, rASRM I to IV) explored the role of CA-19.9 in pelvic endometriosis. Three evaluations were performed in an overlapping population (Harada 2002; Kurdoglu 2009; Mabrouk 2012; Mihalyi 2010; Somigliana 2004; Tuten 2014a; Vodolazkaia 2012). Studies used very diverse cut-off thresholds, ranging from > 7.5 U/ml to > 37.0 U/ml, while two studies did not report the cut-off. In view of inconsistencies in the methods, a meta-analysis was legitimate only for three studies with a total of 309 participants that assessed CA-19.9 for a cut-off value > 37.0 U/ml. The summary sensitivity was 0.36 (95% CI 0.26 to 0.45) and the summary specificity was 0.87 (95% CI 0.75 to 0.99) (Harada 2002; Kurdoglu 2009; Somigliana 2004) (Figure 20). One study from this subgroup (80 participants, all cycle phases, rASRM I to IV) demonstrated that the serum levels of CA-19.9 were comparable between the control and endometriosis groups (Somigliana 2004). Other evaluations of this biomarker were reported separately, and none presented clinically meaningful diagnostic estimates, with a sensitivity ranging from 0.36 to 0.73 and a specificity from 0.56 to 0.90 (see Summary of findings 1; Figure 21). An additional study (118 participants, follicular cycle phase; Guerriero 1996a) addressed only ovarian endometriosis versus other benign ovarian cysts and is reported separately (see 'Blood biomarkers that could differentiate ovarian endometrioma from other benign ovarian cysts in women of reproductive age').

Figure 20. Summary ROC plot of CA-19.9 with a cut-off value > 37 U/ml for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different sets of antibodies tested. The bars correspond to 95% CIs of each individual evaluation. The solid black circle represents the summary sensitivity and specificity.

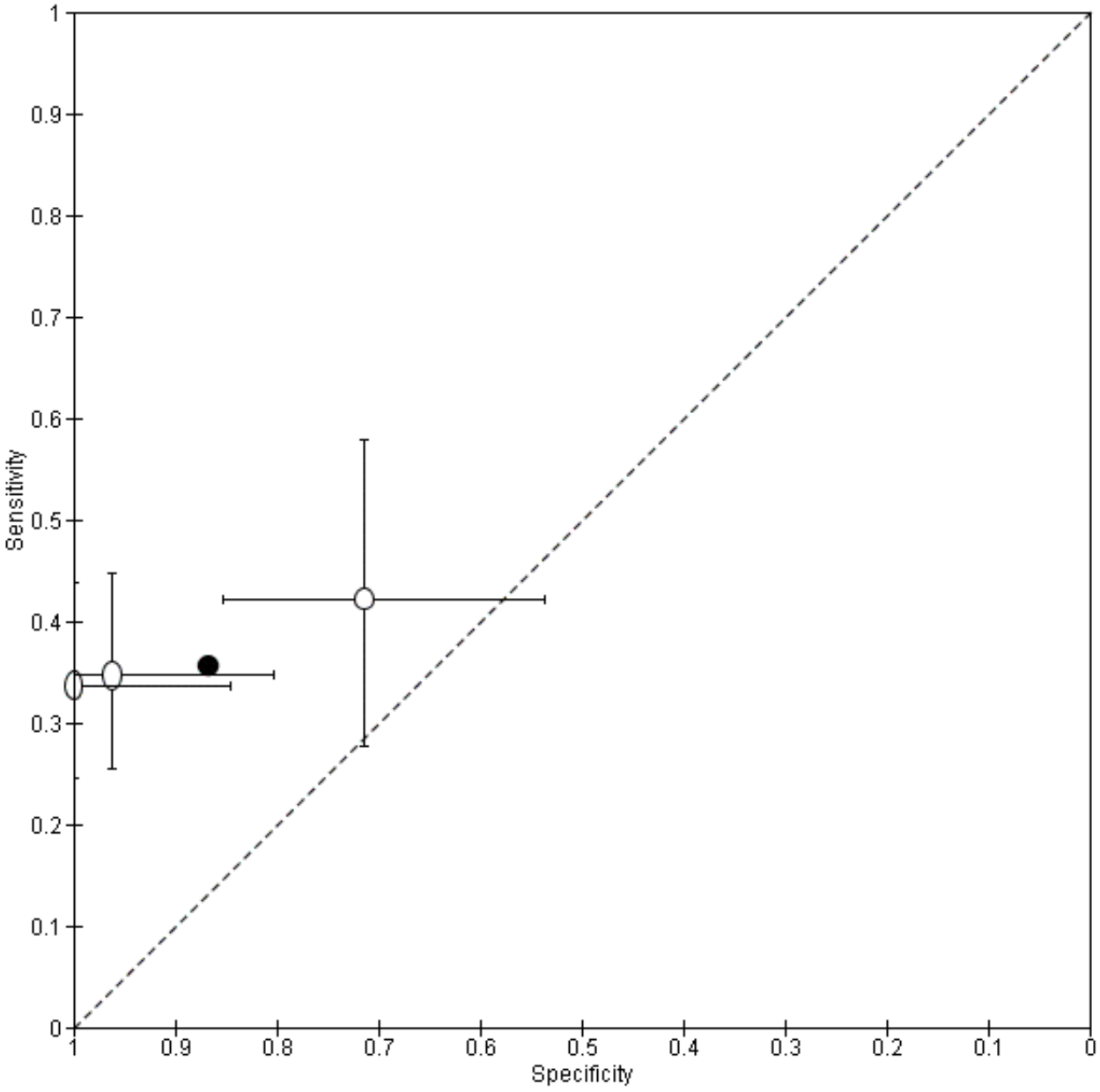
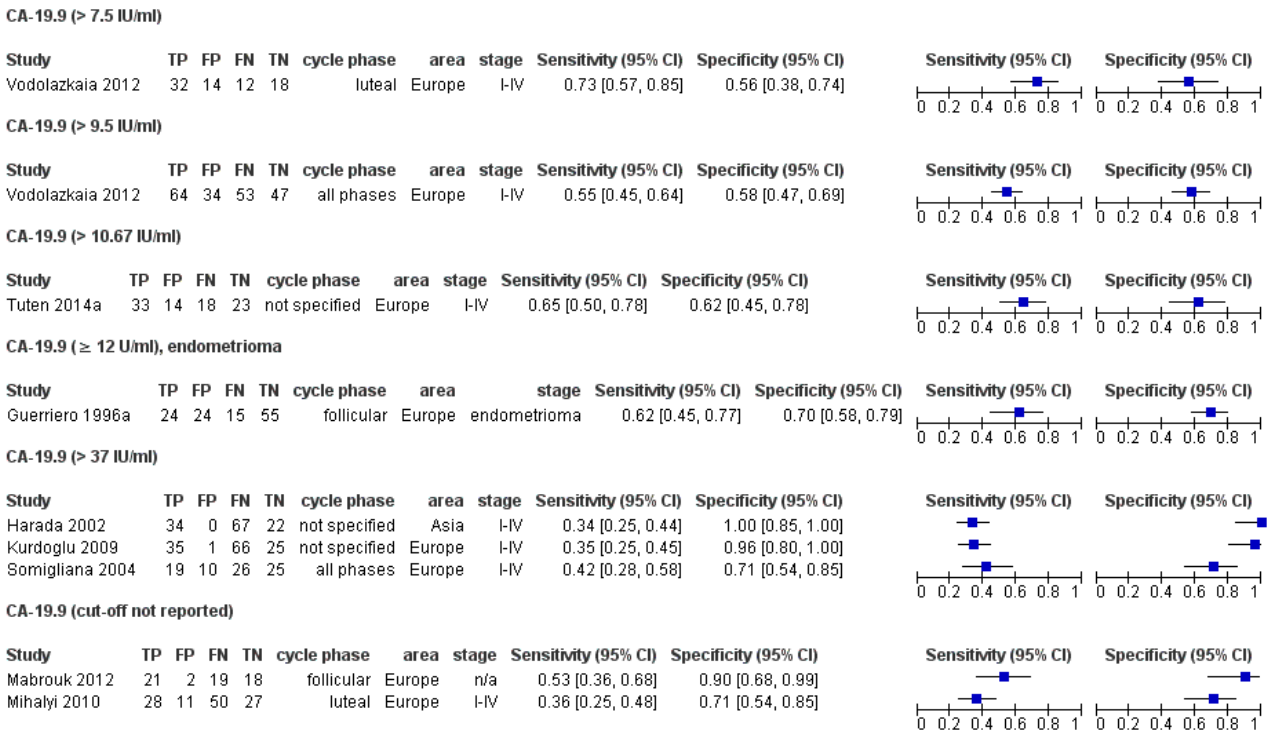


Figure 21. Forest plot of CA-19.9 (all the evaluations) for detection of endometriosis. Plot shows estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.



13.3. CA-72 (TAG-72) (cancer antigen-72 or tumour associated glycoprotein-72)

Two studies (Molo 1994; Muscatello 1992) evaluated the role of CA-72 in detecting pelvic endometriosis in varying phases of menstrual cycle, using different cut-off values. One study (35 participants in the follicular cycle phase, rASRM stage not reported) reported a sensitivity 0.05 (95% CI 0.00 to 0.26) and a specificity 0.75 (95% CI 0.48 to 0.93) for the cut-off > 4.0 U/ml; Molo 1994. A second study (119 participants in luteal cycle phase, rASRM I to IV) demonstrated a sensitivity of 0.09 (95% CI 0.04 to 0.17) and a specificity of 0.89 (95% CI 0.75 to 0.97) for the cut-off value > 6.0 U/ml; Muscatello 1992. A meta-analysis was not performed as the methodology was heterogeneous, but both presented unacceptably low sensitivities indicating no clinically applicable alteration of blood CA-72 levels in the presence of endometriosis, which shows that this biomarker is not suitable for detecting disease.

13.4. CA-125 (cancer antigen-125)

Forty-five studies including 60 data sets with a total of 5534 participants explored the accuracy of CA-125 in the diagnosis of endometriosis. The included evaluations were performed in varying phases of the menstrual cycle for different spectra of the disease and using a broad range of cut-off thresholds, from > 10.0 U/ml to > 42.0 U/ml (Summary of findings 1; Figure 22). Since a sufficient number of studies assessed CA-125 for most of the diagnostic cut-offs, the studies for overall pelvic and ovarian endometriosis were included in the analysis for each cut-off threshold, with a subsequent sensitivity analyses after excluding the data for ovarian endometriosis. We grouped the tests by clinically relevant target cut-off ranges as follows.

Figure 22. Forest plot of CA-125 (all the included evaluations) for detection of endometriosis. Plot shows the estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease

assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

CA-125 (> 10 IU/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Rosa E Silva 2007	95	10	53	43	follicular	South America	I-IV	0.64 [0.56, 0.72]	0.81 [0.68, 0.91]

CA-125 (> 11 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Szubert 2012	30	5	14	10	follicular	Europe	I-IV	0.68 [0.52, 0.81]	0.67 [0.38, 0.88]

CA-125 (> 11.5 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Vodolazkaia 2012	24	6	4	11	follicular	Europe	I-IV	0.86 [0.67, 0.96]	0.65 [0.38, 0.86]

CA-125 (> 12.5 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Vodolazkaia 2012	48	24	10	17	all phases	Europe	I-IV	0.83 [0.71, 0.91]	0.41 [0.26, 0.58]

CA-125 (> 12.8 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Gagne 2003a	133	127	40	68	luteal	North America	I-IV	0.77 [0.70, 0.83]	0.35 [0.28, 0.42]

CA-125 (> 13.5 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Vodolazkaia 2012	15	11	4	5	luteal	Europe	I-IV	0.79 [0.54, 0.94]	0.31 [0.11, 0.59]

CA-125 (> 14.7 IU/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Salehpour 2009	23	8	12	17	follicular	Asia	I-IV	0.66 [0.48, 0.81]	0.68 [0.46, 0.85]

CA-125 (> 16 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Ferreira 1994	8	5	15	13	not specified	South America	I-IV	0.35 [0.16, 0.57]	0.72 [0.47, 0.90]
Gurgan 1990	12	6	5	15	luteal	Europe	I	0.71 [0.44, 0.90]	0.71 [0.48, 0.89]
Pittaway 1989	66	5	16	76	follicular	North America	I-IV	0.80 [0.70, 0.88]	0.94 [0.86, 0.98]
Wild 1991a	21	1	51	20	not specified	North America	I-IV	0.29 [0.19, 0.41]	0.95 [0.76, 1.00]

CA-125 (> 17.6 IU/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Foda 2012	39	0	26	30	follicular	Middle East	I-IV	0.60 [0.47, 0.72]	1.00 [0.88, 1.00]

CA-125 (> 20 IU/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Bilibio 2014	32	3	31	31	luteal	South America	I-IV	0.51 [0.38, 0.64]	0.91 [0.76, 0.98]
Kitawaki 2005	347	141	86	201	follicular or luteal	Asia	I-IV	0.80 [0.76, 0.84]	0.59 [0.53, 0.64]
Rosa E Silva 2007	45	1	103	52	follicular	South America	I-IV	0.30 [0.23, 0.38]	0.98 [0.90, 1.00]
Yang 1994	20	6	8	8	luteal	Asia	I-IV	0.71 [0.51, 0.87]	0.57 [0.29, 0.82]

CA-125 (> 20 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996b	23	32	6	40	follicular	Europe	endometrioma	0.79 [0.60, 0.92]	0.56 [0.43, 0.67]
Tokmak 2011	37	17	5	29	follicular	Europe	endometrioma	0.88 [0.74, 0.96]	0.63 [0.48, 0.77]

CA-125 (> 26 IU/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Kitawaki 2005	307	108	126	234	follicular or luteal	Asia	I-IV	0.71 [0.66, 0.75]	0.68 [0.63, 0.73]
Tuten 2014a	44	5	6	32	not specified	Europe	I-IV	0.88 [0.76, 0.95]	0.86 [0.71, 0.95]

CA-125 (> 30 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Acien 1989	23	1	31	13	menstrual	Europe	I-IV	0.43 [0.29, 0.57]	0.93 [0.66, 1.00]
Dayangan Sayan 2013	32	6	18	44	menstrual	Europe	I-IV	0.64 [0.49, 0.77]	0.88 [0.76, 0.95]
Kitawaki 2005	275	89	158	253	follicular or luteal	Asia	I-IV	0.64 [0.59, 0.68]	0.74 [0.69, 0.79]

CA-125 (> 33 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Matalliotakis 2004	20	0	30	50	not specified	Europe	I-IV	0.40 [0.26, 0.55]	1.00 [0.93, 1.00]

CA-125 (> 35 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)

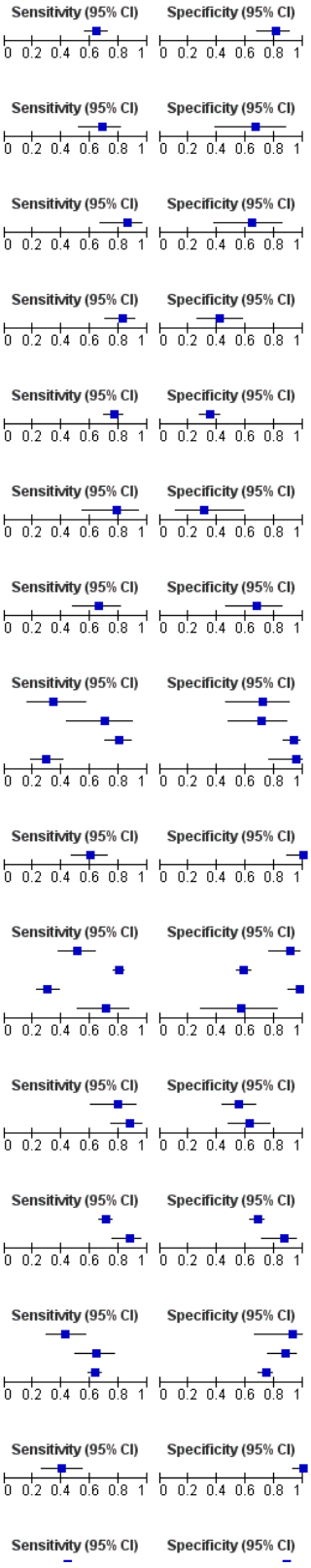
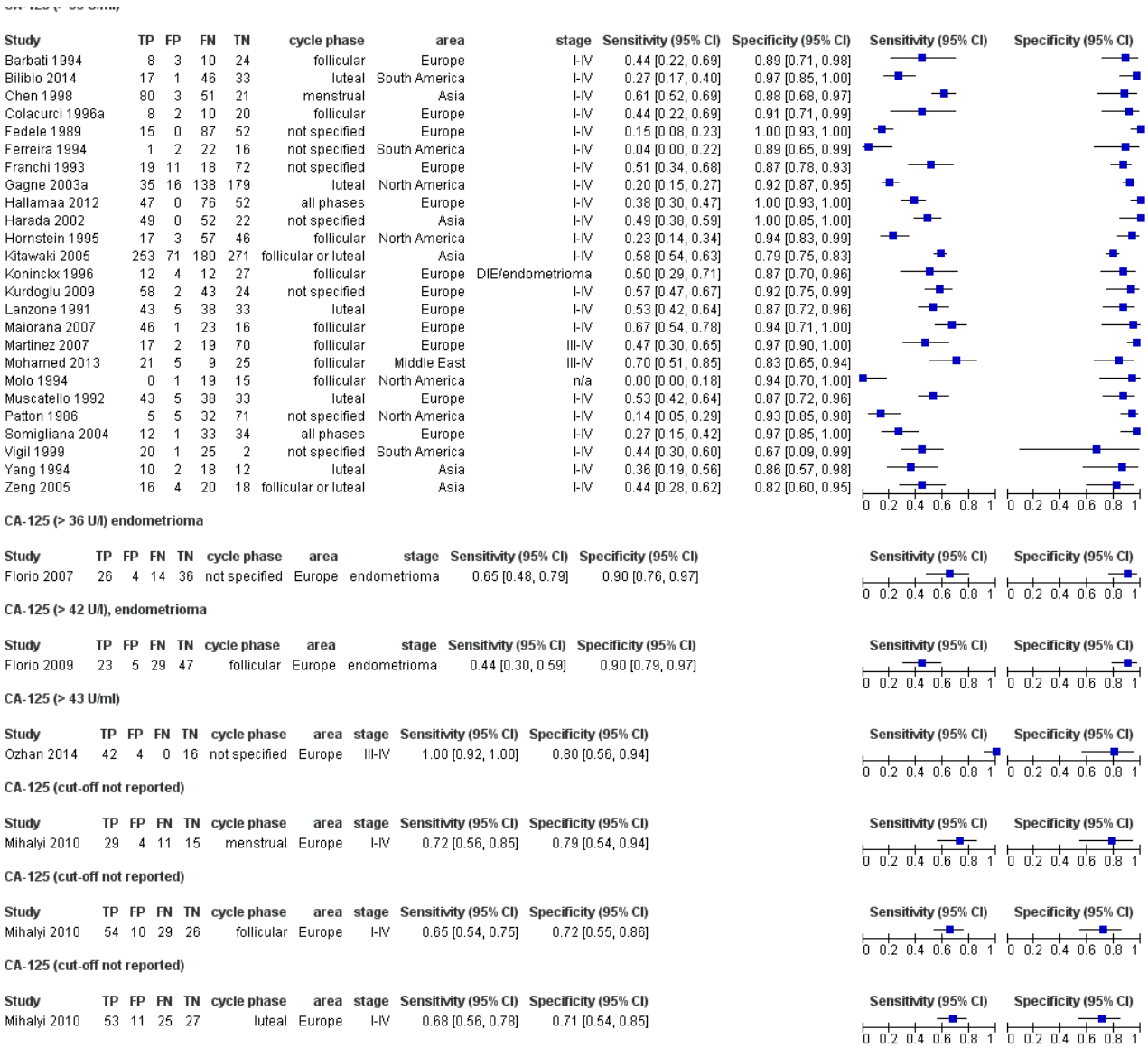


Figure 22. (Continued)



Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Barbati 1994	8	3	10	24	follicular	Europe	I-IV	0.44 [0.22, 0.69]	0.89 [0.71, 0.98]
Bilibio 2014	17	1	46	33	luteal	South America	I-IV	0.27 [0.17, 0.40]	0.97 [0.85, 1.00]
Chen 1998	80	3	51	21	menstrual	Asia	I-IV	0.61 [0.52, 0.69]	0.88 [0.68, 0.97]
Colacurci 1996a	8	2	10	20	follicular	Europe	I-IV	0.44 [0.22, 0.69]	0.91 [0.71, 0.99]
Fedele 1989	15	0	87	52	not specified	Europe	I-IV	0.15 [0.08, 0.23]	1.00 [0.93, 1.00]
Ferreira 1994	1	2	22	16	not specified	South America	I-IV	0.04 [0.00, 0.22]	0.89 [0.65, 0.99]
Franchi 1993	19	11	18	72	not specified	Europe	I-IV	0.51 [0.34, 0.68]	0.87 [0.78, 0.93]
Gagne 2003a	35	16	138	179	luteal	North America	I-IV	0.20 [0.15, 0.27]	0.92 [0.87, 0.95]
Hallamaa 2012	47	0	76	52	all phases	Europe	I-IV	0.38 [0.30, 0.47]	1.00 [0.93, 1.00]
Harada 2002	49	0	52	22	not specified	Asia	I-IV	0.49 [0.38, 0.59]	1.00 [0.85, 1.00]
Hornstein 1995	17	3	57	46	follicular	North America	I-IV	0.23 [0.14, 0.34]	0.94 [0.83, 0.99]
Kitawaki 2005	253	71	180	271	follicular or luteal	Asia	I-IV	0.58 [0.54, 0.63]	0.79 [0.75, 0.83]
Koninckx 1996	12	4	12	27	follicular	Europe	DIE/endometrioma	0.50 [0.29, 0.71]	0.87 [0.70, 0.96]
Kurdoglu 2009	58	2	43	24	not specified	Europe	I-IV	0.57 [0.47, 0.67]	0.92 [0.75, 0.99]
Lanzone 1991	43	5	38	33	luteal	Europe	I-IV	0.53 [0.42, 0.64]	0.87 [0.72, 0.96]
Maiorana 2007	46	1	23	16	follicular	Europe	I-IV	0.67 [0.54, 0.78]	0.94 [0.71, 1.00]
Martinez 2007	17	2	19	70	follicular	Europe	III-IV	0.47 [0.30, 0.65]	0.97 [0.90, 1.00]
Mohamed 2013	21	5	9	25	follicular	Middle East	III-IV	0.70 [0.51, 0.85]	0.83 [0.65, 0.94]
Molo 1994	0	1	19	15	follicular	North America	n/a	0.00 [0.00, 0.18]	0.94 [0.70, 1.00]
Muscatello 1992	43	5	38	33	luteal	Europe	I-IV	0.53 [0.42, 0.64]	0.87 [0.72, 0.96]
Patton 1986	5	5	32	71	not specified	North America	I-IV	0.14 [0.05, 0.29]	0.93 [0.85, 0.98]
Somigliana 2004	12	1	33	34	all phases	Europe	I-IV	0.27 [0.15, 0.42]	0.97 [0.85, 1.00]
Vigil 1999	20	1	25	2	not specified	South America	I-IV	0.44 [0.30, 0.60]	0.67 [0.09, 0.99]
Yang 1994	10	2	18	12	luteal	Asia	I-IV	0.36 [0.19, 0.56]	0.86 [0.57, 0.98]
Zeng 2005	16	4	20	18	follicular or luteal	Asia	I-IV	0.44 [0.28, 0.62]	0.82 [0.60, 0.95]

CA-125 (> 36 U/l) endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2007	26	4	14	36	not specified	Europe	endometrioma	0.65 [0.48, 0.79]	0.90 [0.76, 0.97]

CA-125 (> 42 U/l), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2009	23	5	29	47	follicular	Europe	endometrioma	0.44 [0.30, 0.59]	0.90 [0.79, 0.97]

CA-125 (> 43 U/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Ozhan 2014	42	4	0	16	not specified	Europe	III-IV	1.00 [0.92, 1.00]	0.80 [0.56, 0.94]

CA-125 (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	29	4	11	15	menstrual	Europe	I-IV	0.72 [0.56, 0.85]	0.79 [0.54, 0.94]

CA-125 (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	54	10	29	26	follicular	Europe	I-IV	0.65 [0.54, 0.75]	0.72 [0.55, 0.86]

CA-125 (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	53	11	25	27	luteal	Europe	I-IV	0.68 [0.56, 0.78]	0.71 [0.54, 0.85]

1. CA-125 with the cut-off values of > 10.0 to 14.7 U/ml (5 studies 769 participants, cycle phase varied, rASRM stage varied) had a mean sensitivity of 0.70 (95% CI 0.63 to 0.77) and mean sensitivity of 0.64 (95% CI 0.47 to 0.82), excluding two overlapping evaluations (Figure 23).
2. CA-125 with a cut-off threshold of > 16.0 to 17.6 U/ml (5 studies, 430 participants, cycle phase varied, rASRM stage varied) had a mean sensitivity of 0.56 (95% CI 0.24 to 0.88) and mean specificity of 0.91 (95% CI 0.75 to 1.00) (Figure 24).
3. CA-125 with a cut-off threshold of > 20.0 U/ml (6 studies, 1304 participants, cycle phase varied, rASRM stage varied) had a mean sensitivity of 0.67 (95% CI 0.50 to 0.85) and mean specificity of 0.69 (95% CI 0.58 to 0.80) (Figure 25). This group included two studies that specifically aimed to differentiate ovarian endometriosis from the other benign ovarian masses.
4. CA-125 with a cut-off of > 25.0 to 26.0 U/ml (3 studies, 963 participants, cycle phase varied, rASRM stage varied) had a summary sensitivity of 0.73 (95% CI 0.67 to 0.79) and specificity of 0.70 (95% CI 0.63 to 0.77) (Figure 26). In this group, two studies assessed overall pelvic endometriosis, whilst one study looked at ovarian endometriosis versus other ovarian cysts.
5. CA-125 with a cut-off of > 30.0 to 33.0 U/ml (6 studies, 1206 participants, cycle phase varied, rASRM stage varied) had a mean sensitivity of 0.62 (95% CI 0.45 to 0.79) and specificity of 0.76 (95% CI 0.53 to 1.00) (Figure 27). Two studies included in the analysis focused on differentiation of ovarian endometriosis from the other ovarian cysts.
6. CA-125 with a cut-off threshold of > 35.0 to 36.0 U/ml (27 studies, 3276 participants, cycle phase varied, rASRM stage varied) had a mean sensitivity of 0.40 (95% CI 0.32 to 0.49) and specificity of 0.91 (95% CI 0.88 to 0.94) (Figure 28). Meta-analysis included

two studies differentiating ovarian endometrioma from other ovarian cysts.

7. Only two studies reported CA-125 with a cut-off of > 42.0 to 43.0 U/ml, of which one (104 participants, cut-off value > 42.0 U/ml) assessed the performance of CA-125 in differentiating ovarian endometrioma from other benign ovarian cysts (presented separately). The second study (62 participants, cycle phase not reported, rASRM III to IV, cut-off value > 43.0 U/ml) was not confined to ovarian disease and included any type of endometriosis, demonstrating a sensitivity of 1.00 (95% CI 0.92 to 1.00) and a specificity of 0.80 (95% CI 56 to 0.94). The studies were not combined in a meta-analysis due to the heterogeneity of the included populations and paucity of the data.
8. The cut-off thresholds were not reported in four evaluations of CA-125, three of which had overlapping populations and were presented separately (Summary of findings 1).

Figure 23. Summary ROC plot of CA-125 with cut-off values ranging > 10-14.7 U/ml for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The solid black circle represents the mean sensitivity and specificity, which is surrounded by a 95% confidence region (dotted line) and by 95% prediction region (dashed line). Meta-analysis was performed for 5 studies (the data for 2 evaluations (CA-125 > 11.5 U/ml and CA-125 > 13.5 U/ml were not included as overlapping populations with already included study).

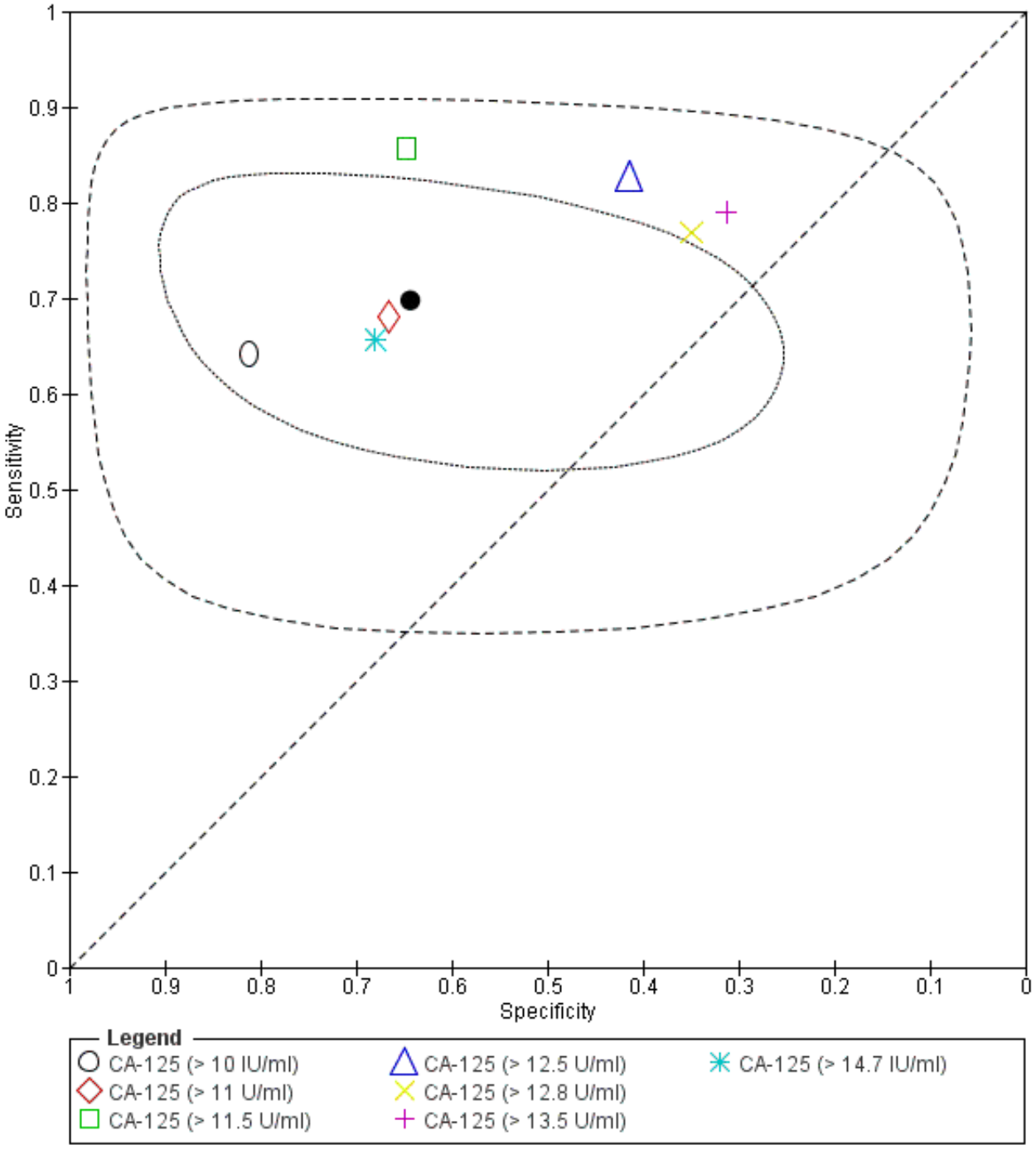


Figure 24. Summary ROC plot of CA-125 with cut-off values ranging > 16-17.6 U/ml for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The solid black circle represents the mean sensitivity and specificity, which is surrounded by a 95% confidence region (dotted line) ad by 95% prediction region (dashed line).

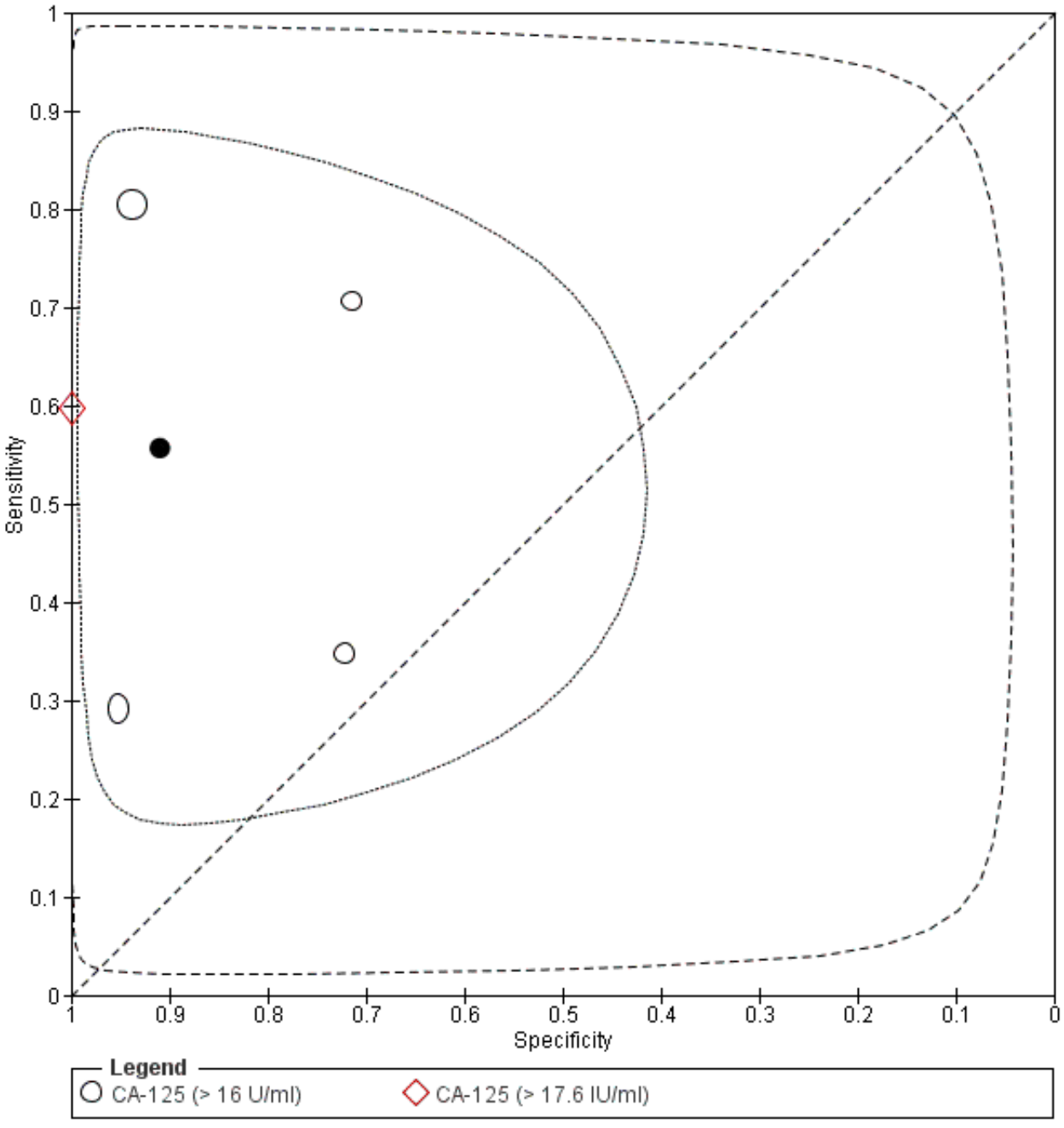


Figure 25. Summary ROC plot of CA-125 with cut-off values > 20 U/ml for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size. The solid black circle represents the mean sensitivity and specificity, which is surrounded by a 95% confidence region (dotted line) and by 95% prediction region (dashed line).

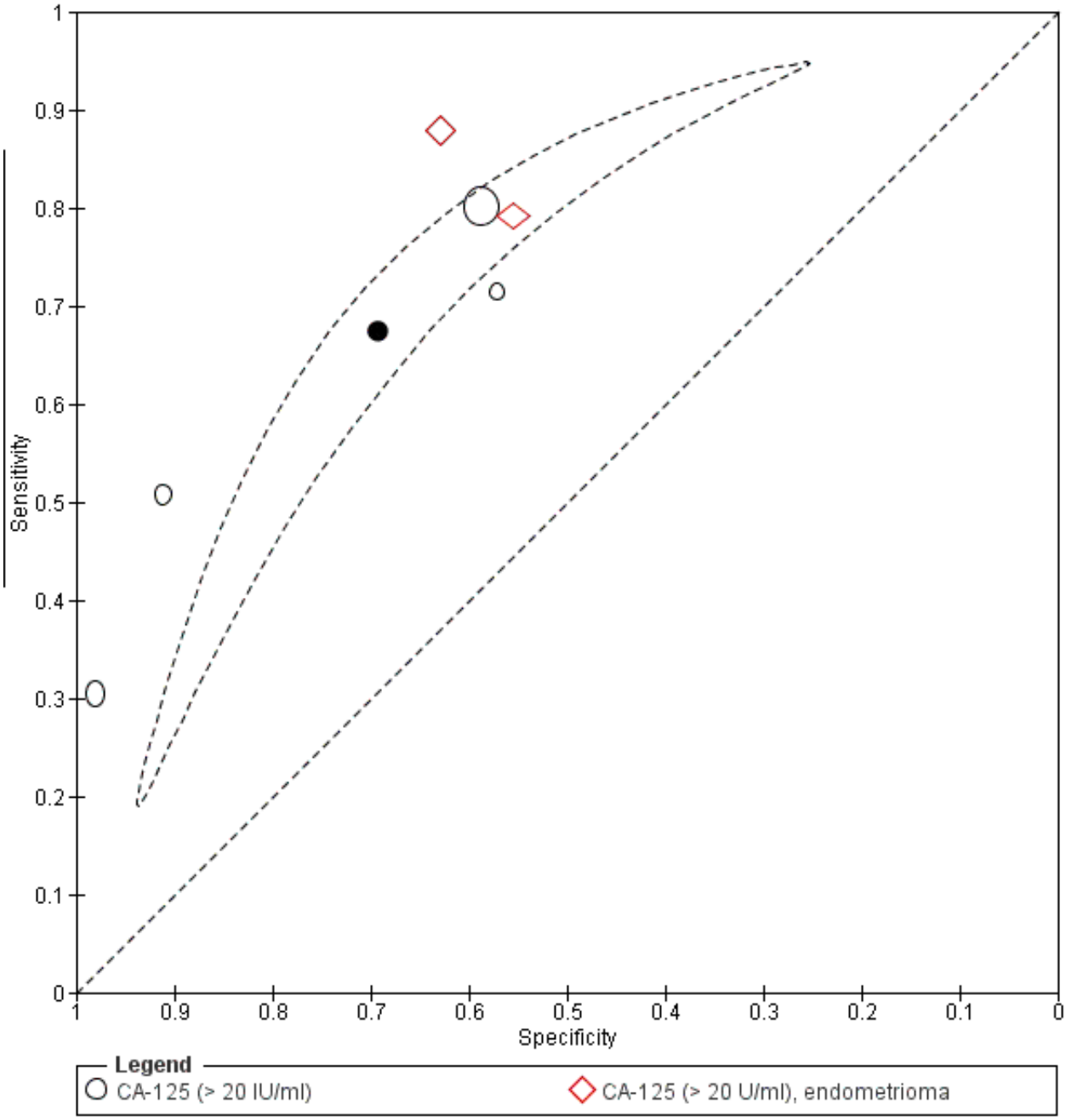


Figure 26. Summary ROC plot of CA-125 with cut-off values ranging > 25-26 U/ml for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The bars correspond to 95% CIs of each individual evaluation. The solid black circle represents the summary sensitivity and specificity.

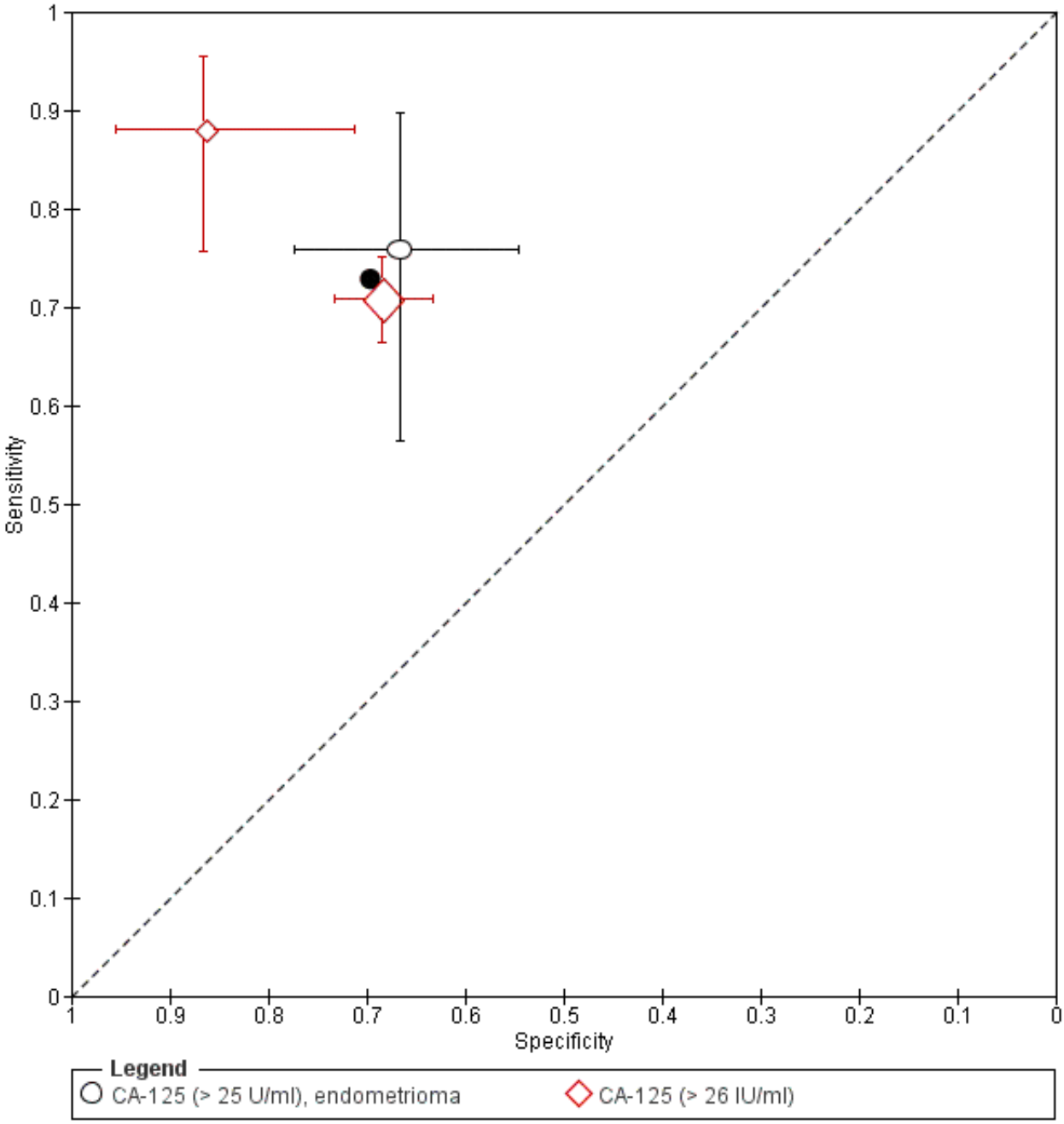


Figure 27. Summary ROC plot of CA-125 with cut-off values ranging > 30-33 U/ml for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The solid black circle represents the mean sensitivity and specificity, which is surrounded by a 95% confidence region (dotted line) and by 95% prediction region (dashed line).

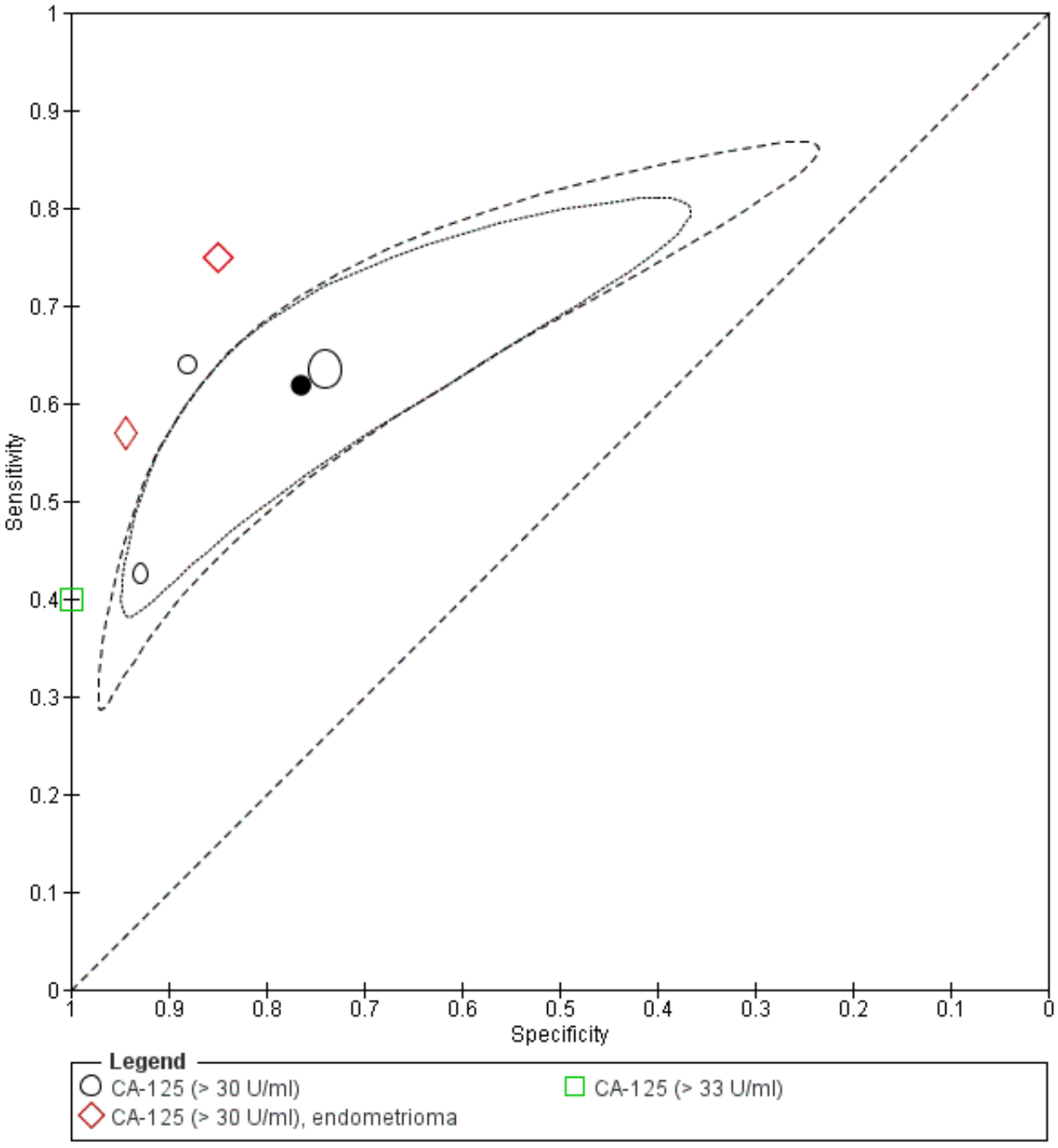
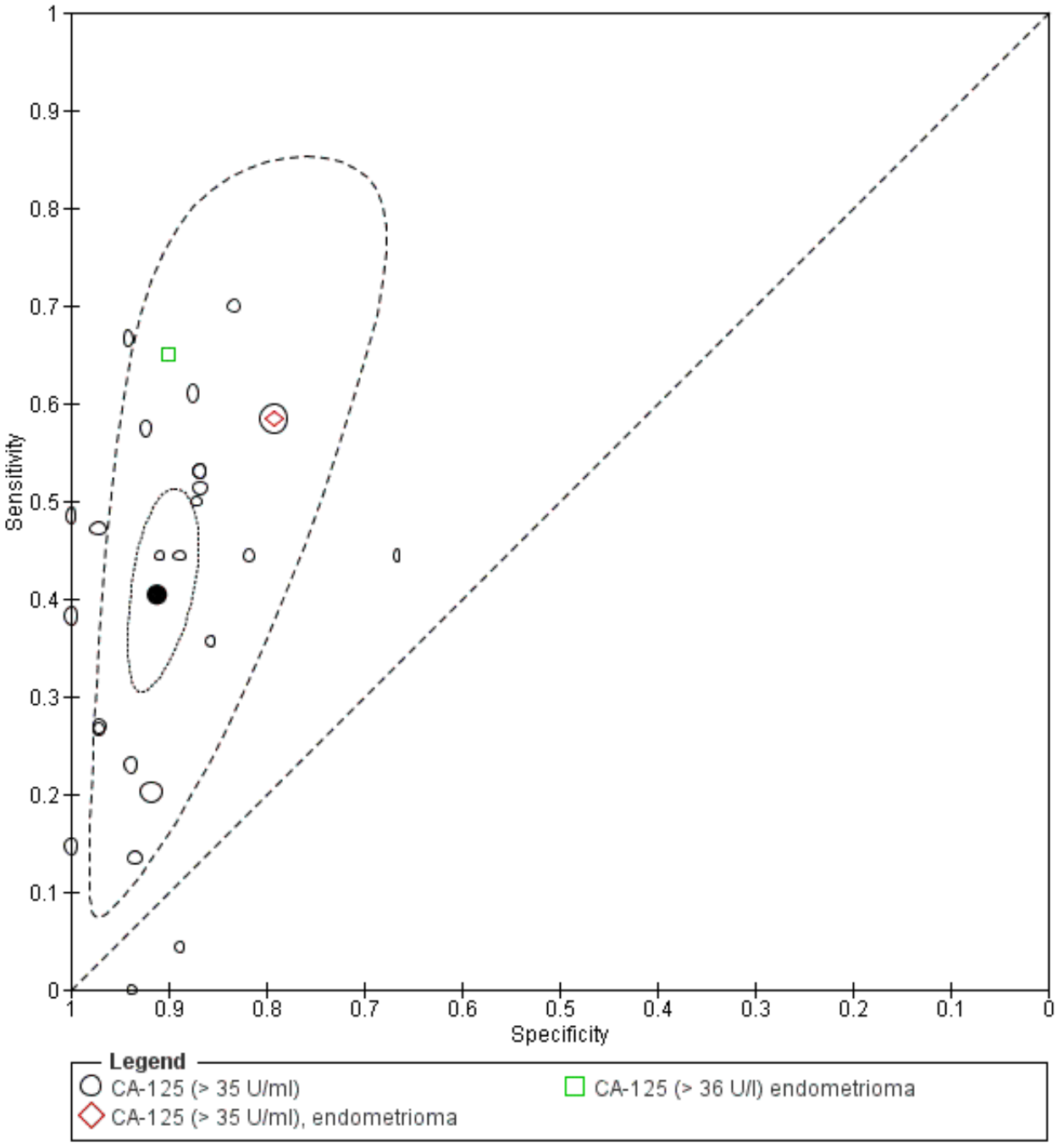


Figure 28. Summary ROC plot of CA-125 with cut-off values ranging > 35-36 U/ml for detection of endometriosis. Each point represents the pair of sensitivity and specificity from each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The solid black circle represents the mean sensitivity and specificity, which is surrounded by a 95% confidence region (dotted line) and by 95% prediction region (dashed line).



Overall, none of the cut-off thresholds for CA-125 subjected to a meta-analysis met the criteria for either a replacement or triage test. Only CA-125, with a cut-off value > 16.0 to 17.6 U/ml approached the criteria for a SpIN triage test, but results showed a substantial degree of heterogeneity and wide confidence intervals. Even though the reported diagnostic estimates for CA-125 with a cut-off of > 43.0 U/ml met the criteria for a replacement test, this cut-off value came from an individual study and only for moderate-severe forms of endometriosis. This is consistent with the commonly reported observation that CA-125 levels were

significantly increased in advanced stages of endometriosis and minimally altered in minimal-mild disease. Further large, well-designed diagnostic studies are required to evaluate the role of CA-125 with a cut-off > 43.0 U/ml in a population with a wide spectrum of endometriosis.

Two further studies (112 participants, follicular or follicular and luteal cycle phase) showed no association between CA-125 and endometriosis when assessing the full spectrum of the disease (rASRM I to IV) (Riley 2007) or only minimal-mild endometriosis (Barbosa 2009), as presented in Appendix 7.

A meta-analysis was undertaken for each specific cut-off value of CA-125 and included the studies that assessed its ability to detect pelvic endometriosis as well as the studies that aimed to determine if an ovarian mass was an endometrioma. The estimates from the studies that specifically evaluated ovarian endometrioma are also reported separately under 'Blood biomarkers that could differentiate ovarian endometrioma from other benign ovarian cysts in women of reproductive age' with the aim of evaluating the role of the test in differential diagnosis of ovarian masses in reproductive-aged women.

Direct comparisons for CA-125

Eight studies presented direct head-to-head comparisons between different cut-off thresholds for CA-125.

1. > 20.0 U/ml versus > 35.0 U/ml (Bilibio 2014).
2. > 16.0 U/ml versus >35.0 U/ml (Ferreira 1994).
3. > 30.0 U/ml versus > 36.0 U/ml (Florio 2007).
4. > 12.8 U/ml versus > 35.0 U/ml (Gagne 2003a).
5. > 20.0 U/ml versus > 25.0 U/ml versus > 35.0 U/ml (Guerriero 1996b).
6. > 20.0 U/ml versus > 26.0 U/ml versus > 30.0 U/ml versus > 35.0 U/ml (Kitawaki 2005).
7. > 10.0 U/ml versus > 20.0 U/ml (Rosa E Silva 2007).
8. > 20.0 U/ml versus >35.0 U/ml (Yang 1994).

Neither threshold appeared to be superior in most studies, and even when the diagnostic performance was improved when a different threshold was utilised, none of the threshold levels met the criteria for an adequate replacement or triage diagnostic test for endometriosis (Figure 29). Two studies performed head-to-head comparisons between different phases of menstrual cycle: > 11.5 U/ml follicular versus > 13.5 U/ml luteal versus > 12.5 U/ml all cycle phases (Vodolazkaia 2012); and menstrual versus follicular versus luteal, no cut-off reported (Mihalyi 2010). The test performance appeared to be improved in the follicular phase in one study (Vodolazkaia 2012) and in the menstrual or follicular phases in another study (Mihalyi 2010); however, the estimates were still lower than the criteria for an adequate replacement or triage test (Figure 30). Twenty-one studies directly compared the diagnostic performance of CA-125 with other blood biomarkers (Bilibio 2014; Dayangan Sayan 2013; Florio 2007; Florio 2009; Foda 2012; Harada 2002; Kurdoglu 2009; Mabrouk 2012; Martinez 2007; Mihalyi 2010; Mohamed 2013; Molo 1994; Muscatello 1992; Ohata 2008; Ozhan 2014; Somigliana 2004; Tokmak 2011; Tuten 2014a; Vodolazkaia 2012; Wild 1991a; Yang 1994). In view of the unsatisfactory diagnostic performance of CA-125 as a diagnostic or triage test, we do not discuss these comparisons in detail.

Figure 29. Forest plot of direct comparisons of CA-125 for detection of endometriosis performed between different cut-off values in 8 separate studies. Plot shows the estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which

the test was performed and severity of the disease assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

CA-125 (> 20 U/ml), Bilibio 2014

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Bilibio 2014	32	3	31	31	luteal	South America	I-IV	0.51 [0.38, 0.64]	0.91 [0.76, 0.98]

CA-125 (> 35 U/ml), Bilibio 2014

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Bilibio 2014	17	1	46	33	luteal	South America	I-IV	0.27 [0.17, 0.40]	0.97 [0.85, 1.00]

CA-125 (> 16 U/ml), Ferreira 1994

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Ferreira 1994	8	5	15	13	not specified	South America	I-IV	0.35 [0.16, 0.57]	0.72 [0.47, 0.90]

CA-125 (> 35 U/ml), Ferreira 1994

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Ferreira 1994	1	2	22	16	not specified	South America	I-IV	0.04 [0.00, 0.22]	0.89 [0.65, 0.99]

CA-125 (> 30 U/ml), Florio 2007

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2007	30	6	10	34	not specified	Europe	endometrioma	0.75 [0.59, 0.87]	0.85 [0.70, 0.94]

CA-125 (> 36 U/ml), Florio 2007

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2007	26	4	14	36	not specified	Europe	endometrioma	0.65 [0.48, 0.79]	0.90 [0.76, 0.97]

CA-125 (> 12.8 U/ml), Gagne 2003a

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Gagne 2003a	133	127	40	68	luteal	North America	I-IV	0.77 [0.70, 0.83]	0.35 [0.28, 0.42]

CA-125 (> 35 U/ml), Gagne 2003a

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Gagne 2003a	35	16	138	179	luteal	North America	I-IV	0.20 [0.15, 0.27]	0.92 [0.87, 0.95]

CA-125 (> 20 U/ml), Guerriero 1996b

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996b	23	32	6	40	follicular	Europe	endometrioma	0.79 [0.60, 0.92]	0.56 [0.43, 0.67]

CA-125 (≥ 25 U/ml), Guerriero 1996b

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996b	22	24	7	48	follicular	Europe	endometrioma	0.76 [0.56, 0.90]	0.67 [0.55, 0.77]

CA-125 (> 35 U/ml), Guerriero 1996b

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996b	17	15	12	57	follicular	Europe	endometrioma	0.59 [0.39, 0.76]	0.79 [0.68, 0.88]

CA-125 (> 20 U/ml), Kitawaki 2005

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Kitawaki 2005	347	141	86	201	follicular or luteal	Asia	I-IV	0.80 [0.76, 0.84]	0.59 [0.53, 0.64]

CA-125 (> 26 U/ml), Kitawaki 2005

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Kitawaki 2005	307	108	126	234	follicular or luteal	Asia	I-IV	0.71 [0.66, 0.75]	0.68 [0.63, 0.73]

CA-125 (> 30 U/ml), Kitawaki 2005

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Kitawaki 2005	275	89	158	253	follicular or luteal	Asia	I-IV	0.64 [0.59, 0.68]	0.74 [0.69, 0.79]

CA-125 (> 35 U/ml), Kitawaki 2005

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Kitawaki 2005	253	71	180	271	follicular or luteal	Asia	I-IV	0.58 [0.54, 0.63]	0.79 [0.75, 0.83]

Figure 29. (Continued)

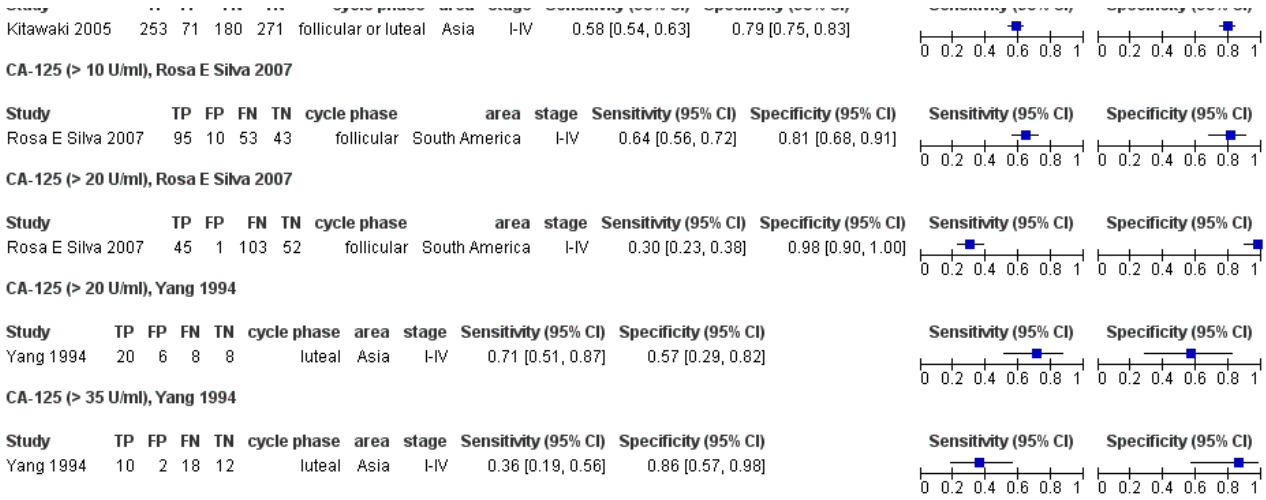
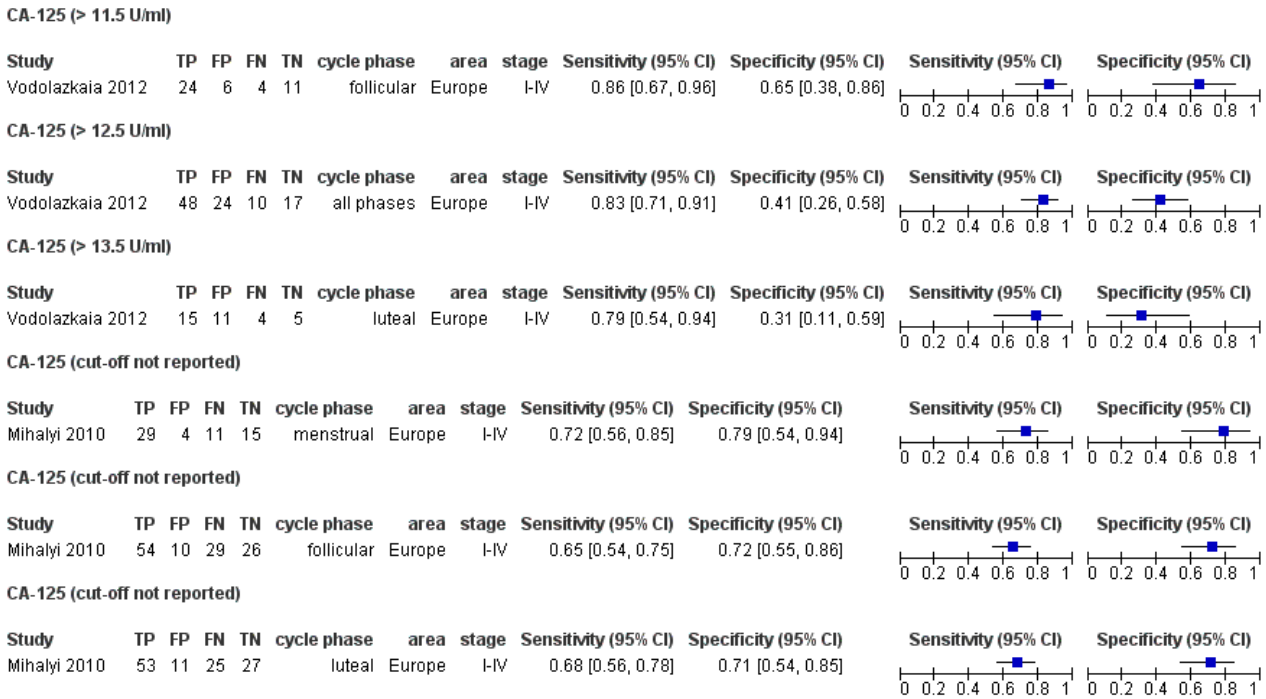


Figure 30. Forest plot of direct comparisons of CA-125 for detection of endometriosis performed between different phases of menstrual cycle in 2 separate studies. Plot shows the estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.



13.5. Tumour markers that exhibited no differential expression in endometriosis

There was no significant difference in the serum levels for several other tumour markers in women with and without endometriosis (Appendix 7), including AFP, c-erbB-2 (HER-2/neu) (Philippoussis 2004; 72 participants) and HE4 (Hallamaa 2012; 175 participants). Additional studies need to confirm these data.

14. Combined blood tests

There were 28 combined tests, comprised of two to six blood biomarkers that were evaluated as diagnostic tests for endometriosis and two other tests that attempted to discriminate ovarian endometriosis from other benign masses. We present the data for all the evaluated combined biomarkers, including the cut-off values and the analytical methods, in Summary of findings 1 and Figure 31. Twenty-three tests combined CA-125 with other blood biomarkers (Figure 32). Each set of biomarkers was tested in individual clinical trials that varied with respect to the selection of the biomarkers constituting the test and the population studied. The most promising results were reported for eight combined tests (Figure 33).

Figure 31. Forest plot of the combined tests (all the included evaluations) for detection of endometriosis, which consist of the combinations of 2-6 blood biomarkers. Plot shows the estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported as rASRM

stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Combined test (CA-125 ≥ 25 U/ml +/or CA-19.9 ≥ 12 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996a	35	47	4	32	follicular	Europe	endometrioma	0.90 [0.76, 0.97]	0.41 [0.30, 0.52]

Combined test (CA-125 ≥ 25 U/ml + Ca-19.9 ≥ 12 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996a	21	8	18	71	follicular	Europe	endometrioma	0.54 [0.37, 0.70]	0.90 [0.81, 0.96]

Combined test (CA-125 > 19.8 U/l + Prolactin > 14.8 ng/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Bilibio 2014	49	4	14	30	luteal	South America	I-IV	0.78 [0.66, 0.87]	0.88 [0.73, 0.97]

Combined test (CA-125 > 35 U/l + Prolactin > 20 ng/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Bilibio 2014	28	1	35	33	luteal	South America	I-IV	0.44 [0.32, 0.58]	0.97 [0.85, 1.00]

Combined test (CA-125 > 17.6 IU/ml + VEGF > 236 pg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Foda 2012	50	2	15	28	follicular	Middle East	I-IV	0.77 [0.65, 0.86]	0.93 [0.78, 0.99]

Combined test (CA-125 > 20 U/l + Anti-endometrial Abs > 0.3 A-value)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Yang 1994	17	3	11	11	luteal	Asia	I-IV	0.61 [0.41, 0.78]	0.79 [0.49, 0.95]

Combined test (CA-125 x NLR; (> 43.1)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Dayangan Sayan 2013	40	7	10	43	menstrual	Europe	I-IV	0.80 [0.66, 0.90]	0.86 [0.73, 0.94]

Combined test (CA-125 > 30 U/ml +/or IL-8 ≥ 25 pg/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Ohata 2008	56	5	9	13	follicular or luteal	Asia	endometrioma	0.86 [0.75, 0.93]	0.72 [0.47, 0.90]

Combined test (CA-125 + IL-8) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	143	27	58	66	all phases	Europe	I-IV	0.71 [0.64, 0.77]	0.71 [0.61, 0.80]

Combined test (IL-6 > 12.2 pg/ml + TNF-α > 12.45 pg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Foda 2012	46	0	20	30	follicular	Middle East	I-IV	0.70 [0.57, 0.80]	1.00 [0.88, 1.00]

Combined test (IL-6 > 12.2 pg/ml + CRP > 438 µg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Foda 2012	49	0	16	30	follicular	Middle East	I-IV	0.75 [0.63, 0.85]	1.00 [0.88, 1.00]

Combined test (TNF-α > 12.45 pg/ml + CRP > 438 µg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Foda 2012	48	0	17	30	follicular	Middle East	I-IV	0.74 [0.61, 0.84]	1.00 [0.88, 1.00]

Combined test (miR-199a + miR-542-3p) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Wang 2013a	58	3	2	22	follicular or luteal	Asia	I-IV	0.97 [0.88, 1.00]	0.88 [0.69, 0.97]

Combined test (miR-199a + miR-122) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Wang 2013a	48	5	12	20	follicular or luteal	Asia	I-IV	0.80 [0.68, 0.89]	0.80 [0.59, 0.93]

Combined test (Ca-125 + Ca 19-9 + Survivin) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mabrouk 2012	35	2	5	18	follicular	Europe	n/a	0.88 [0.73, 0.96]	0.90 [0.68, 0.99]

Figure 31. (Continued)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mabrouk 2012	35	2	5	18	follicular	Europe	n/a	0.88 [0.73, 0.96]	0.90 [0.68, 0.99]

Combined test (CA-125 + STX-5 + LN-1) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Ozhan 2014	57	6	3	14	not specified	Europe	I-IV	0.95 [0.86, 0.99]	0.70 [0.46, 0.88]

Combined test (CA-125 > 35 IU/ml +/or CA-19.9 > 37 IU/ml +/or IL-6 > 2 pg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Somigliana 2004	19	10	26	25	all phases	Europe	I-IV	0.42 [0.28, 0.58]	0.71 [0.54, 0.85]

Combined test (CA-125 > 50 IU/mL +/ or CCR1 > 1.16 +/or MCP-1 > 140 pg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Agic 2008	94	9	8	40	follicular	Europe	I-IV	0.92 [0.85, 0.97]	0.82 [0.68, 0.91]

Combined test (Ca-125 > 20 mIU/ml + MCP-1 > 152.74 pg/ml + Leptin > 3.14 ng/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Seeber 2008	31	5	32	73	all phases	North America	II-IV	0.49 [0.36, 0.62]	0.94 [0.86, 0.98]

Combined test CA-125 + IL-8 + TNF-α) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	70	11	8	27	luteal	Europe	I-IV	0.90 [0.81, 0.95]	0.71 [0.54, 0.85]

Combined test (IL-6 > 12.2 pg/ml + TNF-α > 12.45 pg/ml + CRP > 438 µg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Foda 2012	41	0	24	30	follicular	Middle East	I-IV	0.63 [0.50, 0.75]	1.00 [0.88, 1.00]

Combined test (CA-125 + VEGF + annexin V + glycodelin] - MLR (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Vodolazkaia 2012	9	2	2	6	menstrual	Europe	I-IV	0.82 [0.48, 0.98]	0.75 [0.35, 0.97]

Combined test (CA-125 + VEGF + annexin V + glycodelin] - LS-SVM (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Vodolazkaia 2012	9	3	2	5	menstrual	Europe	I-IV	0.82 [0.48, 0.98]	0.63 [0.24, 0.91]

Combined test (CA-125 + VEGF + annexin V + sICAM-1) - MLR or LS-SVM (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Vodolazkaia 2012	9	2	2	6	menstrual	Europe	I-IV	0.82 [0.48, 0.98]	0.75 [0.35, 0.97]

Combined test (CA-125 > 20 mIU/ml + MCP-1 > 53.5 pg/ml + Leptin > 29.1 ng/ml + MIF > 14.7 ng/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Seeber 2008	63	51	0	27	all phases	North America	II-IV	1.00 [0.94, 1.00]	0.35 [0.24, 0.46]

Combined test (miR-199a + miR-122 + miR-145* + miR-542-3p) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Wang 2013a	56	1	4	24	follicular or luteal	Asia	I-IV	0.93 [0.84, 0.98]	0.96 [0.80, 1.00]

Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	181	44	20	49	all phases	Europe	I-IV	0.90 [0.85, 0.94]	0.53 [0.42, 0.63]

Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	36	5	4	14	menstrual	Europe	I-IV	0.90 [0.76, 0.97]	0.74 [0.49, 0.91]

Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	48	10	35	26	follicular	Europe	I-IV	0.58 [0.46, 0.69]	0.72 [0.55, 0.86]

Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)

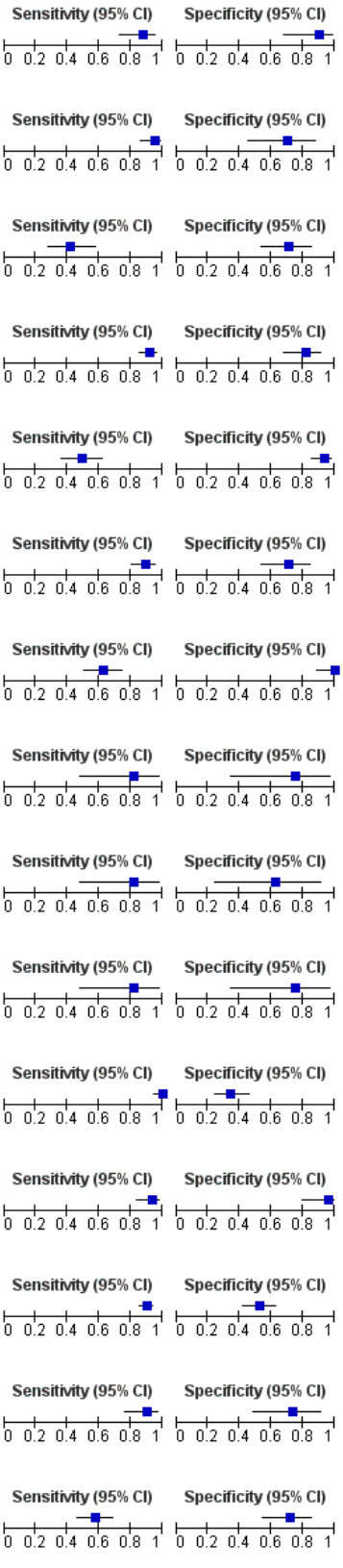


Figure 31. (Continued)

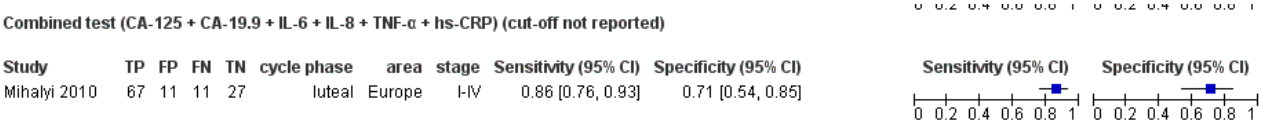


Figure 32. Forest plot of the combined tests for detection of endometriosis, which consist of the combinations of CA-125 with other blood biomarkers. Plot shows the estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which

the test was performed and severity of the disease assessed by each study, reported as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Combined test (CA-125 ≥ 25 U/ml +/or CA-19.9 ≥ 12 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996a	35	47	4	32	follicular	Europe	endometrioma	0.90 [0.76, 0.97]	0.41 [0.30, 0.52]

Combined test (CA-125 ≥ 25 U/ml + Ca-19.9 ≥ 12 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996a	21	8	18	71	follicular	Europe	endometrioma	0.54 [0.37, 0.70]	0.90 [0.81, 0.96]

Combined test (CA-125 > 19.8 U/l + Prolactin > 14.8 ng/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Bilibio 2014	49	4	14	30	luteal	South America	I-IV	0.78 [0.66, 0.87]	0.88 [0.73, 0.97]

Combined test (CA-125 > 35 U/l + Prolactin > 20 ng/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Bilibio 2014	28	1	35	33	luteal	South America	I-IV	0.44 [0.32, 0.58]	0.97 [0.85, 1.00]

Combined test (CA-125 > 17.6 IU/ml + VEGF > 236 pg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Foda 2012	50	2	15	28	follicular	Middle East	I-IV	0.77 [0.65, 0.86]	0.93 [0.78, 0.99]

Combined test (CA-125 > 20 U/l + Anti-endometrial Abs > 0.3 A-value)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Yang 1994	17	3	11	11	luteal	Asia	I-IV	0.61 [0.41, 0.78]	0.79 [0.49, 0.95]

Combined test (CA-125 x NLR; (> 43.1)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Dayangan Sayan 2013	40	7	10	43	menstrual	Europe	I-IV	0.80 [0.66, 0.90]	0.86 [0.73, 0.94]

Combined test (CA-125 > 30 U/ml +/or IL-8 ≥ 25 pg/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Ohata 2008	56	5	9	13	follicular or luteal	Asia	endometrioma	0.86 [0.75, 0.93]	0.72 [0.47, 0.90]

Combined test (CA-125 + IL-8) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	143	27	58	66	all phases	Europe	I-IV	0.71 [0.64, 0.77]	0.71 [0.61, 0.80]

Combined test (Ca-125 + Ca 19-9 + Survivin) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mabrouk 2012	35	2	5	18	follicular	Europe	n/a	0.88 [0.73, 0.96]	0.90 [0.68, 0.99]

Combined test (CA-125 + STX-5 + LN-1) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Ozhan 2014	57	6	3	14	not specified	Europe	I-IV	0.95 [0.86, 0.99]	0.70 [0.46, 0.88]

Combined test (CA-125 > 35 IU/ml +/or CA-19.9 > 37 IU/ml +/or IL-6 > 2 pg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Somigliana 2004	19	10	26	25	all phases	Europe	I-IV	0.42 [0.28, 0.58]	0.71 [0.54, 0.85]

Combined test (CA-125 > 50 IU/mL +/ or CCR1 > 1.16 +/or MCP-1 > 140 pg/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Agic 2008	94	9	8	40	follicular	Europe	I-IV	0.92 [0.85, 0.97]	0.82 [0.68, 0.91]

Combined test (Ca-125 > 20 mIU/ml + MCP-1 > 152.74 pg/ml + Leptin > 3.14 ng/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Seeber 2008	31	5	32	73	all phases	North America	II-IV	0.49 [0.36, 0.62]	0.94 [0.86, 0.98]

Combined test CA-125 + IL-8 + TNF-α) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	70	11	8	27	luteal	Europe	I-IV	0.90 [0.81, 0.95]	0.71 [0.54, 0.85]

Figure 32. (Continued)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	70	11	8	27	luteal	Europe	I-IV	0.90 [0.81, 0.95]	0.71 [0.54, 0.85]

Combined test (CA-125 + VEGF + annexin V + glycodelin] - MLR (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Vodolazkaia 2012	9	2	2	6	menstrual	Europe	I-IV	0.82 [0.48, 0.98]	0.75 [0.35, 0.97]

Combined test (CA-125 + VEGF + annexin V + glycodelin] - LS-SVM (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Vodolazkaia 2012	9	3	2	5	menstrual	Europe	I-IV	0.82 [0.48, 0.98]	0.63 [0.24, 0.91]

Combined test (CA-125 + VEGF + annexin V + sICAM-1) - MLR or LS-SVM (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Vodolazkaia 2012	9	2	2	6	menstrual	Europe	I-IV	0.82 [0.48, 0.98]	0.75 [0.35, 0.97]

Combined test (CA-125 > 20 mIU/ml + MCP-1 > 53.5 pg/ml + Leptin > 29.1 ng/ml + MIF > 14.7 ng/ml)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Seeber 2008	63	51	0	27	all phases	North America	II-IV	1.00 [0.94, 1.00]	0.35 [0.24, 0.46]

Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	181	44	20	49	all phases	Europe	I-IV	0.90 [0.85, 0.94]	0.53 [0.42, 0.63]

Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	36	5	4	14	menstrual	Europe	I-IV	0.90 [0.76, 0.97]	0.74 [0.49, 0.91]

Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	48	10	35	26	follicular	Europe	I-IV	0.58 [0.46, 0.69]	0.72 [0.55, 0.86]

Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Mihalyi 2010	67	11	11	27	luteal	Europe	I-IV	0.86 [0.76, 0.93]	0.71 [0.54, 0.85]

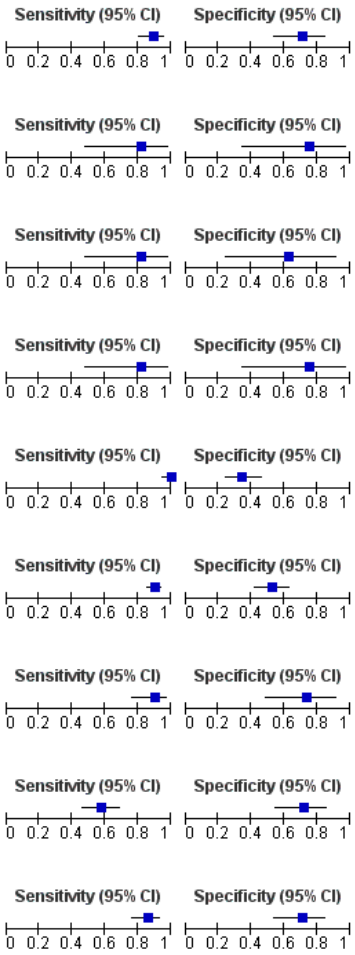
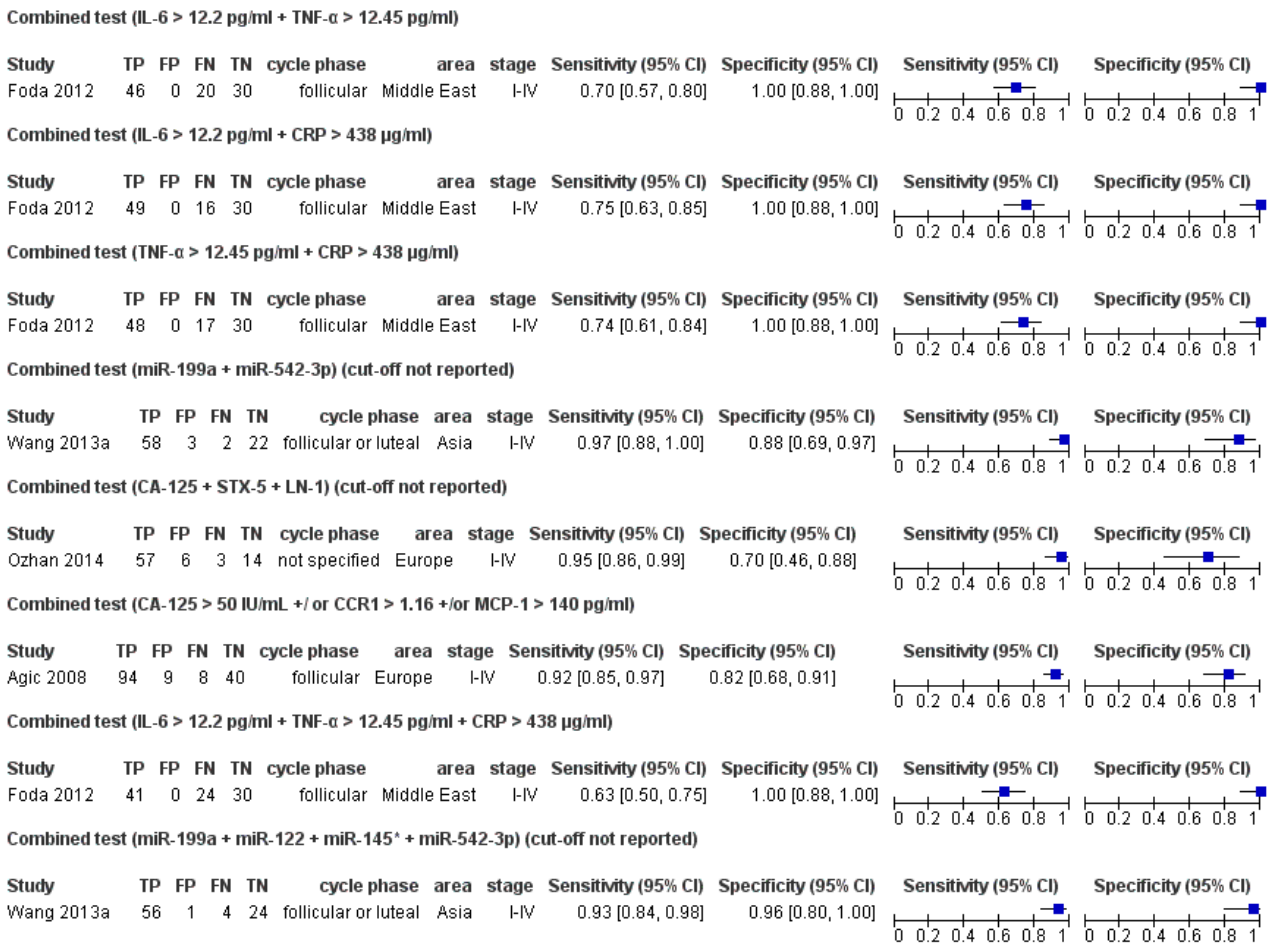


Figure 33. Forest plot of the most promising combined tests of blood biomarkers for detection of endometriosis. Plot shows the estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported as rASRM stage. FN: false negative; FP: false positive; TN: true negative; TP: true positive.



1. miR-199a + miR-542-3p (85 participants, follicular or luteal cycle phase, rASRM I to IV) with a sensitivity of 0.97 (95% CI 0.88 to 1.00) and a specificity of 0.88 (95% CI 0.69 to 0.97), demonstrating estimates that reached the criteria for either replacement or SnOUT triage test (Wang 2013a).
2. CA-125 +/CCR1 +/MCP-1 (151 participants, follicular cycle phase; rASRM I to IV) with a sensitivity of 0.92 (95% CI 0.85 to 0.97) and a specificity of 0.82 (95% CI 0.68 to 0.91), demonstrating diagnostic estimates that approached that of a replacement or SnOUT triage test (Agic 2008).
3. miR-199a + miR-122 + miR-145* + miR-542-3p (85 participants, follicular or luteal cycle phase; rASRM I to IV) with a sensitivity of 0.93 (95% CI 0.84 to 0.98) and a specificity of 0.96 (95% CI 0.80 to 1.00), demonstrating diagnostic estimates that approached that of a replacement or SnOUT triage test (Wang 2013a).
4. CA-125 + STX-5 + LN-1, cut-off not reported (80 participants, cycle phase not reported, rASRM I to IV) with a sensitivity of 0.95 (95% CI 0.86 to 0.99) and a specificity of 0.70 (95% CI 0.46 to 0.88), meeting the criteria for SnOUT triage test (Ozhan 2014).
5. IL-6 > 12.20 pg/ml + TNF-α >12.45 pg/ml (96 participants, follicular cycle phase, rASRM I to IV) with a sensitivity of 0.70 (95% CI 0.57 to 0.80) and a specificity of 1.00 (95% CI 0.88 to 1.00), meeting the criteria for SpIN triage test (Foda 2012).
6. IL-6 > 12.20 pg/ml + CRP > 438 µg/ml (95 participants, follicular cycle phase, rASRM I to IV) with a sensitivity of 0.75 (95% CI 0.63 to 0.85) and a specificity of 1.00 (95% CI 0.88 to 1.00), meeting the criteria for SpIN triage test (Foda 2012).
7. TNF-α > 12.45 pg/ml + CRP > 438 µg/ml (95 participants, follicular cycle phase, rASRM I to IV) with a sensitivity of 0.74 (95% CI 0.61 to 0.84) and a specificity of 1.00 (95% CI 0.88 to 1.00), meeting the criteria for SpIN triage test (Foda 2012).
8. IL-6 > 12.20 pg/ml + TNF-α > 12.45 pg/ml + CRP > 438 µg/ml (95 participants, follicular cycle phase, rASRM I to IV) with a sensitivity of 0.63 (95% CI 0.50 to 0.75) and a specificity of 1.00 (95% CI 0.88 to 1.00), meeting the criteria for SpIN triage test (Foda 2012).

With a few exceptions, the panels of multiple biomarkers did not appear to be superior to some single biomarker tests. These findings need to be confirmed in large, well-designed diagnostic studies and independent test populations.

Blood biomarkers that could differentiate ovarian endometrioma from other benign ovarian cysts in women of reproductive age

Seven studies evaluated blood biomarkers for their potential to distinguish ovarian endometrioma from other benign ovarian masses, with six formally evaluating diagnostic test performance and one study presenting negative findings. We summarise the evaluated biomarkers in a forest plot (Figure 34) and describe them here.

Figure 34. Forest plot of the tests for detection of ovarian endometriosis performed through comparisons in women with endometriosis versus other benign ovarian cysts in 6 studies. Plot shows the estimates of sensitivity and specificity (squares) with 95% CI (black line) specific for each evaluation, country in which the study was conducted, menstrual cycle phase at which the test was performed and severity of the disease assessed by each study, reported

as rASRM stage. The studies are ordered according to the study names. FN: false negative; FP: false positive; TN: true negative; TP: true positive.

Urocortin (> 29 pg/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2007	39	6	1	34	not specified	Europe	endometrioma	0.97 [0.87, 1.00]	0.85 [0.70, 0.94]

Urocortin (> 33 pg/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2007	35	4	5	36	not specified	Europe	endometrioma	0.88 [0.73, 0.96]	0.90 [0.76, 0.97]

Urocortin (> 41.6 pg/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Tokmak 2011	32	25	10	21	follicular	Europe	endometrioma	0.76 [0.61, 0.88]	0.46 [0.31, 0.61]

IL-8 (≥ 25 pg/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Ohata 2008	50	4	20	17	follicular or luteal	Asia	endometrioma	0.71 [0.59, 0.82]	0.81 [0.58, 0.95]

Follistatin (> 1433 pg/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2009	48	4	4	48	follicular	Europe	endometrioma	0.92 [0.81, 0.98]	0.92 [0.81, 0.98]

CA-19.9 (≥ 12 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996a	24	24	15	55	follicular	Europe	endometrioma	0.62 [0.45, 0.77]	0.70 [0.58, 0.79]

CA-125 (> 20 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996b	23	32	6	40	follicular	Europe	endometrioma	0.79 [0.60, 0.92]	0.56 [0.43, 0.67]
Tokmak 2011	37	17	5	29	follicular	Europe	endometrioma	0.88 [0.74, 0.96]	0.63 [0.48, 0.77]

CA-125 (> 25 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996b	22	24	7	48	follicular	Europe	endometrioma	0.76 [0.56, 0.90]	0.67 [0.55, 0.77]

CA-125 (> 30 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2007	30	6	10	34	not specified	Europe	endometrioma	0.75 [0.59, 0.87]	0.85 [0.70, 0.94]
Ohata 2008	37	1	28	17	follicular or luteal	Asia	endometrioma	0.57 [0.44, 0.69]	0.94 [0.73, 1.00]

CA-125 (> 35 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996b	17	15	12	57	follicular	Europe	endometrioma	0.59 [0.39, 0.76]	0.79 [0.68, 0.88]

CA-125 (> 36 U/l) endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2007	26	4	14	36	not specified	Europe	endometrioma	0.65 [0.48, 0.79]	0.90 [0.76, 0.97]

CA-125 (> 42 U/l), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Florio 2009	23	5	29	47	follicular	Europe	endometrioma	0.44 [0.30, 0.59]	0.90 [0.79, 0.97]

Combined test (CA-125 ≥ 25 U/ml +/or CA-19.9 ≥ 12 U/ml), endometrioma

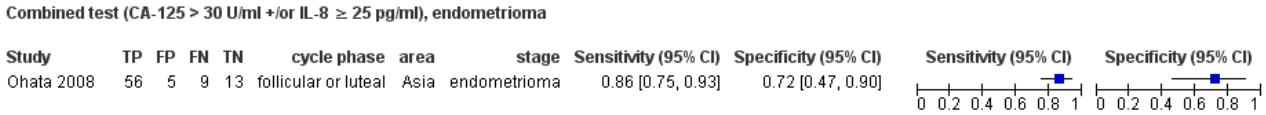
Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996a	35	47	4	32	follicular	Europe	endometrioma	0.90 [0.76, 0.97]	0.41 [0.30, 0.52]

Combined test (CA-125 ≥ 25 U/ml + Ca-19.9 ≥ 12 U/ml), endometrioma

Study	TP	FP	FN	TN	cycle phase	area	stage	Sensitivity (95% CI)	Specificity (95% CI)
Guerriero 1996a	21	8	18	71	follicular	Europe	endometrioma	0.54 [0.37, 0.70]	0.90 [0.81, 0.96]

Combined test (CA-125 > 30 U/ml +/or IL-8 ≥ 25 pg/ml), endometrioma

Figure 34. (Continued)

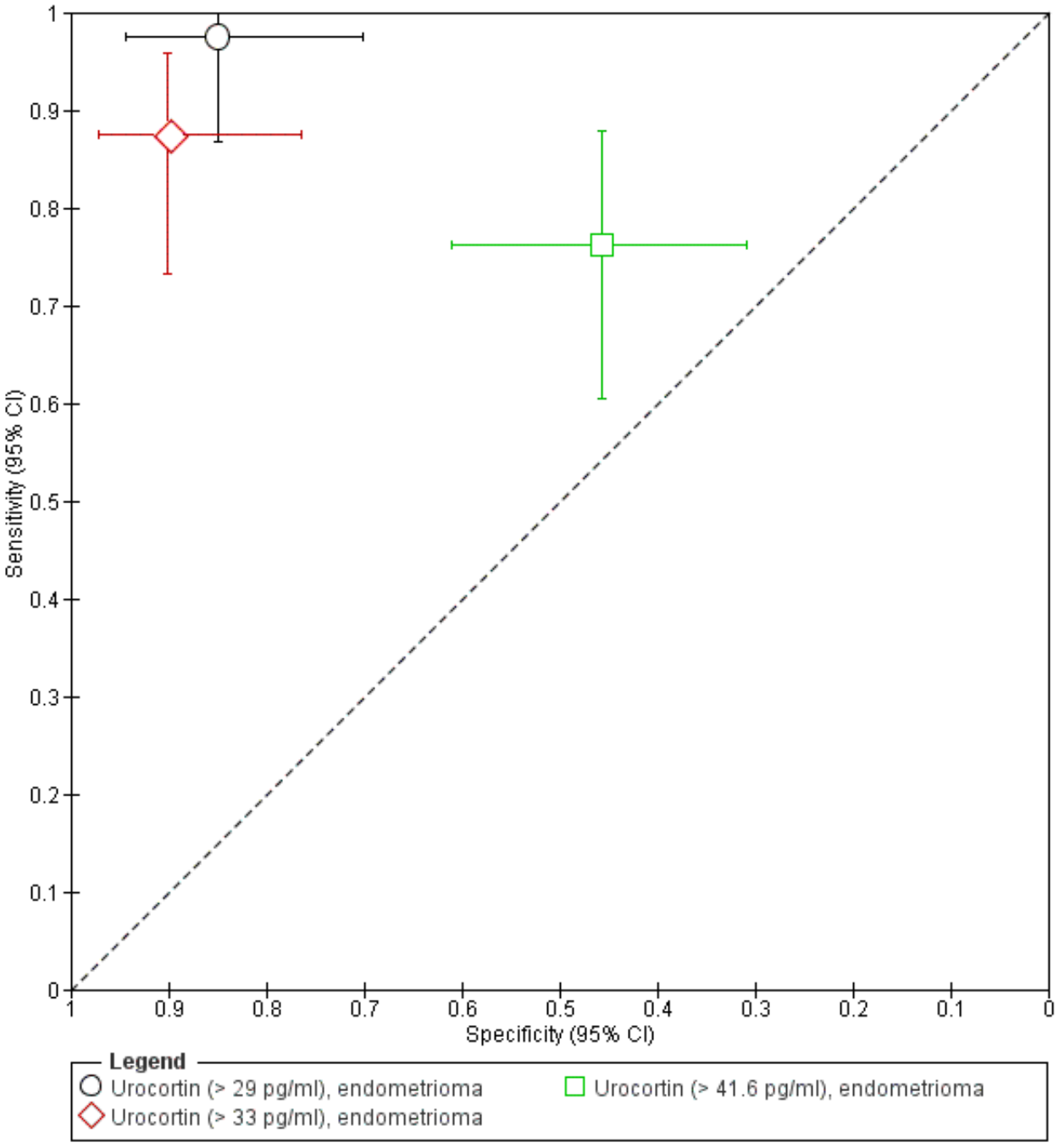


1. Angiogenesis/growth factors and their receptors

1.1. Urocortin

Two studies, including three data sets with a total of 168 participants, assessed the accuracy of urocortin in detecting ovarian endometriosis (Figure 35). One study evaluated two different cut-offs in the same population (80 participants, cycle phase not reported, rASRM III to IV; Florio 2007): urocortin with a cut-off of > 29.00 pg/ml had a sensitivity of 0.97 (95% CI 0.87 to 1.00) and a specificity of 0.85 (95% CI 0.70 to 0.94), meeting the criteria for a replacement test; and urocortin with a cut-off of > 33.00 pg/ml had a sensitivity of 0.88 (95% CI 0.73 to 0.96) and a specificity of 0.90 (95% CI 0.76 to 0.97), approaching the criteria for a SpIN triage test. Another study (88 participants, follicular cycle phase, rASRM III to IV; Tokmak 2011) demonstrated urocortin levels that were not statistically different in women with and without endometriosis. The test was still performed at a cut-off of > 41.60 pg/ml, demonstrating a sensitivity of 0.76 (95% CI 0.61 to 0.88) and a specificity of 0.46 (95% CI 0.31 to 0.61). We did not perform a meta-analysis in view of the heterogeneity of the cut-off thresholds between studies. Further evaluation of urocortin across the spectrum of endometriosis may help to clarify its diagnostic role in endometriosis.

Figure 35. Summary ROC plot of urocortin for detection of endometriosis. Each point represents the pair of sensitivity and specificity for each evaluation. The size of each point is proportional to the sample size and the shape designates the tests with different cut-off values. The bars correspond to 95% CIs of each individual evaluation. Two evaluations (> 29 pg/ml and > 33 pg/ml) were performed in the same population. The data were not assessed by meta-analysis.



2. Immune system and inflammatory markers

2.1. IL-8 (interleukin-8)

One study assessed the performance of IL-8 in detecting ovarian endometriosis (91 participants, follicular or luteal cycle phase, cut-off value >25.00 pg/ml; Ohata 2008), demonstrating a sensitivity of 0.71 (95% CI 0.59 to 0.82) and a specificity of 0.81 (95% CI 0.58 to 0.95) (Summary of findings 1; Figure 15). The diagnostic estimates were higher compared to those reported for overall pelvic endometriosis but remained far below the criteria for either

replacement or triage test, and there were insufficient data for meaningful comparisons.

2.2. Immune system and inflammatory markers that exhibited no differential expression in endometriosis

One study (95 participants, cycle phase not reported) demonstrated no significant difference in peripheral levels of IL-6 and sCD163 when women with ovarian endometrioma were compared to a group with other benign ovarian cysts (Jee 2008). This supports the negative findings reported for IL-6 in overall pelvic endometriosis (see above). The data for sCD163 is insufficient to comment on its diagnostic role.

3. Other peptides/proteins shown to influence key events implicated in endometriosis

3.1. Follistatin

One study evaluated follistatin (104 participants, follicular cycle phase, ovarian endometriosis, rASRM III to IV; Florio 2009) and showed a high sensitivity of 0.92 (95% CI 0.81 to 0.98) and high specificity of 0.92 (95% CI 0.81 to 0.98), using a cut-off value of > 1433.00 pg/ml. The diagnostic estimates approached the criteria for either a replacement or both SnOUT and SpIN triage test, but further validation in larger studies that evaluate a wider spectrum of disease is required.

4. Tumour markers

4.1. CA-19.9 (cancer antigen-19.9)

One study (118 participants, follicular cycle phase; Guerriero 1996a) evaluated the performance of CA-19.9 at a cut-off value > 12 U/ml in differentiating ovarian endometriosis from other benign ovarian cysts. The reported diagnostic estimates (sensitivity 0.62, 95% CI 0.45 to 0.77 and specificity 0.70, 95% CI 0.58 to 0.79) were below the diagnostic thresholds for either replacement or triage test, and this was similar to the findings reported for CA-19.9 in overall pelvic endometriosis (Figure 21).

4.2. CA-125 (cancer antigen-125)

Seven studies assessed CA-125 in differentiating ovarian endometriosis from other ovarian cysts, using several cut-off values.

1. CA-125 with a cut-off value of > 20 U/ml (Guerriero 1996b; Tokmak 2011; 189 women, follicular cycle phase) had sensitivities of 0.79 and 0.88 (95% CI 0.60 to 0.92 and 0.74 to 0.96) and specificities of 0.56 and 0.63 (95% CI 0.43 to 0.67 and 0.48 to 0.77).
2. CA-125 with a cut-off value of ≥ 25 U/ml (Guerriero 1996b; 101 women, follicular cycle phase) demonstrated a sensitivity of 0.76 (95% CI 0.56 to 0.90) and a specificity of 0.67 (95% CI 0.55 to 0.77).
3. CA-125 with a cut-off value of > 30 U/ml (Florio 2007; Ohata 2008; 171 women, various cycle phases) exhibited sensitivities of 0.75 and 0.57 (95% CI 0.59 to 0.87 and 0.44 to 0.69) and specificities of 0.85 and 0.94 (95% CI 0.70 to 0.94 and 0.73 to 1.00).
4. CA-125 with a cut-off value of > 35 U/ml (Guerriero 1996b; 101 women, follicular cycle phase) had a sensitivity of 0.59 (95% CI 0.39 to 0.76) and a specificity of 0.79 (95% CI 0.68 to 0.88).
5. CA-125 with a cut-off value of > 36 U/ml (Florio 2009; 80 women, follicular cycle phase) had a sensitivity of 0.65 (95% CI 0.98 to 0.79) and a specificity of 0.90 (95% CI 0.76 to 0.97).
6. CA-125 with a cut-off value of > 42 U/ml (Florio 2009, 104 women, follicular cycle phase) had a sensitivity of 0.44 (95% CI 0.30 to 0.59) and a specificity of 0.90 (95% CI 0.79 to 0.97). None of the tests met the criteria for a replacement or triage test and CA-125 with a cut-off value > 36 U/ml only approached the criteria of a SpIN triage test; however, there were insufficient data to perform meaningful analyses specific for ovarian endometrioma for any of the cut-offs.

5. Combined blood tests

Two combinations of biomarkers were specifically assessed for their ability to distinguish ovarian endometrioma from other ovarian cysts.

1. CA-125 + CA-19.9 with the cut-off values ≥ 25 U/ml and ≥ 12 U/ml, respectively (Guerriero 1996a; 118 women, follicular cycle phase) demonstrated a sensitivity of 0.90 (95% CI 0.76 to 0.97) with a specificity of 0.41 (95% CI 0.3 to 0.52) when either positive biomarker was considered and a sensitivity of 0.54 (95% CI 0.37 to 0.70) with a specificity of 0.90 (95% CI 0.81 to 0.96) when both positive biomarkers were included.
2. CA-125 + IL-8 with the cut-off values > 30 U/ml and ≥ 25 U/ml, respectively (Ohata 2008; 91 women, follicular or luteal cycle phase) had a sensitivity of 0.86 (95% CI 0.75 to 0.93) and a specificity of 0.72 (95% CI 0.47 to 0.90).

Small numbers of studies assessed each test evaluated for their ability to distinguish ovarian endometriosis from other benign masses, and we could not draw a firm conclusion. The available evidence is scant; however, several biomarkers showed some diagnostic potential as summarised in 'Summary of main results' under 'Tests to be validated for their diagnostic potential'. Further evaluations of these biomarkers are necessary to improve the certainty with regard to their diagnostic value in ovarian endometriosis.

Investigation of heterogeneity and sensitivity analyses

The potential sources of heterogeneity are outlined in Secondary objectives. Although we attempted to assess these sources, there were not enough studies evaluating each test to make this a meaningful analysis, except for the meta-analysis comprised of 27 studies for CA-125 with a cut-off value > 35 to 36 U/ml.

1. Two studies were published between 1986 and 1989; 13 studies, between 1990 and 1999; 9 studies, between 2000 and 2009; and 3 studies, between 2010 and 2014.
2. Fourteen studies took place in Europe, five studies in Asia, four studies in North America, three studies in South America, and one study in the Middle East.
3. Nineteen studies used a single-gate design, and eight studies had a two-gate design.
4. One study assessed only minimal-mild endometriosis (rASRM I to II); 3 studies, only moderate-severe disease (rASRM III to IV); one study did not provide information on the severity, and 22 studies evaluated a wide spectrum of endometriosis (rASRM I to IV). Of these 22 studies, 11 presented separate diagnostic estimates for different rASRM stages in addition to the data for the entire group, but we did not include this information in the review and did not consider it in the assessment of heterogeneity.

5. Thirteen studies used histopathology in adjunct to laparoscopy as a reference standard, while 14 studies relied on visual inspection of pelvic cavity.
6. Two studies were specific for the diagnosis of ovarian endometrioma, one study assessed only peritoneal endometriosis, and the remaining 24 studies evaluated overall pelvic endometriosis.
7. Nine studies evaluated the diagnostic performance of CA-125 in the follicular cycle phase; six studies, in the luteal phase; two studies, in the follicular or luteal phase; and three studies, in all cycle phases. Seven studies did not report the cycle phase.
8. Nineteen studies included various clinical presentations (pain ± infertility ± ovarian mass), of which one study reported separate estimates for populations with infertility and with pelvic pain, one study included only participants with pelvic pain, three studies were confined to women presenting with infertility, three studies evaluated only women with ovarian mass, and one study did not specify clinical presentation.

There was no significant difference in sensitivity or specificity between the studies with regard to the study design (single-gate versus two-gate studies), the rASRM stages of endometriosis, the reference standard (histological confirmation versus laparoscopic visualisation alone), the target condition (ovarian versus overall pelvic endometriosis), the menstrual cycle phase of testing or the clinical presentations (pain, infertility, ovarian mass versus infertility only or pain only).

With regards to the geographical areas of the studies, studies based in North America reported higher sensitivity compared to the other continents (P = 0.0003), but we were unable to identify the reason for this difference. The other significant factor was year of publication. Studies published after 2010 reported lower estimates of sensitivity compared to the studies published before 2000 (P = 0.026), which is likely to be an indicator that other things have changed in the laboratory methodology including sample processing and types of assays.

We were unable to explore the effect of the following potential sources of heterogeneity.

- Age (adolescents versus later reproductive years): only one study presented separate data for different age groups (younger than 25 years old and 25 to 41 years) in addition to the estimates for the entire included population, and all the remaining studies reported data for the whole reproductive age group.
- Methodological quality: low versus unclear or high risk: all the studies were of low methodological quality with high or unclear risk of bias.

We could not formally assess observer variability bias or bias related to interpretation of results in this review.

DISCUSSION

Summary of main results

We evaluated the diagnostic performance for 47 of the 122 blood biomarkers included in this review. Only four biomarkers were assessed in a sufficient number of studies for a meta-analysis: CA-125 for different cut-offs, CA-19.9 for a cut-off value of > 37 U/ml, IL-6 for a cut-off value of > 1.90 to 2.00 pg/ml and anti-endometrial antibodies. None of the meta-analyses revealed a test with the diagnostic accuracy for a suitable replacement test (sensitivity ≥ 0.94 and specificity ≥ 0.79) or triage test (either a sensitivity ≥ 0.95 with specificity ≥ 0.50, SnOUT, or a sensitivity ≥ 0.50 with a specificity ≥ 0.95, SpIN).

CA-125 was the most studied biomarker, and studies analysed multiple cut-off values within the following groups: > 10.0 to 14.7 U/ml, >16.0 to 17.6 U/ml, > 20.0 U/ml, > 25.0 to 26.0 U/ml, > 30.0 to 33.0 U/ml, > 35.0 to 36.0 U/ml, > 42.0 to 43.0 U/ml. None of these tests were sensitive or specific enough to be considered as a replacement or triage test. The summary estimates of the mean sensitivity and the mean specificity of CA-125 did not all show the expected pattern (higher sensitivity and lower specificity with lower thresholds), but this was likely related to the indirect nature of the comparisons and heterogeneous study groups from different populations. The cut-off > 16.0 to 17.6 U/ml was the best performing of all the CA-125 thresholds subjected to meta-analysis, but it only approached the criteria for a SpIN triage test and showed substantial heterogeneity. CA-125 with a cut-off of > 43.0 U/ml reached the criteria for a replacement test for detecting advanced endometriosis, but only one study demonstrated this, and the data for a wide spectrum of disease was lacking.

The sensitivity of CA-19.9 in detecting endometriosis was too low to meet the criteria for a replacement or triage test. Although only the cut-off value of > 37.0 U/ml was adequately assessed for this biomarker, other thresholds reported in individual studies did not show promising results.

In this review anti-endometrial antibodies and IL-6 with a cut-off value of > 1.90 to 2.00 pg/ml displayed unsatisfactory diagnostic estimates to qualify for either a replacement or triage test. There were too few studies to perform a meaningful evaluation for other cut-off values of IL-6. Although, IL-6 with a cut-off of > 12.20 pg/ml, had a sufficiently high sensitivity and specificity to satisfy the criteria for a replacement test, it was explored in only one study and warrants further validation.

Readers should interpret the findings of the meta-analyses presented in this review with caution. Considering both the level of heterogeneity and the high/unclear risk of bias of the included studies, the results do not seem to be reliable enough to inform clinical practice.

The remaining biomarkers were classified as follows.

- Tests to be validated for their diagnostic potential. This group included:
  - those with an adequate diagnostic performance, but insufficient data to confidently comment on their diagnostic role (less than three studies with the diagnostic estimates meeting the criteria for either a replacement or triage test); and
  - tests where the diagnostic estimates approached the criteria for replacement or triage tests in a small number of studies and where it is possible that they would reach this criteria in further studies (less than three studies with the diagnostic estimates within 5% of the criteria for either replacement or triage tests). These tests are presented in Table 4.
- Tests of limited diagnostic value (at least three studies demonstrating low diagnostic estimates that do not meet or approach the criteria for either replacement or triage test, or report negative findings). We advise against further evaluation

of these biomarkers in the diagnosis of endometriosis. We present these tests in Appendix 8.

- Tests that appear to have limited diagnostic value, but where there is insufficient data to confidently comment on their diagnostic role (less than three studies with low diagnostic estimates or negative findings). We present the full list of tests from this group in Appendix 9. We advise considering further investigation with a focus of specific phases of menstrual cycle, specific types of endometriosis, different cut-off values or different laboratory methods.

Strengths and weaknesses of the review

This review is part of a comprehensive review series on minimally invasive biomarkers for the diagnosis of endometriosis.

The strengths of this review are the following.

1. A very large number of studies, including data for 15,141 women from 141 studies, which allowed meta-analyses for some blood biomarker tests.
2. A very thorough search of the current literature, including studies written in languages other than English.
3. Data extraction by two independent reviewers and use of a modified QUADAS-2 tool to perform quality assessments.
4. Stringent selection criteria, ensuring that eligible studies used prospectively collected samples and only included women of reproductive age, which minimised the risk of bias in interpreting the reference standard and index test.
5. Attempts to contact study authors to obtain any missing information required to assess eligibility and critically appraise the studies.
6. The inclusion of studies that reported negative findings (i.e. demonstrated that biomarker levels did not significantly differ in endometriosis), which provided a more comprehensive evaluation of diagnostic role of the biomarkers and identified the tests of no value in diagnosing the disease.

The main limitation of this review is that there were a low number of small, heterogeneous studies for the majority of the evaluated index tests. This may undermine the reliability of the summary estimates from the meta-analyses and is likely to have contributed to the marked variability in sensitivity and specificity seen for most index tests. For the vast majority of minimally invasive diagnostic tests (or combinations of tests), no meta-analysis was possible. The studies varied with respect to the included populations, severity of endometriosis, menstrual cycle phase at testing, laboratory methods and the cut-off thresholds for index tests. We could not formally explore sources of heterogeneity for the majority of tests due to the low number of studies in most evaluations. Also, most of the included studies evaluated the diagnostic cut-off thresholds using a ROC analysis without any subsequent validation in an independent cohort. Lack of validation of the diagnostic data in conjunction with the low number of studies for the majority of the presented tests contributed to the low quality of evidence presented in this review. We now have an available standardised methodology for fluid biospecimen collection, processing and storage, and we recommend adhering to these standards in future diagnostic studies (Rahimoglu 2014).

Additional weaknesses of this review series are the following.

1. The variation in the selection of the case and control groups with inclusion of participants that may not reflect a clinically representative population. The reported prevalence of endometriosis in this and the other reviews was generally higher (16% to 84%) than previously reported (6% to 10% in the general female population and 35% to 50% in symptomatic women) (Giudice 2004). This may reflect a high level of surgical diagnostic expertise but could be due to pre-selection of more challenging cases in tertiary referral centres, and there is a high risk of patient selection bias in most of the studies. Selection bias appeared to be reduced but not eliminated by consecutively enrolling participants; however, the information on the method of enrolment was missing in most of the included studies. More than a third of the included studies (61/141, 43%) had a two-gate design and included a wide group of participants who underwent surgery for various indications. Inclusion of healthy asymptomatic individuals or participants with other pathological conditions represents a potential selection bias with regard to the control group, which could have biased the test outcomes. Thirty-four studies included either women with a limited spectrum of endometriosis (N = 26) or they did not provide information on the severity of target condition (N = 8). we included these studies to avoid omission of potentially valuable diagnostic information, but each of the above factors could skew the diagnostic estimates in either direction and subsequently interfere with the interpretation of the index test results. It was not possible to evaluate population and disease spectrum effects on the data because there were too few reports for most of the blood biomarkers.
2. We could not rule out inappropriate assignment to the endometriosis and control groups in many studies. Surgical misdiagnosis is a potential cause of bias as most of the included studies did not adequately describe the number and experience of the surgical team, the surgical diagnostic criteria and the surgical methods. We now have a standardised technique for performing laparoscopy, and we recommend that any future studies use this method (Becker 2014). Additionally, we did not confine the studies included in this review to those that reported histological confirmation of endometriotic lesions. Although a recent ESHRE guideline stated that evidence is lacking to support laparoscopy without histology to confirm endometriosis (Dunselman 2014), the clinical significance of histological verification remains debatable. Diagnosis by surgical visualisation only remains a common clinical practice and can be considered reliable when an accurate inspection of the abdominal cavity is performed by experienced surgeons. We chose to include the 66 studies that only reported surgical visualisation as the reference standard, and we did not wish to lose this potentially valuable information. However, this could impact the accuracy of assignment to the case and control groups.
3. The methodology of systematic reviews of diagnostic test accuracy is still emerging, and there are no well-established criteria for replacement or triage diagnostic tests, therefore we chose criteria that were both realistic and clinically applicable to assist in the interpretation of the complex results. For a replacement test, we considered the threshold reported by the one and the only systematic review on accuracy of the reference standard (laparoscopy) in detecting endometriosis to be the most objective (Wykes 2004). The meta-analysis was published in 2004 and included four eligible studies comprising

433 women. We acknowledge the limitations associated with emphasising a single review, particularly if it does not present the latest and possibly more accurate data that reflect advances in surgical expertise and technology. Several studies on accuracy of laparoscopy in detecting endometriosis have been published in the last decade; however, their results were not addressed in a systematic way. A further systematic analysis to evaluate the accuracy of laparoscopy was beyond the scope of this review. The criteria for triage tests utilised the common concepts of SnOUT and SpIN in medical statistics, and the cut-offs were set at levels we considered to be clinically relevant (see Role of index test(s)). We encourage the readers to apply independent interpretations of the presented diagnostic estimates by using thresholds that may be more applicable to specific populations and clinical circumstances.

Applicability of findings to the review question

Based on our use of the QUADAS-2, we assigned a low rank (high concern) to clinical applicability with respect to patient selection in 51% of the studies (72/141). This occurred when the set of participants in the study was broader that seen in clinical practice or when the spectrum of the target condition was limited and the findings may not be applicable to the review question and to clinical practice. We judged the applicability of the index test and reference standard to be satisfactory using the QUADAS-2 tool for all studies. However, the majority of included studies took place in academic institutions with a high level of expertise in laboratory techniques, and the index test outcome measures may not be able to be reproduced in all institutions or extrapolated to general practice.

We excluded some potentially relevant well-designed studies, as they did not directly address the review question. For example, we excluded studies that reported on biomarkers with differential expression in endometriosis, but that did not provide enough information to assess the diagnostic performance of the biomarker. Additionally, we excluded most of the studies that compared endometrioma with other ovarian masses, as they either did not meet our inclusion criteria for reproductive age or assessed the numbers of cysts rather than the number of participants. Therefore we could not fully address the review question on non-invasive diagnosis of ovarian endometriosis. We also excluded some forms of endometriosis, such as bladder, ureteric or endometriosis involving the extrapelvic sites (e.g. umbilicus, hernia sacs, abdominal wall, lung, kidney, etc.), as they are informed predominantly by case reports or small case series, and diagnostic laparoscopy is not an applicable reference test for these conditions. Although these target conditions are rare, from a clinical perspective the diagnostic options for these forms of endometriosis remain unclear.

AUTHORS' CONCLUSIONS

Implications for practice

CA-125 was the most studied technique, but showed only moderate sensitivity and moderate specificity for pelvic endometriosis, which did not meet the criteria for a replacement or triage test. This is consistent with international guidelines, which do not recommend CA-125 testing in women with suspected endometriosis (ACOG 2010; Dunselman 2014; SOGC 2010).

CA-19.9 (cut-off > 37.0 U/ml), IL-6 (cut-off > 1.90 to 2.00 pg/ml) and anti-endometrial antibodies demonstrated an unsatisfactory diagnostic performance in detecting endometriosis and hence have no role in clinical practice.

We suggest cautious interpretation of the presented results. Although studies demonstrated diagnostic potential for a number of tests, the level of heterogeneity, wide confidence intervals and high/unclear risk of bias in most studies included from this review series undermine reliability of the presented results, and hence these data are insufficient to confidently inform clinical practice.

Additional biomarkers, reported in individual studies, displayed diagnostic estimates that qualified for either replacement or triage tests; however, there were not enough data for a meaningful recommendation on the use of any of these tests.

As there is an absence of well-established criteria for an adequate diagnostic test, the diagnostic criteria for replacement and triage tests were determined by the authors of this review in a way that we believe will aid the interpretation for clinically active readers. However, we encourage readers to apply different criteria according to each clinical population and situation.

There is wide recognition that an accurate non-invasive test for endometriosis is likely to confer several advantages over a surgical diagnosis for women with symptoms of endometriosis. These potential advantages include a reduction in cost (both in direct medical costs and in time off work), reduced discomfort, shorter recovery times and a reduction in the rare but serious complications of anaesthesia and surgery. Another benefit of an accurate, non-invasive diagnostic test for endometriosis is the prospect of early diagnosis and timely therapeutic interventions to minimise progression of disease, which can occur in up to 50% of women (D'Hooghe 2002).

An accurate 'negative' non-invasive test is expected to reduce the need for diagnostic surgery in 50 - 70% of women with chronic pelvic pain or infertility (Giudice 2004), although it is likely that some women with a negative test would still require surgery to explore other pathologies. An accurate 'positive' non-invasive test for endometriosis is likely to increase the need for surgery in women with mild symptoms or subfertility (D'Hooghe 2006). Thus, until an accurate non-invasive diagnostic test is developed and tested in large clinical populations, it is impossible to accurately predict its impact on surgical uptake and the number of women that would benefit from performing the test.

Implications for research

Currently, randomised controlled treatment trials require women with and without endometriosis to have had diagnostic surgery for accurate group allocation. For ethical reasons, therapeutic surgery is usually performed at the same time, potentially biasing treatment trial outcomes. Thus our current inability to diagnose and assess the progression of endometriosis in a non-invasive way is a significant limitation to the advancement of clinical research in endometriosis.

Several blood biomarkers reported in this review showed promisingly high diagnostic estimates for detecting endometriosis, but there were too few evaluations to determine their value as replacement or triage tests for a laparoscopic diagnosis. Further

well-designed diagnostic studies are necessary to establish the diagnostic test accuracy and clinical utility of these blood tests.

In this review we identified a list of biomarkers that have no value in detecting endometriosis and hence are not recommended for evaluation in future diagnostic studies. This is important for appropriate allocation of research resources and to guide clinically relevant experimental work in the field. These biomarkers comprise: glycodelin, IGFBP-3, leptin, sICAM-1, MCP-1, hs-CRP, IFN-γ, MIF, TNF-α, WBC, IL-1β, IL-2, IL-4, IL-8, IL-10, IL-12, IL-18, sGM-CSF and the above-mentioned tests evaluated in the meta-analyses.

The QUADAS-2 quality assessment of the included studies identified several weakness in study design that can impede an objective evaluation of the findings. We recommend that future authors consider:

1. including large cohorts after pre-defining the sample size via a power calculation (Liu 2005);
2. focusing on a single-gate design that only includes a clinically relevant population (Rutjes 2005);
3. utilising a diagnostic accuracy study design that adheres to the recommendations of the Standards for Reporting of Diagnostic Accuracy (STARD) initiative (Bossuyt 2003);
4. incorporating the QUADAS checklist into the study design (Whiting 2011);
5. formally assessing inter- and intraobserver variability of the laboratory methods;
6. establishing universally acceptable laboratory methodologies and a diagnostic criteria for a positive test (Rahimoglu 2014);
7. utilising universally acceptable methods of performing laparoscopy as the reference standard test (Becker 2014);
8. implementing validation techniques to assess how the results of a statistical analysis will generalise to an independent data set;
9. undertaking direct comparisons of promising tests in conjunction with a cost-effectiveness analysis;
10. applying testing to different clinical phenotypes rather than to women classified according to rASRM staging (Vitonis 2014); and
11. assessing the long term outcomes and lifetime healthcare costs of women that have participated in diagnostic test accuracy trials of specific diagnostic tests.

Specific opportunities for further research identified by this review include:

1. assessing the diagnostic potential of anti-endometrial antibodies and the tests identified as promising replacement or triage tests in detecting pelvic endometriosis in larger, high quality studies;
2. exploring the value of sequential testing, implementing SnOUT and SpIN triage tests in diagnosing endometriosis in conjunction with a cost-effectiveness evaluation of such testing;
3. directly comparing promising biomarkers in well-designed diagnostic accuracy studies;
4. evaluation of the whole spectrum of disease across all phases of menstrual cycle, aiming to identify the most appropriate target population and the best time of testing;
5. attempting testing in the populations that differ by clinical phenotype rather than by rASRM staging in view of the poor correlation of this classification with clinical presentations and treatment outcomes;
6. adding separate evaluations of blood biomarkers, particularly urocortin and follistatin, CA-125 with the cut-off values above 30 to 42 U/ml and a combination of CA-125 and CA-19.9 to determine if ovarian endometrioma can be distinguished from other ovarian masses in reproductive-aged women; and
7. assessing the long-term outcomes and lifetime healthcare costs of women in diagnostic test accuracy trials that have evaluated specific diagnostic blood tests.

ACKNOWLEDGEMENTS

We would like to thank Associate Professor Petra Macaskill for her valuable comments and substantial contribution to development of the statistical methods for the review. Sincere thanks to the late Professor Ali Akoum and Professor Ian Fraser for their intellectual input and help with drafting of the protocol. We are grateful to Marian Showell, the Trials Search Co-ordinator of the Cochrane Gynaecology and Fertility Group, for her help in designing and conducting the literature search and in locating the full texts of the relevant studies. We thank Dr Deepika Arora for her assistance in study selection, quality appraisal and data extraction. We gratefully acknowledge the help of Ms Erika Ota for translation of the studies published in Japanese language and help of Dr Wai Sun Lam and Dr Minglan Li for translation from Chinese language. We also thank the authors of the review series Emily Liu, Devashana Gupta and Lucy Prentice for their dedicated assistance in studies' selection process. Finally, we thank all contacted authors who contributed information to this review.

REFERENCES

References to studies included in this review

Acien 1989 *{published data only}*

Acien P, Shaw RW, Irvine L, Burford G, Gardner R. CA 125 levels in endometriosis patients before, during and after treatment with danazol or LHRH agonists. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* 1989;**32**(3):241-6.

Agic 2008 *{published data only}*

Agic A, Djalali S, Wolfler MM, Halis G, Diedrich K, Hornung D. Combination of CCR1 mRNA, MCP1, and CA125 measurements in peripheral blood as a diagnostic test for endometriosis. *Reproductive Sciences* 2008;**15**(9):906-11.

Akoum 1996 *{published data only}*

Akoum A, Lemay A, McColl SR, Paradis I, Maheux R. Increased monocyte chemotactic protein-1 level and activity in the peripheral blood of women with endometriosis. Le Groupe d'Investigation en Gynecologie. *American Journal of Obstetrics and Gynecology* 1996;**175**(6):1620-5.

Andreoli 2011 *{published data only}*

Andreoli CG, Genro VK, Souza CA, Michelon T, Bilibio JP, Scheffel C, et al. T helper (Th)1, Th2, and Th17 interleukin pathways in infertile patients with minimal/mild endometriosis. *Fertility and Sterility* 2011;**95**(8):2477-80.

Barbati 1994 *{published data only}*

Barbati A, Cosmi EV, Spaziani R, Ventura R, Montanino G. Serum and peritoneal fluid CA-125 levels in patients with endometriosis. *Fertility and Sterility* 1994;**61**(3):438-42.

Barbosa 2009 *{published data only}*

Barbosa CP, Souza AM, Bianco B, Christofolini D, Bach FAM, Lima GR. Frequency of endometriotic lesions in peritoneum samples from asymptomatic fertile women and correlation with CA125 values. *Sao Paulo Medical Journal = Revista Paulista de Medicina* 2009;**127**(6):342-5.

Barcz 2002 *{published data only}*

Barcz E, Rozewska ES, Kaminski P, Demkow U, Bobrowska K, Marianowski L. Angiogenic activity and IL-8 concentrations in peritoneal fluid and sera in endometriosis. *International Journal of Gynaecology and Obstetrics* 2002;**79**(3):229-35.

Bedaiwy 2002 *{published data only}*

Bedaiwy MA, Falcone T, Sharma RK, Goldberg JM, Attaran M, Nelson DR, et al. Prediction of endometriosis with serum and peritoneal fluid markers: a prospective controlled trial. *Human Reproduction* 2002;**17**(2):426-31.

Bilibio 2014 *{published data only}*

Bilibio JP, Souza CA, Rodini GP, Andreoli CG, Genro VK, de Conto E, et al. Serum prolactin and CA-125 levels as biomarkers of peritoneal endometriosis. *Gynecologic and Obstetric Investigation* 2014;**78**(1):45-52.

Borkowski 2008 *{published data only}*

Borkowski J, Gmyrek GB, Madej JP, Nowacki W, Goluda M, Gabrys M, et al. Serum and peritoneal evaluation of vitamin D-binding protein in women with endometriosis. *Postepy Higieny i Medycyny Doswiadczalnej (Online)* 2008;**62**:103-9.

Braun 1996 *{published data only}*

Braun DP, Gebel H, House R, Rana N, Dmowski WP. Spontaneous and induced synthesis of cytokines by peripheral blood monocytes in patients with endometriosis. *Fertility and Sterility* 1996;**65**(6):1125-9.

Calienno 2008 *{published data only}*

Calienno C, Trio C, Monti B, Ieda N, Cortese M, Varisco E. MCP-1 and IL-8 in serum and ectopic endometrium in women with endometriosis. *Italian Journal of Gynaecology and Obstetrics* 2008;**20**(1):12-8.

Chen 1998 *{published data only}*

Chen FP, Soong YK, Lee N, Lo SK. The use of serum CA-125 as a marker for endometriosis in patients with dysmenorrhea for monitoring therapy and for recurrence of endometriosis. *Acta Obstetricia et Gynecologica Scandinavica* 1998;**77**(6):665-70.

Cho 2007 *{published data only}*

Cho SH, Oh YJ, Nam A, Kim HY, Park JH, Kim JH, et al. Evaluation of serum and urinary angiogenic factors in patients with endometriosis. *American Journal of Reproductive Immunology* 2007;**58**(6):497-504.

Colacurci 1996a *{published data only}*

Colacurci N, Fortunato N, De Franciscis P, Cardone A. Relevance of CA-125 in the evaluation of endometriosis. *Clinical and Experimental Obstetrics & Gynecology* 1996;**23**(3):150-4.

Da Silva 2014 *{published data only}*

Da Silva CM, Vilaca Belo A, Passos Andrade S, Peixoto Campos P, Cristina Franca Ferreira M, Lopes da Silva-Filho A, et al. Identification of local angiogenic and inflammatory markers in the menstrual blood of women with endometriosis. *Biomedicine & Pharmacotherapy* 2014;**68**(7):899-904.

Dayangan Sayan 2013 *{published data only}*

Dayangan Sayan C, Ozaksit MG, Sarikaya E, Gun Eryilmaz O, Mollamahmutoglu L, Deveer R. Serum interleukin-8, CA-125 levels, neutrophil-to-lymphocyte ratios, and combined markers in the diagnosis of endometriosis. *Turkish Journal of Medical Sciences* 2013;**43**(3):417-23.

De Placido 1998 *{published data only}*

De Placido G, Alviggi C, Di P, Carravetta C, Matarese G, Landino G, et al. Serum concentrations of soluble human leukocyte class I antigens and of the soluble intercellular adhesion molecule-1 in endometriosis: Relationship with stage and non-pigmented peritoneal lesions. *Human Reproduction* 1998;**13**(11):3206-10.

Drosdzol-Cop 2012a *{published data only}*

Drosdzol-Cop A, Skrzypulec-Plinta V. Selected cytokines and glycodelin A levels in serum and peritoneal fluid in girls with endometriosis. *Journal of Obstetrics and Gynaecology Research* 2012;**38**(10):1245-53.

Drosdzol-Cop 2012b *{published data only}*

Drosdzol-Cop A, Skrzypulec-Plinta V, Stojko R. Serum and peritoneal fluid immunological markers in adolescent girls with chronic pelvic pain. *Obstetrical & Gynecological Survey* 2012;**67**(6):374-81.

Elgafor el Sharkwy 2013 *{published data only}*

Elgafor el Sharkwy IA. Combination of non-invasive and semi-invasive tests for diagnosis of minimal to mild endometriosis. *Archives of Gynecology and Obstetrics* 2013;**288**(4):793-7.

Fairbanks 2009 *{published data only}*

Fairbanks F, Abrao MS, Podgaec S, Dias JA Jr, de Oliveira RM, Rizzo LV. Interleukin-12 but not interleukin-18 is associated with severe endometriosis. *Fertility and Sterility* 2009;**91**(2):320-4.

Fassbender 2009 *{published data only}*

Fassbender A, D'Hooghe T, Mihalyi A, Kyama C, Simsa P, Lessey BA. Plasma C3a-des-Arg levels in women with and without endometriosis. *American Journal of Reproductive Immunology* 2009;**62**(3):187-95.

Fassbender 2012 *{published data only}*

Fassbender A, Waelkens E, Verbeeck N, Kyama CM, Bokor A, Vodolazkaia A, et al. Proteomics analysis of plasma for early diagnosis of endometriosis. *Obstetrics and Gynecology* 2012;**119**(2 Pt 1):276-85.

Fedele 1989 *{published data only}*

Fedele L, Arcaini L, Vercellini P, Marchini M, Baglioni A, Bianchi S. Serum Ca-125 concentrations in endometriosis. *Acta Europaea Fertilitatis* 1989;**20**(3):137-9.

Ferreira 1994 *{published data only}*

Ferreira CA, Camargos AF. Assessement of serum concentration CA-125 in the diagnosis of endometriosis [Avaliação da dosagem do CA-125 sérico no diagnóstico da endometriose]. *Jornal Brasileiro de Ginecologia* 1994;**104**(9):311-4.

Ferrero 2005a *{published data only}*

Ferrero S, Gillott DJ, Anserini P, Remorgida V, Price KM, Ragni N, et al. Vitamin D binding protein in endometriosis. *Journal of the Society for Gynecologic Investigation* 2005;**12**(4):272-7.

Florio 2007 *{published data only}*

Florio P, Reis FM, Torres PB, Calonaci F, Toti P, Bocchi C, et al. Plasma urocortin levels in the diagnosis of ovarian endometriosis. *Obstetrics and Gynecology* 2007;**110**(3):594-600.

Florio 2009 *{published data only}*

Florio P, Reis FM, Torres PB, Calonaci F, Abrao MS, Nascimento LL, et al. High serum follistatin levels in women with ovarian endometriosis. *Human Reproduction* 2009;**24**(10):2600-6.

Foda 2012 *{published data only}*

Foda AA, Aal IAA. Role of some biomarkers in chronic pelvic pain for early detection of endometriosis in infertile women. *Middle East Fertility Society Journal* 2012;**17**(3):187-94.

Franchi 1993 *{published data only}*

Franchi M, Beretta P, Zanaboni F, Donadello N, Ghezzi F. Use of serum CA125 measurement in patients with endometriosis. *Italian Journal of Gynaecology and Obstetrics* 1993;**5**(4):149-53.

Gagne 2003a *{published data only}*

Gagne D, Rivard M, Page M, Lepine M, Platon C, Shazand K, et al. Development of a nonsurgical diagnostic tool for endometriosis based on the detection of endometrial leukocyte subsets and serum CA-125 levels. Fertility and Sterility 2003; Vol. 80, issue 4:876-85.

Gagne 2003b *{published data only}*

Gagne D, Page M, Robitaille G, Hugo P, Gosselin D. Levels of vascular endothelial growth factor (VEGF) in serum of patients with endometriosis. *Human Reproduction* 2003;**18**(8):1674-80.

Gazvani 1998 *{published data only}*

Gazvani MR, Christmas S, Quenby S, Kirwan J, Johnson PM, Kingsland CR. Peritoneal fluid concentrations of interleukin-8 in women with endometriosis: relationship to stage of disease. *Human Reproduction* 1998;**13**(7):1957-61.

Glitz 2009 *{published data only}*

Glitz C, Souza CA, Rodini GP, Genro V, Bilibio JP, Senger M, et al. Peritoneal and serum interleukin-18 levels are not increased in women with minimum or mild endometriosis. *Brazilian Journal of Medical and Biological Research* 2009;**42**(11):1039-43.

Gogacz 2014 *{published data only}*

Gogacz M, Winkler I, Bojarska-Junak A, Tabarkiewicz J, Semczuk A, Rechberger T, et al. T regulatory lymphocytes in patients with endometriosis. *Molecular Medicine Reports* 2014;**10**(2):1072-6.

Goluda 1998 *{published data only}*

Goluda M, Kuliczkowski K, Jedryka M. The concentration of exfoliative adhesion molecules (ICAM-1 and E-Selectin) in serum and peritoneal fluid of women with endometriosis. *Ginekologia Polska* 1998;**69**(12):1175-8.

Gorai 1993 *{published data only}*

Gorai I, Ishikawa M, Onose R, Hirahara F, Minaguchi H. Antiendometrial autoantibodies are generated in patients with endometriosis. *American Journal of Reproductive Immunology* 1993;**29**(2):116-23.

Guerriero 1996a *{published data only}*

Guerriero S, Ajossa S, Paoletti AM, Mais V, Angiolucci M, Melis GB. Tumor markers and transvaginal ultrasonography in the diagnosis of endometrioma. *Obstetrics and Gynecology* 1996;**88**(3):403-7.

Guerriero 1996b *{published data only}*

Guerriero S, Mais V, Ajossa S, Paoletti AM, Angiolucci M, Melis GB. Transvaginal ultrasonography combined with CA-125

plasma levels in the diagnosis of endometrioma. *Fertility and Sterility* 1996;**65**(2):293-8.

Gurgan 1990 *{published data only}*

Gurgan T, Kisnisci H, Yarali H, Aksu T, Zeyneloglu H, Develioglu O. Serum and peritoneal fluid CA-125 levels in early stage endometriosis. *Gynecologic and Obstetric Investigation* 1990;**30**:105-8.

Gurgan 1999 *{published data only}*

Gurgan T, Bukulmez O, Yarali H, Tanir M, Akyildiz S. Serum and peritoneal fluid levels of IGF I and II and insulinlike growth binding protein-3 in endometriosis. *Journal of Reproductive Medicine* 1999;**44**:450-4.

Hallamaa 2012 *{published data only}*

Hallamaa M, Suvitie P, Huhtinen K, Matomaki J, Poutanen M, Perheentupa A. Serum HE4 concentration is not dependent on menstrual cycle or hormonal treatment among endometriosis patients and healthy premenopausal women. *Gynecologic Oncology* 2012;**125**(3):667-72.

Hapangama 2008 *{published data only}*

Hapangama DK, Turner MA, Drury JA, Quenby S, Saretzki G, Martin-Ruiz C, et al. Endometriosis is associated with aberrant endometrial expression of telomerase and increased telomere length. *Human Reproduction* 2008;**23**(7):1511-9.

Harada 2002 *{published data only}*

Harada T, Kubota T, Aso T. Usefulness of CA19-9 versus CA125 for the diagnosis of endometriosis. *Fertility and Sterility* 2002;**78**(4):733-9.

Hassa 2009 *{published data only}*

Hassa H, Tanir H Mete, Tekin B, Kirilmaz S D, Sahin Mutlu F. Cytokine and immune cell levels in peritoneal fluid and peripheral blood of women with early- and late-staged endometriosis. *Archives of Gynecology and Obstetrics* 2009;**279**(6):891-5.

Hornstein 1995 *{published data only}*

Hornstein MD, Harlow BL, Thomas PP, Check JH. Use of a new CA 125 assay in the diagnosis of endometriosis. *Human Reproduction* 1995;**10**(4):932-4.

Inagaki 2003 *{published data only}*

Inagaki J, Sugiura-Ogasawara M, Nomizu M, Nakatsuka M, Ikuta K, Suzuki N, et al. An association of IgG anti-laminin-1 autoantibodies with endometriosis in infertile patients. *Human Reproduction* 2003;**18**(3):544-9.

Iwasaki 1993 *{published data only}*

Iwasaki K, Makino T, Maruyama T, Matsubayashi H, Nozawa S, Yokokura T. Leukocyte subpopulations and natural killer activity in endometriosis. *International Journal of Fertility and Menopausal Studies* 1993;**38**(4):229-34.

Jee 2008 *{published data only}*

Jee BC, Suh CS, Kim SH, Moon SY. Serum soluble CD163 and interleukin-6 levels in women with ovarian endometriomas. *Gynecologic and Obstetric Investigation* 2008;**66**(1):47-52.

Jia 2013 *{published data only}*

Jia SZ, Yang Y, Lang J, Sun P, Leng J. Plasma miR-17-5p, miR-20a and miR-22 are down-regulated in women with endometriosis. *Human Reproduction* 2013;**28**(2):322-30.

Joshi 1986 *{published data only}*

Joshi SG, Zamah NM, Raikar RS, Buttram VC Jr, Henriques ES, Gordon M. Serum and peritoneal fluid proteins in women with and without endometriosis. *Fertility and Sterility* 1986;**46**(6):1077-82.

Kalu 2007 *{published data only}*

Kalu E, Sumar N, Giannopoulos T, Patel P, Croucher C, Sherriff E, et al. Cytokine profiles in serum and peritoneal fluid from infertile women with and without endometriosis. *Journal of Obstetrics and Gynaecology Research* 2007;**33**(4):490-5.

Khan 2006 *{published data only}*

Khan KN, Masuzaki H, Fujishita A, Kitajima M, Hiraki K, Miura S, et al. Peritoneal fluid and serum levels of hepatocyte growth factor may predict the activity of endometriosis. *Acta Obstetricia et Gynecologica Scandinavica* 2006;**85**(4):458-66.

Khan 2012 *{published data only}*

Khan KN, Kitajima M, Yamaguchi N, Fujishita A, Nakashima M, Ishimaru T, et al. Role of prostaglandin E2 in bacterial growth in women with endometriosis. *Human Reproduction* 2012;**27**(12):3417-24.

Khan 2013 *{published data only}*

Khan KN, Kitajima M, Inoue T, Tateishi S, Fujishita A, Nakashima M, et al. Additive effects of inflammation and stress reaction on Toll-like receptor 4-mediated growth of endometriotic stromal cells. *Human Reproduction* 2013;**28**(10):2794-803.

Khanaki 2012 *{published data only}*

Khanaki K, Nouri M, Ardekani AM, Ghassemzadeh A, Shahnazi V, Sadeghi MR, et al. Evaluation of the relationship between endometriosis and omega-3 and omega-6 polyunsaturated fatty acids. *Iranian Biomedical Journal* 2012;**16**(1):38-43.

Kianpour 2012 *{published data only}*

Kianpour M, Nematbakhsh M, Ahmadi SM. C-reactive protein of serum and peritoneal fluid in endometriosis. *Iranian Journal of Nursing and Midwifery Research* 2012;**17**(2 Suppl 1):S115-9.

Kianpour 2013 *{published data only}*

Kianpour M, Nematbakhsh M, Ahmadi SM, Jafarzadeh M, Hajjarian M, Pezeshki Z, et al. Serum and peritoneal fluid levels of vascular endothelial growth factor in women with endometriosis. *International Journal of Fertility & Sterility* 2013;**7**(2):96-9.

Kim 2008 *{published data only}*

Kim JY, Kim H, Suh CS, Kim SH, Choi YM, Kim JG. The G(-2518)A polymorphism of monocyte chemotactic protein-1 (MCP-1) and its serum and peritoneal fluid levels in Korean women with endometriosis. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2008;**139**(1):106-10.

Kitawaki 2005 *{published data only}*

Kitawaki J, Ishihara H, Koshiba H, Kiyomizu M, Teramoto M, Kitaoka Y. Usefulness and limits of CA-125 in diagnosis of endometriosis without associated ovarian endometriomas. *Human Reproduction* 2005;**20**(7):1999-2003.

Kocbek 2013 *{published data only}*

Kocbek V, Vouk K, Mueller MD, Rizner TL, Bersinger NA. Elevated glycodelin-A concentrations in serum and peritoneal fluid of women with ovarian endometriosis. *Gynecological Endocrinology* 2013;**29**(5):455-9.

Kocbek 2014a *{published data only}*

Kocbek V, Bersinger NA, Brglez V, Mueller MD, Petan T, Rizner TL. Phospholipase A2 group IIA is elevated in endometriomas but not inperitoneal fluid and serum of ovarian endometriosis patients. *Gynecological Endocrinology* 2014;**31**(3):214-8.

Kocbek 2014b *{published data only}*

Kocbek V, Hevir-Kene N, Bersinger NA, Mueller MD, Rizner TL. Increased levels of biglycan in endometriomas and peritoneal fluid samples from ovarian endometriosis patients. *Gynecological Endocrinology* 2014;**30**(7):520-4.

Koninckx 1996 *{published data only}*

Koninckx PR, Meuleman C, Oosterlynck D, Cornillie FJ. Diagnosis of deep endometriosis by clinical examination during menstruation and plasma CA-125 concentration. *Fertility and Sterility* 1996;**65**(2):280-7.

Kubatova 2013 *{published data only}*

Kubatova A, Erdem A, Erdem M, FiratMutlu M, Korucuoglu U. Serum cytokine and growth factor levels in patients with endometriosis. *Central-European Journal of Immunology* 2013;**38**(4):500-4.

Kuessel 2014 *{published data only}*

Kuessel L, Jaeger-Lansky A, Pateisky P, Rossberg N, Schulz A, Schmitz AA, et al. Cytokeratin-19 as a biomarker in urine and in serum for the diagnosis of endometriosis - a prospective study. *Gynecological Endocrinology* 2014;**30**(1):38-41.

Kurdoglu 2009 *{published data only}*

Kurdoglu Z, Gursoy R, Kurdoglu M, Erdem M, Erdem O, Erdem A. Comparison of the clinical value of CA 19-9 versus CA 125 for the diagnosis of endometriosis. *Fertility and Sterility* 2009;**92**(5):1761-3.

Lambrinoudaki 2009 *{published data only}*

Lambrinoudaki IV, Augoulea A, Christodoulakos GE, Economou EV, Kaparos G, Kontoravdis A, et al. Measurable serum markers of oxidative stress response in women with endometriosis. *Fertility and Sterility* 2009;**91**(1):46-50.

Lamp 2012 *{published data only}*

Lamp M, Saare M, Kadastik U, Karro H, Salumets A, Uibo R, et al. Survivin promoter polymorphisms and autoantibodies in endometriosis. *Journal of Reproductive Immunology* 2012;**96**(1-2):95-100.

Lanzone 1991 *{published data only}*

Lanzone A, Marana R, Muscatello R, Fulghesu AM, Dellacqua S, Caruso A, et al. Serum Ca-125 Levels in the Diagnosis and Management of Endometriosis. *Journal of Reproductive Medicine* 1991;**36**(8):603-7.

Li 2005 *{published data only}*

Li JX, Dai SZ, Liu H, Cao YM, Liu SQ. Study on the changes of T-lymphocyte subsets in the patients with endometriosis. *Zhonghua Fuchanke Zazhi [Chinese Journal of Obstetrics and Gynecology]* 2005;**40**(1):17-20. Chinese.

Lima 2006 *{published data only}*

Lima AP, Rosa E Silva AAM, Moura MD. FSH, LH, estradiol, progesterone, and histamine concentrations in serum, peritoneal fluid and follicular fluid of women with and without endometriosis [Concentrações de FSH, LH, estradiol, progesterona e histamina no soro, no fluido peritoneal e no fluido folicular de mulheres com e sem endometriose]. *Revista Brasileira de Ginecologia e Obstetrica* 2006;**28**(11):643-51.

Lin 2005 *{published data only}*

Lin J, Zhang XM, Deng L, Chen ZY, Chen L. Determination of interleukine-16 levels in peritoneal fluid and serum of women with endometriosis. *Zhejiang Daxue Xuebao (Yixueban) [Journal of Zhejiang University (Medical Sciences)]* 2005;**34**(3):260-2. Chinese.

Liu 2009 *{published data only}*

Liu HY, Zheng YH, Zhang JZ, Leng JH, Sun DW, Liu ZF, et al. Establishment of endometriosis diagnostic model using plasma protein profiling. *Chung-Hua Fu Chan Ko Tsa Chih [Chinese Journal of Obstetrics and Gynecology]* 2009;**44**(8):601-4. Chinese.

Mabrouk 2012 *{published data only}*

Mabrouk M, Elmakky A, Caramelli E, Farina A, Mignemi G, Venturoli S, et al. Performance of peripheral (serum and molecular) blood markers for diagnosis of endometriosis. *Archives of Gynecology and Obstetrics* 2012;**285**(5):1307-12.

Maeda 2002a *{published data only}*

Maeda N, Izumiya C, Oguri H, Kusume T, Yamamoto Y, Fukaya T. Aberrant expression of intercellular adhesion molecule-1 and killer inhibitory receptors induces immune tolerance in women with pelvic endometriosis. *Fertility and Sterility* 2002;**77**(4):679-83.

Maeda 2002b *{published data only}*

Maeda N, Izumiya C, Yamamoto Y, Oguri H, Kusume T, Fukaya T. Increased killer inhibitory receptor KIR2DL1 expression among natural killer cells in women with pelvic endometriosis. *Fertility and Sterility* 2002;**77**(2):297-302.

Maiorana 2007 *{published data only}*

Maiorana A, Cicerone C, Niceta A, Allio L. Evaluation of serum CA 125 levels in patients with pelvic pain related to endometriosis. *International Journal of Biological Markers* 2007;**22**:200-2.

Markham 1997a *{published data only}*

Markham R, Fraser IS, Song JY, Young L, Chullapram T. Blood and peritoneal fluid concentrations of TNFalpha and RANTES in

patients with and without endometriosis. *Australian Journal of Medical Science* 1997;**18**(4):116-8.

Martinez 2007 *{published data only}*

Martinez S, Garrido N, Coperias JL, Pardo F, Desco J, Garcia-Velasco JA, et al. Serum interleukin-6 levels are elevated in women with minimal-mild endometriosis. *Human Reproduction* 2007;**22**(3):836-42.

Matalliotakis 2003a *{published data only}*

Matalliotakis IM, Goumenou AG, Koumantakis GE, Neonaki MA, Koumantakis EE, Dionyssopoulou E, et al. Serum concentrations of growth factors in women with and without endometriosis: the action of anti-endometriosis medicines. *International Immunopharmacology* 2003;**3**(1):81-9.

Matalliotakis 2004 *{published data only}*

Matalliotakis IM, Arici A, Goumenou AG, Katassos T, Karkavitsas N, Koumantakis EE. Comparison of the effects of leuprorelin acetate and danazol treatments on serum CA-125 levels in women with endometriosis. *International Journal of Fertility and Women's Medicine* 2004;**49**(2):75-8.

Matveeva 1990 *{published data only}*

Matveeva NK, Volkov NI, Petrenko EP, Mit'kin VV, Pshenichnikova TI, Sukhikh GT. Immunological studies of peripheral blood in patients with external genital endometriosis and infertility. *Akusherstvo i Ginekologiia* 1990;**8**:48-51.

Mier-Cabrera 2011 *{published data only}*

Mier-Cabrera J, Jimenez-Zamudio L, Garcia-Latorre E, Cruz-Orozco O, Hernandez-Guerrero C. Quantitative and qualitative peritoneal immune profiles, T-cell apoptosis and oxidative stress-associated characteristics in women with minimal and mild endometriosis. *BJOG: An International Journal of Obstetrics and Gynaecology* 2011;**118**(1):6-16.

Mihalyi 2010 *{published data only}*

Mihalyi A, Gevaert O, Kyama CM, Simsa P, Pochet N, De Smet F, et al. Non-invasive diagnosis of endometriosis based on a combined analysis of six plasma biomarkers. *Human Reproduction* 2010;**25**(3):654-64.

Mohamed 2013 *{published data only}*

Mohamed ML, El Behery MM, Mansour SAEA. Comparative study between VEGF-A and CA-125 in diagnosis and follow-up of advanced endometriosis after conservative laparoscopic surgery. *Archives of Gynecology and Obstetrics* 2013;**287**(1):77-82.

Molo 1994 *{published data only}*

Molo MW, Kelly M, Radwanska E, Binor Z. Preoperative serum CA-125 and CA-72 in predicting endometriosis in infertility patients. *Journal of Reproductive Medicine* 1994;**39**(12):964-6.

Morin 2005 *{published data only}*

Morin M, Bellehumeur C, Therriault MJ, Metz C, Maheux R, Akoum A. Elevated levels of macrophage migration inhibitory factor in the peripheral blood of women with endometriosis. *Fertility and Sterility* 2005;**83**(4):865-72.

Muscatello 1992 *{published data only}*

Muscatello R, Cucinelli F, Fulghesu A, Lanzone A, Caruso A, Mancuso S. Multiple serum marker assay in the diagnosis of endometriosis. *Gynecological Endocrinology* 1992;**6**:265-9.

Odukoya 1996 *{published data only}*

Odukoya O, Bansal A, Cooke I. Serum endometrial IgG antibodies and soluble CD23 concentrations in patients with endometriosis. *Acta Obstetricia et Gynecologica Scandinavica* 1996;**75**(10):927-31.

Ohata 2008 *{published data only}*

Ohata Y, Harada T, Miyakoda H, Taniguchi F, Iwabe T, Terakawa N. Serum interleukin-8 levels are elevated in patients with ovarian endometrioma. *Fertility and Sterility* 2008;**90**(4):994-9.

Oku 2004 *{published data only}*

Oku H, Tsuji Y, Kashiwamura SI, Adachi S, Kubota A, Okamura H, et al. Role of IL-18 in pathogenesis of endometriosis. *Human Reproduction* 2004;**19**(3):709-14.

Olkowska-Truchanowicz 2013 *{published data only}*

Olkowska-Truchanowicz J, Bocian K, Maksym RB, Bialoszewska A, Wlodarczyk D, Baranowski W, et al. CD4+ CD25+ FOXP3+ regulatory T cells in peripheral blood and peritoneal fluid of patients with endometriosis. *Human Reproduction* 2013;**28**(1):119-24.

Othman 2008 *{published data only}*

Othman EEDR, Homung D, Salem HT, Khalifa EA, El-Metwally TH, Al-Hendy A. Serum cytokines as biomarkers for nonsurgical prediction of endometriosis. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2008;**137**(2):240-6.

Ozhan 2014 *{published data only}*

Ozhan E, Kokcu A, Yanik K, Gunaydin M. Investigation of diagnostic potentials of nine different biomarkers in endometriosis. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2014;**178**:128-33.

Paiva 2014 *{published data only}*

Paiva P, Lappas M, Barker G, Healey M. Using symptom scores, lifestyle measures and biochemical markers to create a test for endometriosis. *Journal of Endometriosis* 2014;**6**(3):135-43.

Patton 1986 *{published data only}*

Patton PE, Field CS, Harms RW, Coulam CB. CA-125 levels in endometriosis. *Fertility and Sterility* 1986;**45**(6):770-3.

Philippoussis 2004 *{published data only}*

Philippoussis F, Gagne D, Hugo P, Gosselin D. Concentrations of alpha-fetoprotein, insulin-like growth factor binding protein-3, c-erbB-2, and epidermal growth factor in serum of patients with endometriosis. *Journal of the Society for Gynecologic Investigation* 2004;**11**(3):175-81.

Pittaway 1989 *{published data only}*

Pittaway DE, Douglas JW. Serum CA-125 in women with endometriosis and chronic pelvic pain. *Fertility and Sterility* 1989;**51**(1):68-70.

Podgaec 2007 *{published data only}*

Podgaec S, Abrao MS, Dias Jr JA, Rizzo LV, de Oliveira RM, Baracat EC. Endometriosis: An inflammatory disease with a Th2 immune response component. *Human Reproduction* 2007;**22**(5):1373-9.

Ramos 2012 *{published data only}*

Ramos IML, Podgaec S, Abrao MS, de Oliveira R, Baracat EC. Evaluation of CA-125 and soluble CD-23 in patients with pelvic endometriosis: A case-control study. *Revista da Associacao Medica Brasileira* 2012;**58**(1):26-32.

Randall 2007 *{published data only}*

Randall GW, Gantt PA, Poe-Zeigler RL, Bergmann CA, Noel ME, Strawbridge WR, et al. Serum antiendometrial antibodies and diagnosis of endometriosis. *American Journal of Reproductive Immunology* 2007;**58**(4):374-82.

Riley 2007 *{published data only}*

Riley CF, Moen MH, Videm V. Inflammatory markers in endometriosis: reduced peritoneal neutrophil response in minimal endometriosis. *Acta Obstetricia et Gynecologica Scandinavica* 2007;**86**(7):877-81.

Rosa E Silva 2007 *{published data only}*

Rosa E Silva ACJS, Rosa E Silva JC, Ferriani RA. Serum CA-125 in the diagnosis of endometriosis. *International Journal of Gynaecology and Obstetrics* 2007;**96**(3):206-7.

Rosa E Silva 2014 *{published data only}*

Rosa ESJC, Do Amara VF, Mendonca JL, Rosa ESACJDS, Nakao LS, Neto OBP, et al. Serum markers of oxidative stress and endometriosis. *Clinical and Experimental Obstetrics & Gynecology* 2014;**41**(4):371-4.

Salehpour 2009 *{published data only}*

Salehpour S, Sene AA, Mehrjerdi EK, Akhoond MR. The correlation between serum and peritoneal fluid CA125 level in women with pelvic endometriosis. *International Journal of Fertility & Sterility* 2009;**3**(1):29-34.

Seeber 2008 *{published data only}*

Seeber B, Sammel MD, Fan X, Gerton GL, Shaunik A, Chittams J, et al. Panel of markers can accurately predict endometriosis in a subset of patients. *Fertility & Sterility* 2008;**89**(5):1073-81.

Seeber 2010 *{published data only}*

Seeber B, Sammel MD, Fan X, Gerton GL, Shaunik A, Chittams J, et al. Proteomic analysis of serum yields six candidate proteins that are differentially regulated in a subset of women with endometriosis. *Fertility & Sterility* 2010;**93**(7):2137-44.

Somigliana 2002 *{published data only}*

Somigliana E, Vigano P, Candiani M, Felicetta I, Di Blasio AM, Vignali M. Use of serum-soluble intercellular adhesion molecule-1 as a new marker of endometriosis. *Fertility & Sterility* 2002;**77**(5):1028-31.

Somigliana 2004 *{published data only}*

Somigliana E, Vigano P, Tirelli AS, Felicetta I, Torresani E, Vignali M, et al. Use of the concomitant serum dosage of CA 125, CA 19-9 and interleukin-6 to detect the presence of endometriosis. Results from a series of reproductive age women undergoing laparoscopic surgery for benign gynaecological conditions. *Human Reproduction* 2004;**19**(8):1871-6.

Steff 2004a *{published data only}*

Steff AM, Gagne D, Page M, Rioux A, Hugo P, Gosselin D. Serum concentrations of insulin-like growth factor-1, soluble tumor necrosis factor receptor-1 and angiogenin in endometriosis patients. *American Journal of Reproductive Immunology* 2004;**51**(2):166-73.

Suen 2014 *{published data only}*

Suen JL, Chang Y, Chiu PR, Hsieh TH, Hsi E, Chen YC, et al. Serum level of IL-10 is increased in patients with endometriosis, and IL-10 promotes the growth of lesions in a murine model. *American Journal of Pathology* 2014;**184**(2):464-71.

Szczepanska 2001a *{published data only}*

Szczepaska M, Skrzypczak J, Kamieniczna M, Kurpisz M. Antizona and antisperm antibodies in women with endometriosis and/or infertility. *Fertility & Sterility* 2001;**75**(1):97-105.

Szczepanska 2001b *{published data only}*

Szczepanska M, Mikolajczyk M, Raczynska P, Skrzypczak J. The evaluation of IL-12 levels in peritoneal fluid and serum of women with endometriosis. *Ginekologia Polska* 2001;**72**(5):408-17.

Szubert 2012 *{published data only}*

Szubert M, Suzin J, Wierzbowski T, Kowalczyk-Amico K. CA-125 concentration in serum and peritoneal fluid in patients with endometriosis - preliminary results. *Archives of Medical Science* 2012;**8**(3):504-8.

Szubert 2014 *{published data only}*

Szubert M, Suzin J, Duechler M, Szuławska A, Czyż M, Kowalczyk-Amico K. Evaluation of selected angiogenic and inflammatory markers in endometriosis before and after danazol treatment. *Reproduction, Fertility, and Development* 2014;**26**(3):414-20.

Thubert 2014 *{published data only}*

Thubert T, Santulli P, Marcellin L, Menard S, M'Baye M, Streuli I, et al. Measurement of hs-CRP is irrelevant to diagnose and stage endometriosis: prospective study of 834 patients. *American Journal of Obstetrics and Gynecology* 2014;**210**(6):533 e1-e10.

Tokmak 2011 *{published data only}*

Tokmak A, Ugur M, Tonguc E, Var T, Moraloglu O, Ozaksit G. The value of urocortin and Ca-125 in the diagnosis of endometrioma. *Archives of Gynecology and Obstetrics* 2011;**283**(5):1075-9.

Tuten 2014a *{published data only}*

Tuten A, Kucur M, Imamoglu M, Kaya B, Acikgoz AS, Yilmaz N, et al. Copeptin is associated with the severity of endometriosis. *Archives of Gynecology and Obstetrics* 2014;**290**(1):75-82.

Vercellini 1993 *{published data only}*

Vercellini P, De Benedetti F, Rossi E, Colombo A, Trespidi L, Crosignani PG, et al. Tumor necrosis factor in plasma and peritoneal fluid of women with and without endometriosis. Gynecologic and Obstetric Investigation 1993; Vol. 36, issue 1:39-41.

Verit 2008 *{published data only}*

Verit FF, Erel O, Celik N. Serum paraoxonase-1 activity in women with endometriosis and its relationship with the stage of the disease. *Human Reproduction* 2008;**23**(1):100-4.

Vigano 2002 *{published data only}*

Vigano P, Somigliana E, Matrone R, Dubini A, Barron C, Vignali M, et al. Serum leptin concentrations in endometriosis. *Journal of Clinical Endocrinology and Metabolism* 2002;**87**(3):1085-7.

Vigil 1999 *{published data only}*

Vigil PP, Aglony IM, Kolbach RM, Rubio AV, Villarroel Del Pino L. Endometriosis and CA 125 in women of reproductive age [Endometriosis y CA 125 en mujeres en edad reproductiva]. Revista Chilena de Obstetrica y Ginecologia 1999; Vol. 64, issue 5:385-8.

Vodolazkaia 2011 *{published data only}*

Vodolazkaia A, Bossuyt X, Fassbender A, Kyama CM, Meuleman C, Peeraer K, et al. A high sensitivity assay is more accurate than a classical assay for the measurement of plasma CRP levels in endometriosis. *Reproductive Biology and Endocrinology* 2011;**9**:113.

Vodolazkaia 2012 *{published data only}*

Vodolazkaia A, El-Aalamat Y, Popovic D, Mihalyi A, Bossuyt X, Kyama CM, et al. Evaluation of a panel of 28 biomarkers for the non-invasive diagnosis of endometriosis. *Human Reproduction* 2012;**27**:2698-711.

Vouk 2012 *{published data only}*

Vouk K, Hevir N, Ribic-Pucelj M, Haarpaintner G, Scherb H, Osredkar J, et al. Discovery of phosphatidylcholines and sphingomyelins as biomarkers for ovarian endometriosis. *Human Reproduction* 2012;**27**(10):2955-65.

Wang 2013a *{published data only}*

Wang WT, Zhao YN, Han BW, Hong SJ, Chen YQ, Wang X, et al. Circulating microRNAs identified in a genome-wide serum microRNA expression analysis as noninvasive biomarkers for endometriosis [Study on polymorphism of human leukocyte antigen I in patients with endometriosis]. *Journal of Clinical Endocrinology and Metabolism* 2013;**98**(1):281-9.

Webster 2013 *{published data only}*

Webster KE, Kennedy SH, Becker CM. Levels of circulating angiogenic cells are not altered in women with endometriosis. *Human Reproduction* 2013;**28**(3):651-7.

Wei 2005 *{published data only}*

Wei XQ, Zhang Y, Tang M. Leptin levels and infertile patients with endometriosis. *Zhongnan Daxue Xuebao (Yixue Ban) [Journal of Central South University (Medical Sciences)]* 2005;**30**(4):487-8. Chinese.

Wild 1991a *{published data only}*

Wild RA, Hirisave V, Bianco A, Podczaski ES, Demers LM. Endometrial antibodies versus CA-125 for the detection of endometriosis. *Fertility and Sterility* 1991;**55**(1):90-4.

Wolfler 2009 *{published data only}*

Wolfler MM, Schwamborn K, Otten D, Hornung D, Liu HY, Rath W. Mass spectrometry and serum pattern profiling for analyzing the individual risk for endometriosis: promising insights?. *Fertility and Sterility* 2009;**91**(6):2331-7.

Wu 1998 *{published data only}*

Wu MH, Yang BC, Hsu CC, Lee YC, Huang KE. The expression of soluble intercellular adhesion molecule-1 in endometriosis. *Fertility and Sterility* 1998;**70**(6):1139-42.

Yagmur 2013 *{published data only}*

Yagmur E, Bastu E, Karamustafaoglu-Balci B, Akhan SE, Buyru F. Non-invasive diagnosis of endometriosis based on a combined analysis of four plasma biomarkers. *Central-European Journal of Immunology* 2013;**38**(2):154-8.

Yang 1994 *{published data only}*

Yang F, Fang AH, Gu RF. Value of CA125 (carbon hydrate tumor-associated antigen) and endometrial antibodies for the detections of endometriosis. *Zhonghua Fuchanke Zazhi [Chinese Journal of Obstetrics and Gynecology]* 1994;**29**(3):147-149, 189. Chinese.

Yavuzcan 2013 *{published data only}*

Yavuzcan A, Caglar M, Ustun Y, Dilbaz S, Ozdemir I, Yildiz E, et al. Evaluation of mean platelet volume, neutrophil/ lymphocyte ratio and platelet/lymphocyte ratio in advanced stage endometriosis with endometrioma [Endometrioma bulunan ileri evre endometrioziste ortalama trombosit hacmi, nötrofil/ lenfosit orani ve trombosit/lenfosit oraninin degerlendirilmesi]. *Journal of the Turkish German Gynecological Association* 2013;**14**(4):210-5.

Zeng 2005 *{published data only}*

Zeng F, Xue M, Zevallos HBV, Lai D, Arthur J, Ng C, et al. Diagnostic value of the detection of aromatase cytochrome P450 and CA125 for endometriosis. *Zhongnan Daxue Xuebao (Yixue Ban) [Journal of Central South University (Medical Sciences)]* 2005;**30**(6):682-5. Chinese.

Zhang 2005a *{published data only}*

Zhang X, Lin J, Deng L, Chen Z, Chen L. Peritoneal fluid and serum concentration of interleukin-16 in women with endometriosis. *Acta Obstetricia et Gynecologica Scandinavica* 2005;**84**(3):297-8.

Zhang 2005b *{published data only}*

Zhang Y, Peng LX, Meng L. Measurements of interleukin-18 in peritoneal fluid and serum of patients with endometriosis.

Zhongnan Daxue Xuebao (Yixue Ban) [Journal of Central South University (Medical Sciences)] 2005;**30**(6):731-2. Chinese.

Zhang 2006a *{published data only}*

Zhang C, Maeda N, Izumiya C, Yamamoto Y, Kusume T, Oguri H, et al. Killer immunoglobulin-like receptor and human leukocyte antigen expression as immunodiagnostic parameters for pelvic endometriosis. *American Journal of Reproductive Immunology* 2006;**55**(2):106-14.

Zhang 2006b *{published data only}*

Zhang R, Luo R. Soluble intercellular adhesion molecule 1 in serum and peritoneal fluid on patients with endometriosis. *Wuhan Yike Daxue Xuebao (Yixueban) [Medical Journal of Wuhan University]* 2006;**27**(1):97-9. Chinese.

References to studies excluded from this review

Abdallah 2006 *{published data only}*

Abdallah MA, Kang S, Nakajima ST, Gercel-Taylor C. Matrix metalloproteinase-9 activity in the plasma of patients after surgical resection of endometriomas. *Fertility and Sterility* 2006;**85**(6):1847-8.

Abrao 1997 *{published data only}*

Abrao MS, Podgaec S, Martorelli B, Ramos LO, Pinotti JA, de Oliveira RM. The use of biochemical markers in the diagnosis of pelvic endometriosis. *Human Reproduction* 1997;**12**:2523-7.

Abrao 1999 *{published data only}*

Abrao MS, Podgaec S, Pinotti JA, de Oliveira RM. Tumor markers in endometriosis. *International Journal of Gynaecology and Obstetrics* 1999;**66**:19-22.

Acien 2007 *{published data only}*

Acien P, Velasco I, Gutierrez M, Martinez-Beltran M. Aromatase expression in endometriotic tissues and its relationship to clinical and analytical findings. *Fertility and Sterility* 2007;**88**(1):32-8.

Adamyan 1993 *{published data only}*

Adamyan LV, Fanchenko ND, Alexeyeva ML, Andreyeva YN, Novikov YA, Jahan I. Hormonal and immunologic methods in the diagnosis and treatment of patients with benign ovarian tumors and endometriotic cysts. *International Journal of Fertility* 1993;**38**(2):92-8.

Agic 2007 *{published data only}*

Agic Ar, Xu H, Rehbein M, Wolfler MM, Ebert AD, Hornung D. Cognate chemokine receptor 1 messenger ribonucleic acid expression in peripheral blood as a diagnostic test for endometriosis. *Fertility and Sterility* 2007;**87**:982-4.

Alcazar 2011 *{published data only}*

Alcazar JL, Guerriero S, Minguez JA, Ajossa S, Paoletti AM, Ruiz-Zambrana A, et al. Adding cancer antigen 125 screening to gray scale sonography for predicting specific diagnosis of benign adnexal masses in premenopausal women: is it worthwhile?. *Journal of Ultrasound in Medicine* 2011;**30**(10):1381-6.

Amaral 2006 *{published data only}*

Amaral VF, Ferriani RA, Sa Marcos FS, Nogueira AA, Rosa e Silva JC, Rosa e Silva AC, et al. Positive correlation between serum and peritoneal fluid CA-125 levels in women with pelvic endometriosis. *Sao Paulo Medical Journal* 2006;**124**(4):223-7.

Ammendola 2008 *{published data only}*

Ammendola M, Bottini N, Pietropolli A, Saccucci P, Gloria-Bottini F. Association between PTPN22 and endometriosis. *Fertility and Sterility* 2008;**89**(4):993-4.

Anastasi 2013 *{published data only}*

Anastasi E, Granato T, Falzarano R, Storelli P, Ticino A, Frati L, et al. The use of HE4, CA125 and CA72-4 biomarkers for differential diagnosis between ovarian endometrioma and epithelial ovarian cancer. *Journal of Ovarian Research* 2013;**6**(1):44.

Andrade 2010 *{published data only}*

Andrade AZd, Rodrigues JK, Dib LA, Romao GS, Ferriani RA, Jordao Junior AA, et al. Serum markers of oxidative stress in infertile women with endometriosis. *Revista Brasileira de Ginecologia e Obstetricia* 2010;**32**(6):279-85.

Andrisani 2014 *{published data only}*

Andrisani A, Dona G, Brunati AM, Clari G, Armanini D, Ragazzi E, et al. Increased oxidation-related glutathionylation and carbonic anhydrase activity in endometriosis. *Reproductive Biomedicine Online* 2014;**28**(6):773-9.

Antsiferova 2005 *{published data only}*

Antsiferova YS, Sotnikova NY, Posiseeva LV, Shor AL. Changes in the T-helper cytokine profile and in lymphocyte activation at the systemic and local levels in women with endometriosis. *Fertility and Sterility* 2005;**84**(6):1705-11.

Arjona Berral 1996 *{published data only}*

Arjona Berral JE, Contreras Puertas PI, Torres Avisbal M, Pacheco Capote C, Vallejo Casas JA, Benitez Velasco A, et al. Serum CA 125 and CA 19.9 levels in the diagnosis of patients with suspected endometriosis [Niveles sericos de CA-125 y CA 19.9 en el diagnostico de pacientes con sospecha clinica de endometriosis]. *Revista Espanola de Medicina Nuclear* 1996;**15**(2):71-6.

Avcioglu 2014 *{published data only}*

Avcioglu SN, Altinkaya SO, Kucuk M, Demircan-Sezer S, Yuksel H. Can platelet indices be new biomarkers for severe endometriosis?. ISRN Obstetrics and Gynecology 2014 Mar 26 [Epub ahead of print].

Ayers 1987 *{published data only}*

Ayers JW, Birenbaum DL, Menon KM. Luteal phase dysfunction in endometriosis: elevated progesterone levels in peripheral and ovarian veins during the follicular phase. *Fertility and Sterility* 1987;**47**(6):925-9.

Badawy 1984 *{published data only}*

Badawy SZ, Cuenca V, Stitzel A, Jacobs RD, Tomar RH. Autoimmune phenomena in infertile patients with endometriosis. *Obstetrics and Gynecology* 1984;**63**(3):271-5.

Badawy 1987 *{published data only}*

Badawy SZ, Cuenca V, Stitzel A, Tice D. Immune rosettes of T and B lymphocytes in infertile women with endometriosis. *Journal of Reproductive Medicine* 1987;**32**(3):194-7.

Badawy 1990 *{published data only}*

Badawy SZA, Cuenca V, Freliech H, Stefanu C. Endometrial antibodies in serum and peritoneal fluid of infertile patients with and without endometriosis. *Fertility and Sterility* 1990;**53**(5):930-2.

Balasch 1985 *{published data only}*

Balasch J, Vanrell JA. Mild endometriosis and luteal function. *International Journal of Fertility* 1985;**30**(3):4-6.

Barbieri 1986 *{published data only}*

Barbieri RL, Niloff JM, Bast RC Jr, Scaetzl E, Kistner RW, Knapp RC. Elevated serum concentrations of CA-125 in patients with advanced endometriosis. *Fertility and Sterility* 1986;**45**(5):630-4.

Barbieri 1987 *{published data only}*

Barbieri RL. CA-125 and endometriosis. *Contributions to Gynecology and Obstetrics* 1987;**16**:103-8.

Barrier 2002 *{published data only}*

Barrier F, Sharpe-Timms KL. Expression of soluble adhesion molecules in sera of women with stage III and IV endometriosis. *Journal of the Society for Gynecologic Investigation* 2002;**9**(2):98-101.

Basta 2009 *{published data only}*

Basta P, Mach P, Pitynski K, Bednarek W, Klimek M, Zietek J, et al. Differences in the blood serum levels of soluble HLA-G concentrations between the menstrual cycle phases and menopause in patients with ovarian endometriosis and uterine leiomyoma. *Neuroendocrinology Letters* 2009;**30**(1):91-8.

Bedaiwy 2006 *{published data only}*

Bedaiwy MA, Falcone T, Mascha EJ, Casper RF. Genetic polymorphism in the fibrinolytic system and endometriosis. *Obstetrics and Gynecology* 2006;**108**(1):162-8.

Berkes 2013 *{published data only}*

Berkes E, Muzinic A, Rigo Jr J, Tinneberg HR, Oehmke F. The analysis of the human plasma N-glycome in endometriosis patients. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2013;**171**(1):107-15.

Bianchi 2003 *{published data only}*

Bianchi M, Macaya R, Durruty G, Manzur A. Correlation between CA-125 marker with the presence and severity of pelvic endometriosis. *Revista Medica de Chile* 2003;**131**(4):367-72.

Bohler 2007 *{published data only}*

Bohler HC, Gercel-Taylor C, Lessey BA, Taylor DD. Endometriosis markers: immunologic alterations as diagnostic indicators for endometriosis. *Reproductive Sciences* 2007;**14**(6):595-604.

Bordin 2010 *{published data only}*

Bordin L, Fiore C, Dona G, Andrisani A, Ambrosini G, Faggian D, et al. Evaluation of erythrocyte band 3 phosphotyrosine level, glutathione content, CA-125, and human epididymal secretory protein E4 as combined parameters in endometriosis. *Fertility and Sterility* 2010;**94**(5):1616-21.

Bourlev 2006a *{published data only}*

Bourlev V, Larsson A, Olovsson M. Elevated levels of fibroblast growth factor-2 in serum from women with endometriosis. *American Journal of Obstetrics and Gynecology* 2006;**194**(3):755-9.

Bourlev 2006b *{published data only}*

Bourlev V, Volkov N, Pavlovitch S, Lets N, Larsson A, Olovsson M. The relationship between microvessel density, proliferative activity and expression of vascular endothelial growth factor-A and its receptors in eutopic endometrium and endometriotic lesions. *Reproduction* 2006;**132**(3):501-9.

Bragatto 2013 *{published data only}*

Bragatto FB, Barbosa CP, Christofolini DM, Peluso C, dos Santos AA, Mafra FA, et al. There is no relationship between Paraoxonase serum level activity in women with endometriosis and the stage of the disease: an observational study. *Reproductive Health* 2013;**10**:32.

Brinton 1996 *{published data only}*

Brinton DA, Quattrociocchi-Longe TM, Kiechle FL. Endometriosis: identification by carbonic anhydrase autoantibodies and clinical features. *Annals of Clinical and Laboratory Science* 1996;**26**(5):409-20.

Brosens 1978 *{published data only}*

Brosens IA, Koninckx PR, Corveleyn PA. A study of plasma progesterone, oestradiol-17beta, prolactin and LH levels, and of the luteal phase appearance of the ovaries in patients with endometriosis and infertility. *British Journal of Obstetrics and Gynaecology* 1978;**85**(4):246-50.

Cai 2005 *{published data only}*

Cai ZH, He YL, Peng DX. Changes of serum epithelial neutrophil-activating peptide-78 in patients with endometriosis. *Di 1 Junyi Daxue Xuebao [Academic Journal of the First Medical College of PLA]* 2005;**25**(4):464-5. Chinese.

Carmona 2012 *{published data only}*

Carmona F, Chapron C, Martinez-Zamora M-A, Santulli P, Rabanal A, Martinez-Florensa M, et al. Ovarian endometrioma but not deep infiltrating endometriosis is associated with increased serum levels of interleukin-8 and interleukin-6. *Journal of Reproductive Immunology* 2012;**95**(1-2):80-6.

Cheng 2002 *{published data only}*

Cheng YM, Wang ST, Chou CY. Serum CA-125 in preoperative patients at high risk for endometriosis. *Obstetrics and Gynecology* 2002;**99**(3):375-80.

Chihal 1986 *{published data only}*

Chihal HJ, Mathur S, Holtz GL, Williamson HO. An endometrial antibody assay in the clinical diagnosis and management of endometriosis. *Fertility and Sterility* 1986;**46**:408-11.

Cho 2008 *{published data only}*

Cho SH, Cho H, Nam A, Kim HY, Choi YS, Park KH, et al. Neutrophil-to-lymphocyte ratio as an adjunct to CA-125 for the diagnosis of endometriosis. *Fertility and Sterility* 2008;**90**(6):2073-9.

Cho 2009 *{published data only}*

Cho S, Ahn YS, Choi YS, Seo SK, Nam A, Kim HY, et al. Endometrial Osteopontin mRNA Expression and Plasma Osteopontin Levels are Increased in Patients with Endometriosis. *American Journal of Reproductive Immunology* 2009;**61**(4):286-93.

Cho 2012 *{published data only}*

Cho S, Choi YS, Jeon YE, Im KJ, Choi YM, Yim SY, et al. Expression of vascular endothelial growth factor (VEGF) and its soluble receptor-1 in endometriosis. *Microvascular Research* 2012;**83**(2):237-42.

Chrobak 2004 *{published data only}*

Chrobak A, Gmyrek GB, Sozanski R, Sieradzka U, Paprocka M, Gabrys M, et al. The influence of extracellular matrix proteins of T-cell proliferation and apoptosis in women with endometriosis or uterine leiomyoma. *American Journal of Reproductive Immunology* 2004;**51**(2):123-9.

Chun 2012 *{published data only}*

Chun S, Kim H, Ku S-Y, Suh CS, Kim SH, Kim JG. The association between endometriosis and polymorphisms in the interleukin-1 family genes in Korean women. *American Journal of Reproductive Immunology* 2012;**68**(2):154-63.

Colacurci 1996b *{published data only}*

Colacurci N, Fortunato N, DeFranciscis P, Fratta M, Cioffi M, Zarcone R, et al. Serum and peritoneal CA-125 levels as diagnostic test for endometriosis. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 1996;**661**:41-3.

Confino 1990 *{published data only}*

Confino E, Harlow L, Gleicher N. Peritoneal fluid and serum autoantibody levels in patients with endometriosis. *Fertility and Sterility* 1990;**53**:242-5.

Cunha-Filho 2001 *{published data only}*

Cunha-Filho JS, Gross JL, Lemos NA, Brandelli A, Castillos M, Passos EP. Hyperprolactinemia and luteal insufficiency in infertile patients with mild and minimal endometriosis. *Hormone and Metabolic Research* 2001;**33**(4):216-20.

D'Amico 2013 *{published data only}*

D'Amico F, Skarmoutsou E, Quaderno G, Malaponte G, La Corte C, Scibilia G, et al. Expression and localisation of osteopontin and prominin-1 (CD133) in patients with endometriosis. *International Journal of Molecular Medicine* 2013;**31**(5):1011-6.

D'Cruz 1996 *{published data only}*

D'Cruz OJ, Wild RA, Haas GG Jr, Reichlin M. Antibodies to carbonic anhydrase in endometriosis: prevalence, specificity, and relationship to clinical and laboratory parameters. *Fertility and Sterility* 1996;**66**(4):547-56.

Darai 2003 *{published data only}*

Darai E, Detchev R, Hugol D, Quang NT. Serum and cyst fluid levels of interleukin (IL) -6, IL-8 and tumour necrosis factor-alpha in women with endometriomas and benign and malignant cystic ovarian tumours. *Human Reproduction* 2003;**18**:1681-5.

Dawood 1988 *{published data only}*

Dawood MY, Khan-Dawood FS, Ramos J. Plasma and peritoneal fluid levels of CA 125 in women with endometriosis. *American Journal of Obstetrics and Gynecology* 1988;**159**:1526-31.

De Sanctis 2011 *{published data only}*

De Sanctis P, Elmakky A, Farina A, Caramelli E, Seracchioli R, Mabrouk M, et al. Matrix metalloproteinase-3 mRNA: a promising peripheral blood marker for diagnosis of endometriosis. Gynecologic and Obstetric Investigation 2011; Vol. 71, issue 2:118-23.

Dias 2006 *{published data only}*

Dias Jr JA, de Oliveira RM, Abrao MS. Antinuclear antibodies and endometriosis. *International Journal of Gynecology and Obstetrics* 2006;**93**(3):262-3.

Dias 2012 *{published data only}*

Dias JA, Podgaec S, de Oliveira RM, Marin MLC, Baracat EC, Abrao MS. Patients with endometriosis of the rectosigmoid have a higher percentage of natural killer cells in peripheral blood. *Journal of Minimally Invasive Gynecology* 2012;**19**(3):317-24.

Di Stefano 1994 *{published data only}*

Di Stefano G, Provinciali M, Muzzioli M, Garzetti GG, Ciavattini A, Fabris N. Correlation between estradiol serum levels and NK cell activity in endometriosis. *Annals of the New York Academy of Sciences* 1994;**741**:197-203.

Dogan 2006 *{published data only}*

Dogan S, Agic A, Frenzel W, Finas D, Diedrich K, Hornung D. Diagnostic tests for endometriosis. New molecular biological investigations. *Gynakologische Endokrinologie* 2006;**4**(3):128-32.

Dutta 2012 *{published data only}*

Dutta M, Joshi M, Srivastava S, Lodh I, Chakravarty B, Chaudhury K. A metabonomics approach as a means for identification of potential biomarkers for early diagnosis of endometriosis. *Molecular Biosystems* 2012;**8**(12):3281-7.

Dutta 2015 *{published data only}*

Dutta M, Subramani E, Taunk K, Gajbhiye A, Seal S, Pendharkar N, et al. Investigation of serum proteome alterations in human endometriosis. *Journal of Proteomics* 2015;**114**:182-96.

Ejzenberg 2013 *{published data only}*

Ejzenberg D, Podgaec S, Dias Jr JA, de Oliveira RM, Baracat EC, Abrao MS. Measurement of serum and peritoneal levels of amyloid protein A and their importance in the diagnosis of pelvic endometriosis. *Journal of Reproductive Medicine* 2013;**58**(9-10):411-6.

Fallat 1997 *{published data only}*

Fallat ME, Siow Y, Marra M, Cook C, Carrillo A. Mullerian-inhibiting substance in follicular fluid and serum: a comparison of patients with tubal factor infertility, polycystic ovary syndrome, and endometriosis. *Fertility and Sterility* 1997;**67**(5):962-5.

Fedele 1988 *{published data only}*

Fedele L, Vercellini P, Arcaini L, Da Dalt MZ, Candiani GB. CA 125 in serum, peritoneal fluid, active lesions, and endometrium of patients with endometriosis. *American Journal of Obstetrics and Gynecology* 1988;**158**(1):166-70.

Fernandez-Shaw 1993 *{published data only}*

Fernandez-Shaw S, Hicks BR, Yudkin PL, Kennedy S, Barlow YH, Starkey PM. Anti-endometrial and anti-endothelial auto-antibodies in women with endometriosis. *Human Reproduction* 1993;**8**(2):310-5.

Fernandez-Shaw 1996 *{published data only}*

Fernandez-Shaw S, Kennedy SH, Hicks BR, Edmonds K, Starkey PM, Barlow DH. Anti-endometrial antibodies in women measured by an enzyme-linked immunosorbent assay. *Human Reproduction* 1996;**11**(6):1180-4.

Ferrero 2005b *{published data only}*

Ferrero S, Gillott DJ, Remorgida V, Anserini P, Price K, Ragni N, et al. Haptoglobin beta chain isoforms in the plasma and peritoneal fluid of women with endometriosis. *Fertility and Sterility* 2005;**83**(5):1536-43.

Fisk 1988 *{published data only}*

Fisk NM, Tan CE. CA 125 in peritoneal fluid and serum of patients with endometriosis. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 1988;**29**(2):153-8.

Flores 2006 *{published data only}*

Flores I, Rivera E, Mousses S, Chen YD, Rozenblum E. Identification of molecular markers for endometriosis in blood lymphocytes by using deoxyribonucleic acid microarrays. *Fertility and Sterility* 2006;**85**(6):1676-83.

Fu 2002 *{published data only}*

Fu C, Lang J. Serum soluble E-cadherin level in patients with endometriosis. *Chinese Medical Sciences Journal* 2002;**17**(2):121-3.

Fujii 2008 *{published data only}*

Fujii EY, Nakayama M, Nakagawa A. Concentrations of receptor for advanced glycation end products, VEGF and CML in plasma, follicular fluid, and peritoneal fluid in women with and without endometriosis. *Reproductive Sciences* 2008;**15**(10):1066-74.

Gagne 2003c *{published data only}*

Gagne D, Rivard M, Page M, Shazand K, Hugo P, Gosselin D, et al. Blood leukocyte subsets are modulated in patients with endometriosis. Fertility and Sterility 2003; Vol. 80, issue 1:43-53.

Gajbhiye 2008 *{published data only}*

Gajbhiye R, Suryawanshi A, Khan S, Meherji P, Warty N, Raut V, et al. Multiple endometrial antigens are targeted in autoimmune endometriosis. *Reproductive Biomedicine Online* 2008;**16**(6):817-24.

Gajbhiye 2012 *{published data only}*

Gajbhiye R, Sonawani A, Khan S, Suryawanshi A, Kadam S, Warty N, et al. Identification and validation of novel serum markers for early diagnosis of endometriosis. *Human Reproduction* 2012;**27**(2):408-17.

Galleri 2009 *{published data only}*

Galleri L, Luisi S, Rotondi M, Romagnani P, Cobellis L, Serio M, et al. Low serum and peritoneal fluid concentration of interferon - induced protein-10 (CXCL10) in women with endometriosis. *Fertility and Sterility* 2009;**91**(2):331-4.

Galo 2005 *{published data only}*

Galo S, Zubor P, Szunyogh N, Kajo K, Machalekova K, Biringer K, et al. TNF-alpha serum levels in women with endometriosis: Prospective clinical study. *Ceska Gynekologie* 2005;**70**(4):286-90.

Garcia-Manero 2007 *{published data only}*

Garcia-Manero M, Alcazar JL, Toledo G. Vascular endothelial growth factor (VEGF) and ovarian endometriosis: correlation between VEGF serum levels, VEGF cellular expression, and pelvic pain. *Fertility and Sterility* 2007;**88**(2):513-5.

Garcia-Velasco 2002 *{published data only}*

Garcia-Velasco JA, Mulayim N, Kayisli UA, Arici A. Elevated soluble Fas ligand levels may suggest a role for apoptosis in women with endometriosis. *Fertility and Sterility* 2002;**78**(4):855-9.

Garza 1991 *{published data only}*

Garza D, Mathur S, Dowd MM, Smith LF, Williamson HO. Antigenic differences between the endometrium of women with and without endometriosis. *Journal of Reproductive Medicine* 1991;**36**(3):177-82.

Garzetti 1994 *{published data only}*

Garzetti GG, Ciavattini A, Tranquilli AL, Arduini D, Romanini C. Serum Ca-125 Concentration in Endometriosis Patients - Role of Pelvic and Peritoneal Irritation. *Gynecological Endocrinology* 1994;**8**(1):27-31.

Gebel 1993 *{published data only}*

Gebel HM, Braun DP, Rotman C, Rana N, Dmowski WP. Mitogen induced production of polyclonal IgG is decreased in women with severe endometriosis. *American Journal of Reproductive Immunology* 1993;**29**(2):124-30.

Gebel 1995 *{published data only}*

Gebel HM, Rana N, Braun DP, Dmowski WP. Differential expression of VLA beta 1 (CD29) on monocytes from patients

with endometriosis. *American Journal of Reproductive Immunology* 1995;**34**(5):317-22.

Giudice 1986 *{published data only}*

Giudice LC, Jacobs A, Pineda J. Serum levels of CA-125 in patients with endometriosis: A preliminary report. *Fertility and Sterility* 1986;**45**(6):876-8.

Gmyrek 2005 *{published data only}*

Gmyrek GB, Sozanski R, Jerzak M, Chrobak A, Wickiewicz D, Skupnik A, et al. Evaluation of monocyte chemotactic protein-1 levels in peripheral blood of infertile women with endometriosis. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2005;**122**(2):199-205.

Gorski 2007 *{published data only}*

Gorski J, Szyllo K, Banasik M, Lewkowicz P, Tchorzewski H. CD4+, CD8+ and CD4+CD25+ T lymphocytes in peripheral blood and peritoneal fluid of women with endometriosis - Preliminary report. *Archives of Medical Science* 2007;**3**(1):37-42.

Guerriero 1997 *{published data only}*

Guerriero S, Mallarini G, Ajossa S, Risalvato A, Satta R, Mais V, et al. Transvaginal ultrasound and computed tomography combined with clinical parameters and CA-125 determinations in the differential diagnosis of persistent ovarian cysts in premenopausal women. *Ultrasound in Obstetrics & Gynecology* 1997;**9**(5):339-43.

Gunev 1981 *{published data only}*

Gunev V, Maleeva A, Spasov S, Melamed V, Trepechov S. FSH, LH, estradiol and testosterone studies of the blood serum in endometriosis externa. *Akusherstvo i Ginekologiia* 1981;**20**(4):312-5.

Hammadeh 2003 *{published data only}*

Hammadeh ME, Fischer-Hammadeh C, Hoffmeister H, Huebner U, Georg T, Rosenbaum P, et al. Fibroblast growth factor (FGF), intracellular adhesion molecule (sICAM-1) level in serum and follicular fluid of infertile women with polycystic ovarian syndrome, endometriosis and tubal damage, and their effect on ICSI outcome. *American Journal of Reproductive Immunology* 2003;**50**(2):124-30.

Han 2009 *{published data only}*

Han YJ, Kim HN, Yoon JK, Yi SY, Moon HS, Ahn JJ, et al. Haplotype analysis of the matrix metalloproteinase-9 gene associated with advanced-stage endometriosis. *Fertility and Sterility* 2009;**91**(6):2324-30.

Hatayama 1996 *{published data only}*

Hatayama H, Imai K, Kanzaki H, Higuchi T, Fujimoto M, Mori T. Detection of antiendometrial antibodies in patients with endometriosis by cell ELISA. *American Journal of Reproductive Immunology* 1996;**35**(2):118-22.

He 1993 *{published data only}*

He YE. Prolactin secretion in patients with endometriosis and its relationship to luteal phase defect and infertility. *Chung-Hua Fu Chan Ko Tsa Chih [Chinese Journal of Obstetrics and Gynecology]* 1993;**28**(1):14-7. Chinese.

Hompes 1996 *{published data only}*

Hompes PGA, Koninckx PR, Kennedy S, Van Kamp GJ, Verstraeten RA, Cornillie F. Serum CA-125 concentrations during midfollicular phase, a clinically useful and reproducible marker in diagnosis of advanced endometriosis. *Clinical Chemistry* 1996;**42**(11):1871-4.

Hornstein 1992 *{published data only}*

Hornstein MD, Thomas PP, Gleason RE, Barbieri RL. Menstrual cyclicity of CA-125 in patients with endometriosis. *Fertility and Sterility* 1992;**58**(2):279-83.

Hrycek 1996 *{published data only}*

Hrycek A, Kalina Z, Cuzytek A. Selected immunologic markers for evaluation of peripheral blood in patients with internal endometriosis. *Wiadomosci lekarskie* 1996;**49**(1-6):10-4.

Hsu 1997 *{published data only}*

Hsu CC, Yang BC, Wu MH, Huang KE. Enhanced interleukin-4 expression in patients with endometriosis. *Fertility and Sterility* 1997;**67**(6):1059-64.

Hsu 2014 *{published data only}*

Hsu CY, Hsieh TH, Tsai CF, Tsai HP, Chen HS, Chang Y, et al. miRNA-199a-5p regulates VEGFA in endometrial mesenchymal stem cells and contributes to the pathogenesis of endometriosis. *Journal of Pathology* 2014;**232**(3):330-43.

Huang 2004 *{published data only}*

Huang HF, Hong LH, Tan Y, Sheng JZ. Matrix metalloproteinase 2 is associated with changes in steroid hormones in the sera and peritoneal fluid of patients with endometriosis. *Fertility and Sterility* 2004;**81**(5):1235-9.

Hwang 2014 *{published data only}*

Hwang JH, Lee KS, Joo JK, Wang T, Son JB, Park JHA, et al. Identification of biomarkers for endometriosis in plasma from patients with endometriosis using a proteomics approach. *Molecular Medicine Reports* 2014;**10**(2):725-30.

Ihlenfeld 2007 *{published data only}*

Ihlenfeld MFK. Determination of cytokines in laboratory diagnosis of endometriosis peritoneal minimum and light [Determinação de citocinas no diagnóstico laboratorial da endometriose peritoneal mínima e leve]. Tese apresentada ao Programa de Pós-Graduação em Ginecologia, Obstetrícia e Mastologia da Faculdade de Medicinade Botucatu – Universidade Estadual Paulista "Júlio de Mesquita Filho" 2007:130.

Illera 2001 *{published data only}*

Illera JC, Silvan G, Illera MJ, Munro CJ, Lessey BA, Illera M. Measurement of serum and peritoneal fluid LH concentrations as a diagnostic tool for human endometriosis. *Reproduction* 2001;**121**(5):761-9.

Izumiya 2003 *{published data only}*

Izumiya C, Maeda N, Kusume T, Masumoto T, Yamashita C, Yamamoto Y, et al. Coordinated but depressed expression of human leukocyte antigen-DR, intercellular adhesion molecule-1, and CD14 on peritoneal macrophages in women

with pelvic endometriosis. *Fertility and Sterility* 2003;**80**(Suppl 2):768-75.

Jackson 2005 *{published data only}*

Jackson LW, Schisterman EF, Dey-Rao R, Browne R, Armstrong D. Oxidative stress and endometriosis. *Human Reproduction* 2005;**20**(7):2014-20.

Jana 2013 *{published data only}*

Jana SK, Dutta M, Joshi M, Srivastava S, Chakravarty B, Chaudhury K. 1H NMR based targeted metabolite profiling for understanding the complex relationship connecting oxidative stress with endometriosis. BioMed Research International 2013 Aug 5 [Epub ahead of print].

Jedryka 2001 *{published data only}*

Jedryka M, Goluda M, Kuliczkowski K, Sozanski L. E-cadherin in the serum and the peritoneal fluid of women with endometriosis. *Ginekologia Polska* 2001;**72**(5):418-21.

Jerzak 2002 *{published data only}*

Jerzak M, Baranowski W, Rechberger T, Gorski A. Enhanced T cells interactions with extracellular matrix proteins in infertile women with endometriosis. *Immunology Letters* 2002;**81**(1):65-70.

Jing 2009 *{published data only}*

Jing JH, Qiao YH, Suginami H, Taniguchi F, Shi HR, Wang XL. Two novel serum biomarkers for endometriosis screened by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry and their change after laparoscopic removal of endometriosis. *Fertility and Sterility* 2009;**92**(4):1221-7.

Kabut 2007 *{published data only}*

Kabut J, Kondera-Anasz Z, Sikora J, Mielczarek-Palacz A. Levels of complement components iC3b, C3c, C4, and SC5b-9 in peritoneal fluid and serum of infertile women with endometriosis. *Fertility and Sterility* 2007;**88**(5):1298-303.

Kadija 2012 *{published data only}*

Kadija S, Stefanovic A, Jeremic K, Radojevic MM, Nikolic L, Markovic I, et al. The utility of human epididymal protein 4, cancer antigen 125, and risk for malignancy algorithm in ovarian cancer and endometriosis. *International Journal of Gynecological Cancer* 2012;**22**(2):238-44.

Kafali 2004 *{published data only}*

Kafali H, Artuc H, Demir N. Use of CA125 fluctuation during the menstrual cycle as a tool in the clinical diagnosis of endometriosis; a preliminary report. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2004;**116**(1):85-8.

Kang 1988 *{published data only}*

Kang JO, Hudak WA, Keller N, Criswell BS. Enzyme-linked immunosorbent-assay of Ca-125 in serum of patients with endometriosis - efficacy in diagnosis. *Clinical Chemistry* 1988;**34**(10):1983-6.

Kataoka 2012 *{published data only}*

Kataoka T, Watanabe Y, Hoshiai H. Retrospective evaluation of tumor markers in ovarian mature cystic teratoma and ovarian endometrioma. *Journal of Obstetrics and Gynaecology Research* 2012;**38**(8):1071-6.

Kharfi 2002 *{published data only}*

Kharfi A, Akoum A. Soluble interleukin-1 receptor type II blocks monocyte chemotactic protein-1 secretion by U937 cells in response to peripheral blood serum of women with endometriosis. *Fertility and Sterility* 2002;**78**(4):836-42.

KhoshdelRad 2014 *{published data only}*

KhoshdelRad N, Salehi Z, Mashayekhi F, Abbasi O, Mirzajani E. Soluble c-Met expression in the peritoneal fluid and serum of patients with different stages of endometriosis. *Archives of Gynecology and Obstetrics* 2014;**289**(5):1107-12.

Kichuchi 1993 *{published data only}*

Kichuchi Y, Ishikawa N, Hirata J, Imaizumi E, Sasa H, Nagata I. Changes of peripheral blood lymphocyte subsets before and after operation of patients with endometriosis. *Acta Obstetricia et Gynecologica Scandinavica* 1993;**72**(3):157-61.

Kiechle 1994 *{published data only}*

Kiechle FL, Quattrociocchi-Longe TM, Brinton DA. Carbonic anhydrase antibody in sera from patients with endometriosis. *American Journal of Clinical Pathology* 1994;**101**(5):611-5.

Kilpatrick 1991 *{published data only}*

Kilpatrick DC, Haining REB, Smith SSK. Are cardiolipin antibody levels elevated in endometriosis?. *Fertility and Sterility* 1991;**55**(2):436-7.

Kim 1995 *{published data only}*

Kim JG, Kim CW, Moon SY, Chang YS, Lee JY. Detection of antiendometrial antibodies in sera of patients with endometriosis by dual-colored, double-labeling immunohistochemical method and western blot. *American Journal of Reproductive Immunology* 1995;**34**(2):80-7.

Kim 2007 *{published data only}*

Kim CM, Oh YJ, Cho SH, Chung DJ, Hwang JY, Park KH, et al. Increased telomerase activity and human telomerase reverse transcriptase mRNA expression in the endometrium of patients with endometriosis. *Human Reproduction* 2007;**22**(3):843-9.

Kim 2014 *{published data only}*

Kim SK, Park JY, Jee BC, Suh CS, Kim SH. Association of the neutrophil-to-lymphocyte ratio and CA 125 with the endometriosis score. *Clinical and Experimental Reproductive Medicine* 2014;**41**(4):151-7.

Kinugasa 2011 *{published data only}*

Kinugasa S, Shinohara K, Wakatsuki A. Increased asymmetric dimethylarginine and enhanced inflammation are associated with impaired vascular reactivity in women with endometriosis. *Atherosclerosis* 2011;**219**(2):784-8.

Kobayashi 1987 *{published data only}*

Kobayashi H, Kanayama N, Hayata T, Kawashima Y. Usefulness of measurement of serum CA125 levels in diagnosing and treating endometriosis. *Nippon Sanka Fujinka Gakkai Zasshi - Acta Obstetrica et Gynaecologica Japonica* 1987;**39**(7):1054-60.

Kondera-Anasz 2004 *{published data only}*

Kondera-Anasz Z, Sikora J, Mertas A, Micinski P, Bednarz B. Concentrations of neopterin and interleukin-10 in peritoneal fluid and in serum of women with endometriosis. *Pteridines* 2004;**15**(1):20-7.

Kondera-Anasz 2005 *{published data only}*

Kondera-Anasz Z, Sikora J, Mielczarek-Palacz A, Jonca M. Concentrations of interleukin (IL)-1alpha, IL-1 soluble receptor type II (IL-1 sRII) and IL-1 receptor antagonist (IL-1 Ra) in the peritoneal fluid and serum of infertile women with endometriosis. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2005;**123**(2):198-203.

Koninckx 1992 *{published data only}*

Koninckx PR, Riittinen L, Seppala M, Cornillie FJ. CA-125 and placental protein 14 concentrations in plasma and peritoneal fluid of women with deeply infiltrating pelvic endometriosis. *Fertility and Sterility* 1992;**57**(3):523-30.

Kopuz 2014 *{published data only}*

Kopuz A, Kurt S, Demirtas O, Toz E, Tasyurt A. Relation of peritoneal fluid and serum vascular endothelial growth factor levels to endometriosis stage. *Clinical and Experimental Obstetrics & Gynecology* 2014;**41**(5):547-50.

Koumantakis 1994 *{published data only}*

Koumantakis E, Matalliotakis I, Neonaki M, Froudarakis G, Georgoulias V. Soluble serum interleukin-2 receptor, interleukin-6 and interleukin-1a in patients with endometriosis and in controls. *Archives of Gynecology and Obstetrics* 1994;**255**(3):107-12.

Kralickova 2007 *{published data only}*

Kralickova M, Ulcova-Gallova Z, Sima R, Vanecek T, Sima P, Krizan J, et al. Association of the leukemia inhibitory factor gene mutation and the antiphospholipid antibodies in the peripheral blood of infertile women. *Folia Microbiologica* 2007;**52**(5):543-8.

Krasnicki 2001 *{published data only}*

Krasnicki D. Serum and peritoneal fluid CA-125 concentration in women with endometriosis. *Ginekologia Polska* 2001;**72**:1365-9.

Kurt 2014 *{published data only}*

Kurt RK, Dogan AC, Yesilyurt H, Karateke A, Okyay AG. Relation of red cell distribution width to the presence and severity of endometriosis. *Clinical and Experimental Obstetrics & Gynecology* 2014;**41**(6):713-6.

Lambert 2014 *{published data only}*

Lambert S, Santulli P, Chouzenoux S, Marcellin L, Borghese B, De Ziegler D, et al. Endometriosis: Increasing concentrations of serum interleukin-1beta and interleukin-1sRII is associated with the deep form of this pathology [Endometriose: L'augmentation des concentrations d'interleukine-1beta et d'interleukine-1sRII dans le serum est associee a la forme profonde de cette pathologie]. *Journal de Gynecologie Obstetrique et Biologie de la Reproduction* 2014;**43**(9):735-43.

Lang 2001 *{published data only}*

Lang GA, Yeaman GR. Autoantibodies in endometriosis sera recognize a Thomsen-Friedenreich-like carbohydrate antigen. *Journal of Autoimmunity* 2001;**16**(2):151-61.

Lee 2014 *{published data only}*

Lee YH, Tan CW, Venkatratnam A, Tan CS, Cui L, Loh SF, et al. Dysregulated sphingolipid metabolism in endometriosis. *Journal of Clinical Endocrinology and Metabolism* 2014;**99**(10):E1913-E21.

Leggieri 2014 *{published data only}*

Leggieri C, D'Agostino G, Tommasi L, Plebani M, Conte L. Is HE4 a useful endometrioma marker?. *European Journal of Gynaecological Oncology* 2014;**35**(4):438-41.

Leng 2002 *{published data only}*

Leng J, Lang J, Zhao D, Liu D. Serum levels of soluble intercellular molecule 1 (sICAM-1) in endometriosis. *Zhonghua Yixue Zazhi [Chinese Medical Journal]* 2002;**82**(3):189-90. Chinese.

Lenhard 2011 *{published data only}*

Lenhard M, Stieber P, Hertlein L, Kirschenhofer A, Furst S, Mayr D, et al. The diagnostic accuracy of two human epididymis protein 4 (HE4) testing systems in combination with CA125 in the differential diagnosis of ovarian masses. *Clinical Chemistry and Laboratory Medicine* 2011;**49**(12):2081-8.

Lermann 2010 *{published data only}*

Lermann J, Mueller A, Korber F, Oppelt P, Beckmann MW, Dittrich R, et al. Evaluation of high-sensitivity C-reactive protein in comparison with C-reactive protein as biochemical serum markers in women with endometriosis. *Fertility and Sterility* 2010;**93**(7):2125-9.

Li 2000 *{published data only}*

Li B, Jin F, Yang L. Evaluation of tumor necrosis factor-alpha and interleukin-6 levels in serum and peritoneal fluid of patients with endometriosis. *Zhonghua Fuchanke Zazhi [Chinese Journal of Obstetrics and Gynecology]* 2000;**35**(3):166-8. Chinese.

Li 2010 *{published data only}*

Li LA, Wang N, Liang TT, Li AL, Li YL. Establishment of serum peptidome pattern-based diagnostic model of endometriosis. *Nanfang Yeie Daxue Xueba [Journal of Southern Medical University]* 2010;**30**(4):851-4. Chinese.

Linghu 2004 *{published data only}*

Linghu H, Xu X, Luo J, Zhuang L. Changes of soluble fas and soluble fas ligand in serum and peritoneal fluid of infertile patients with endometriosis. *Chinese Medical Sciences Journal* 2004;**19**(1):56-9.

Liu 2007 *{published data only}*

Liu HY, Lang JH, Zhou QF, Shan D, Li QZ. Detection of endometriosis with the use of plasma protein profiling by surf ace-enhanced laser desorption/ionization time-of-flight mass spectrometry. *Fertility and Sterility* 2007;**87**(4):988-90.

Liu 2013 *{published data only}*

Liu F, He L, Liu Y, Shi Y, Du H. The expression and role of oxidative stress markers in the serum and follicular fluid of patients with endometriosis. *Clinical and Experimental Obstetrics & Gynecology* 2013;**40**(3):372-6.

Long 2013 *{published data only}*

Long X, Jiang P, Zhou L, Zhang W. Evaluation of novel serum biomarkers and the proteomic differences of endometriosis and adenomyosis using MALDI-TOF-MS. *Archives of Gynecology and Obstetrics* 2013;**288**(1):201-5.

Luo 2005 *{published data only}*

Luo, M, He YL, Peng DX, Lui MB, Chen YY. Evaluation of tumor necrosis factor-alpha and tumor necrosis factor-beta in serum of patients with endometriosis. *Zhongnan Daxue Xuebao (Yixue Ban) [Journal of Central South University (Medical Sciences)]* 2005;**30**(3):304-6. Chinese.

Maeda 2004 *{published data only}*

Maeda N, Izumiya C, Kusum T, Masumoto T, Yamashita C, Yamamoto Y, et al. Killer inhibitory receptor CD158a overexpression among natural killer cells in women with endometriosis is undiminished by laparoscopic surgery and gonadotropin releasing hormone agonist treatment. *American Journal of Reproductive Immunology* 2004;**51**(5):364-72.

Mahdian 2015 *{published data only}*

Mahdian S, Aflatoonian R, Yazdi RS, Yaghmaei P, Ramazanali F, Afsharian P, et al. Macrophage migration inhibitory factor as a potential biomarker of endometriosis. *Fertility and Sterility* 2015;**103**(1):153-9.

Malvezzi 2013 *{published data only}*

Malvezzi H, Aguiar VG, De Paz CCP, Tanus-Santos JE, De Araujo Penna IA, Navarro PA. Increased circulating MMP-2 levels in infertile patients with moderate and severe pelvic endometriosis. *Reproductive Sciences* 2013;**20**(5):557-62.

Manero 2009 *{published data only}*

Manero MG, Olartecoechea B, Royo P, Alcazar JL. Thrombospondin-1 serum levels do not correlate with pelvic pain in patients with ovarian endometriosis. *Journal of Ovarian Research* 2009;**2**(1):18.

Manero 2010 *{published data only}*

Manero MG, Alcazar JL. Interleukin-8 serum levels do not correlate with pelvic pain in patients with ovarian endometriomas. *Fertility and Sterility* 2010;**94**(2):450-2.

Markham 1997b *{published data only}*

Markham R, Fraser IS, Song JY, Jansen RPS. The measurement of tumour necrosis factor alpha in patients with endometriosis. *Australian Journal of Medical Science* 1997;**18**(2):56-9.

Masahashi 1988 *{published data only}*

Masahashi T, Matsuzawa K, Ohsawa M, Narita O, Asai T, Ishihara M. Serum CA 125 levels in patients with endometriosis: changes in CA 125 levels during menstruation. *Obstetrics and Gynecology* 1988;**72**:328-31.

Matalliotakis 1994 *{published data only}*

Matalliotakis I, Makrigiannakis A, Karkavitsas N, Psaroudakis E, Froudarakis G, Koumantakis E. Use of CA-125 in the diagnosis and management of endometriosis: influence of treatment with danazol. *International Journal of Fertility and Menopausal Studies* 1994;**39**:100-4.

Matalliotakis 1997 *{published data only}*

Matalliotakis I, Neonaki M, Zolindaki A, Hassan E, Georgoulias V, Koumantakis E. Changes in immunologic variables (TNF-a, sCD8 and sCD4) during danazol treatment in patients with endometriosis. *International Journal of Fertility and Womens Medicine* 1997;**42**(3):211-4.

Matalliotakis 2000 *{published data only}*

Matalliotakis IM, Koumantaki YG, Neonaki MA, Goumenou AG, Koumantakis GE, Kyriakou DS, et al. Increase in serum leptin concentrations among women with endometriosis during danazol and leuprolide depot treatments. *American Journal of Obstetrics and Gynecology* 2000;**183**(1):58-62.

Matalliotakis 2001a *{published data only}*

Matalliotakis IM, Vassiliadis S, Goumenou AG, Athanassakis I, Koumantakis GE, Neonaki MA, et al. Soluble ICAM-1 levels in the serum of endometriotic patients appear to be independent of medical treatment. *Journal of Reproductive Immunology* 2001;**51**(1):9-19.

Matalliotakis 2001b *{published data only}*

Matalliotakis IM, Athanassakis I, Goumenou AG, Neonaki MA, Koumantakis EE, Vassiliadis S. The possible anti-inflammatory role of circulating human leukocyte antigen levels in women with endometriosis after treatment with danazol and leuprorelin acetate depot. *Mediators of Inflammation* 2001;**10**(2):75-80.

Matalliotakis 2003b *{published data only}*

Matalliotakis LM, Goumenou AG, Koumantakis GE, Athanassakis I, Dionyssopoulou E, Neonaki MA, et al. Expression of serum human leukocyte antigen and growth factor levels in a greek family with familial endometriosis. *Journal of the Society for Gynecologic Investigation* 2003;**10**(2):118-21.

Matarese 2000 *{published data only}*

Matarese G, Alviggi C, Sanna V, Howard JK, Lord GM, Carravetta C, et al. Increased leptin levels in serum and peritoneal fluid of patients with pelvic endometriosis. *Journal of Clinical Endocrinology and Metabolism* 2000;**85**(7):2483-7.

Mathur 1982 *{published data only}*

Mathur S, Peress MR, Williamson HO. Autoimmunity to endometrium and ovary in endometriosis. *Clinical and Experimental Immunology* 1982;**50**(2):259-66.

Mathur 1990 *{published data only}*

Mathur S, Garza DE, Smith LF. Endometrial autoantigens eliciting immunoglobulin (Ig) G, IgA, and IgM responses in endometriosis. *Fertility and Sterility* 1990;**54**(1):56-63.

Mathur 1998 *{published data only}*

Mathur SP, Holt VL, Lee JH, Jiang H, Rust PF. Levels of antibodies to transferrin and alpha 2-HS glycoprotein in women with and without endometriosis. *American Journal of Reproductive Immunology* 1998;**40**(2):69-73.

Mathur 1999 *{published data only}*

Mathur SP, Lee JH, Jiang H, Arnaud P, Rust PF. Levels of transferrin and alpha 2-HS glycoprotein in women with and without endometriosis. *Autoimmunity* 1999;**29**(2):121-7.

Mathur 2000 *{published data only}*

Mathur SP. Autoimmunity in endometriosis: relevance to infertility. *American Journal of Reproductive Immunology* 2000;**44**(2):89-95.

Matsuoka 2005 *{published data only}*

Matsuoka S, Maeda N, Izumiya C, Yamashita C, Nishimori Y, Fukaya T. Expression of inhibitory-motif killer immunoglobulin-like receptor, KIR2DL1, is increased in natural killer cells from women with pelvic endometriosis. *American Journal of Reproductive Immunology* 2005;**53**(5):249-54.

Medl 1997 *{published data only}*

Medl M, Ogris E, Peters-Engl C, Mierau M, Buxbaum P, Leodolter S. Serum levels of the tumour-associated trypsin inhibitor in patients with endometriosis. *British Journal of Obstetrics and Gynaecology* 1997;**104**(1):78-81.

Michaud 2014 *{published data only}*

Michaud N, Al-Akoum M, Akoum A. Blood soluble interleukin 1 receptor accessory protein levels are consistently low throughout the menstrual cycle of women with endometriosis. *Reproductive Biology and Endocrinology* 2014;**12**(1):1.

Moloney 1989 *{published data only}*

Moloney MD, Thornton JG, Cooper EH. Serum CA 125 antigen levels and disease severity in patients with endometriosis. *Obstetrics and Gynecology* 1989;**73**(1):767-9.

Moretuzzo 1988 *{published data only}*

Moretuzzo RW, DiLauro S, Jenison E, Chen SL, Reindollar RH, McDonough PG. Serum and peritoneal lavage fluid CA-125 levels in endometriosis. *Fertility and Sterility* 1988;**50**(3):430-3.

Nabeta 2009 *{published data only}*

Nabeta M, Abe Y, Kagawa L, Haraguchi R, Kito K, Ueda N, et al. Identification of anti--enolase autoantibody as a novel serum marker for endometriosis. *Proteomics Clinical Applications* 2009;**3**(10):1201-10.

Nabeta 2011 *{published data only}*

Nabeta M, Abe Y, Takaoka Y, Kusanagi Y, Ito M. Identification of anti-syntaxin 5 autoantibody as a novel serum marker of endometriosis. *Journal of Reproductive Immunology* 2011;**91**(1-2):48-55.

Nagamani 1992 *{published data only}*

Nagamani M, Kelver ME, Smith ER. CA 125 levels in monitoring therapy for endometriosis and in prediction of recurrence. *International Journal of Fertility* 1992;**37**(4):227-31.

Nalbanski 2008 *{published data only}*

Nalbanski A, Kiurkchiev D. The use of biochemical markers in diagnosis of endometriosis. *Akusherstvo i Ginekologiia* 2008;**47**(4):19-22.

Nomiyama 1997 *{published data only}*

Nomiyama M, Hachisuga T, Sou H, Nakamura K, Matsumoto Y, Iwasaka T, et al. Local immune response in infertile patients with minimal endometriosis. *Gynecologic and Obstetric Investigation* 1997;**44**(1):32-7.

O'Shaughnessy 1993 *{published data only}*

O'Shaughnessy A, Check JH, Nowroozi K, Lurie D. CA 125 levels measured in different phases of the menstrual cycle in screening for endometriosis. *Obstetrics and Gynecology* 1993;**81**(1):99-103.

Odukoya 1995a *{published data only}*

Odukoya OA, Bansal A, Wilson AP, Weetman AP, Cooke ID. Serum-soluble CD23 in patients with endometriosis and the effect of treatment with danazol and leuprolide acetate depot injection. *Human Reproduction* 1995;**10**(4):942-6.

Odukoya 1995b *{published data only}*

Odukoya OA, Wheatcroft N, Weetman AP, Cooke ID. The prevalence of endometrial immunoglobulin G antibodies in patients with endometriosis. *Human Reproduction* 1995;**10**(5):1214-9.

Ota 1990 *{published data only}*

Ota H, Maki M. Evaluation of autoantibody and CA125 in the diagnosis of endometriosis or adenomyosis. *Medical Science Research* 1990;**18**(8):309-10.

Ota 1991 *{published data only}*

Ota H. Autoantibody sensitivity in the diagnosis of endometriosis and adenomyosis. *Japanese Journal of Fertility and Sterility* 1991;**36**(4):799-804.

Ozaksit 1995 *{published data only}*

Ozaksit G, Caglar T, Cicek N, Kuscu E, Batioglu S, Gokmen O. Serum Ca-125 Levels before, during and after treatment for endometriosis. *International Journal of Gynecology and Obstetrics* 1995;**50**(3):269-73.

Ozasa 1987 *{published data only}*

Ozasa H, Noda Y, Mori T. Progesterone increases serum CA-125 in endometriosis. *Fertility and Sterility* 1987;**47**(4):699-701.

Perwira 2009 *{published data only}*

Perwira I. Comparison of TNF-alpha and IL-6 between the blood and peritoneal fluid in patients with endometriosis [Perbandingan nilai TNF-alpha dan IL-6 antara cairan peritoneal dan darah pada pasien endometriosis]. Thesis: Universitas Gadjah Mada, Yogyakarta 2009.

Pittaway 1986 *{published data only}*

Pittaway DE, Fayez JA. The use of CA-125 in the diagnosis and management of endometriosis. *Fertility and Sterility* 1986;**46**(5):790-5.

Pittaway 1987a *{published data only}*

Pittaway DE, Fayez JA, Douglas JW. Serum CA-125 in the evaluation of benign adnexal cysts. *American Journal of Obstetrics and Gynecology* 1987;**157**(6):1426-8.

Pittaway 1987b *{published data only}*

Pittaway DE, Fayez JA. Serum CA-125 antigen levels increase during menses. *American Journal of Obstetrics and Gynecology* 1987;**156**(1):75-6.

Pizzo 2002 *{published data only}*

Pizzo A, Salmeri FM, Ardita FV, Sofo V, Tripepi M, Marsico S. Behaviour of cytokine levels in serum and peritoneal fluid of women with endometriosis. *Gynecologic and Obstetric Investigation* 2002;**54**(2):82-7.

Podgaec 2010 *{published data only}*

Podgaec S, Dias Junior JA, Chapron C, Oliveira RMd, Baracat EC, Abrao MS. Th1 and Th2 ummune responses related to pelvic endometriosis. *Revista Da Associacao Medica Brasileira* 2010;**56**(1):92-8.

Pupo-Nogueira 2007 *{published data only}*

Pupo-Nogueira A, De Oliveira RM, Petta CA, Podgaec S, Dias JA Jr, Abrao MS. Vascular endothelial growth factor concentrations in the serum and peritoneal fluid of women with endometriosis. *International Journal of Gynaecology and Obstetrics* 2007;**99**(1):33-7.

Quaranta 2006 *{published data only}*

Quaranta MG, Porpora MG, Mattioli B, Giordani L, Libri I, Ingelido AM, et al. Impaired NK-cell-mediated cytotoxic activity and cytokine production in patients with endometriosis: a possible role for PCBs and DDE. *Life Sciences* 2006;**79**(5):491-8.

Rajkumar 1992 *{published data only}*

Rajkumar K, Malliah V, Simpson CW. Identifying the presence of antibodies against endometrial antigens. A preliminary study. *Journal of Reproductive Medicine* 1992;**37**(6):552-6.

Ramos 2011 *{published data only}*

Ramos IML, Podgaec S. Evaluation of CA-125 and CD-23 soluble in patients with pelvic endometriosis [Avaliação do CA-125 e do CD-23 solúvel em pacientes com endometriose pélvica]. Biblioteca Digitais de Teses e Dissertações da USP 2011.

Reis 2012 *{published data only}*

Reis FM, Luisi S, Abro MS, Rocha ALL, Vigan P, Rezende CP, et al. Diagnostic value of serum activin A and follistatin levels in women with peritoneal, ovarian and deep infiltrating endometriosis. *Human Reproduction* 2012;**27**(5):1445-50.

Santulli 2015 *{published data only}*

Santulli P, Streuli I, Melonio I, Marcellin L, M'Baye M, Bititi A, et al. Increased serum cancer antigen-125 is a marker for severity of deep endometriosis. *Journal of Minimally Invasive Gynecology* 2015;**22**(2):275-84.

Sengul 2014 *{published data only}*

Sengul O, Dilbaz B, Halici Z, Ferah I, Cadirci E, Yilmaz F. Decreased serum nesfatin-1 levels in endometriosis. European Journal of Obstetrics, Gynecology, & Reproductive Biology 2014 Mar 15 [Epub ahead of print].

Seo 2010 *{published data only}*

Seo SK, Yang HI, Lee KE, Kim HY, Cho S, Choi YS, et al. The roles of thioredoxin and thioredoxin-binding protein-2 in endometriosis. *Human Reproduction* 2010;**25**(5):1251-8.

Sha 2009 *{published data only}*

Sha GH, Zhang Y, Zhang CY, Wan YP, Zhao ZM, Li CY, et al. Elevated levels of gremlin-1 in eutopic endometrium and peripheral serum in patients with endometriosis. *Fertility and Sterility* 2009;**91**(2):350-8.

Shanti 1999 *{published data only}*

Shanti A, Santanam N, Morales AJ, Parthasarathy S, Murphy AA. Autoantibodies to markers of oxidative stress are elevated in women with endometriosis. *Fertility and Sterility* 1999;**71**(6):1115-8.

Sharma 2010 *{published data only}*

Sharma I, Dhaliwal LK, Saha SC, Sangwan S, Dhawan V. Role of 8-iso-prostaglandin F2alpha and 25-hydroxycholesterol in the pathophysiology of endometriosis. *Fertility and Sterility* 2010;**94**(1):63-70.

Sharpe-Timms 1998 *{published data only}*

Sharpe-Timms KL, Keisler LW, McIntush EW, Keisler DH. Tissue inhibitor of metalloproteinase-1 concentrations are attenuated in peritoneal fluid and sera of women with endometriosis and restored in sera by gonadotropin-releasing hormone agonist therapy. *Fertility and Sterility* 1998;**69**(6):1128-34.

Signorile 2014 *{published data only}*

Signorile PG, Baldi A. Serum Biomarker for Diagnosis of Endometriosis. *Journal of Cellular Physiology* 2014;**229**(11):1731-5.

Slabe 2013 *{published data only}*

Slabe N, Meden-Vrtovec H, Verdenik I, Kosir-Pogacnik R, Ihan A. Cytotoxic T-cells in peripheral blood in women with endometriosis. *Geburtshilfe und Frauenheilkunde* 2013;**73**(10):1042-8.

Socolov 2011 *{published data only}*

Socolov R, Butureanu S, Angioni S, Sindilar A, Boiculese L, Cozma L, et al. The value of serological markers in the diagnosis and prognosis of endometriosis: a prospective case-control study. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2011;**154**(2):215-7.

Steff 2004b *{published data only}*

Steff AM, Gagne D, Page M, Hugo P, Gosselin D. Concentration of soluble intercellular adhesion molecule-1 in serum samples from patients with endometriosis collected during the luteal phase of the menstrual cycle. *Human Reproduction* 2004;**19**:172-8.

Suryawanshi 2013 *{published data only}*

Suryawanshi S, Vlad AM, Lin HM, Mantia-Smaldone G, Laskey R, Lee M, et al. Plasma MicroRNAs as novel biomarkers for

endometriosis and endometriosis-associated ovarian cancer. *Clinical Cancer Research* 2013;**19**(5):1213-24.

Szyllo 2001 *{published data only}*

Szyllo K, Lewy J, Tchorzewski H, Glowacka E, Banasik M, Lewkowicz P, et al. Evaluation of the generation of interleukin-10 by peripheral blood lymphocytes in women with endometriosis. *Ginekologia Polska* 2001;**72**(5):437-41.

Takahashi 1987 *{published data only}*

Takahashi K, Yamane Y, Kijima S, Yoshino K, Shibukawa T, Kitao M. CA 125 antigen is an effective diagnostic for external endometriosis. *Gynecologic and Obstetric Investigation* 1987;**23**(4):257-60.

Takahashi 1988 *{published data only}*

Takahashi K, Nagata H, Kijima S, Kusakari M, Shirai T, Yoshino K, et al. Clinical usefulness of determination of CA 125 levels in the serum and menstrual blood. *Gynecologic and Obstetric Investigation* 1988;**26**(1):63-5.

Takahashi 1989 *{published data only}*

Takahashi K, Kijima S, Yoshino K, Shibukawa T, Kitao M. Serum CA 125 as a marker for patients with external endometriosis. *International Journal of Fertility* 1989;**34**:143-8.

Takemura 2005 *{published data only}*

Takemura Y, Osuga Y, Harada M, Hirata T, Koga K, Morimoto C, et al. Serum adiponectin concentrations are decreased in women with endometriosis. *Human Reproduction* 2005;**20**(12):3510-3.

Tanaka 2000 *{published data only}*

Tanaka T, Umesaki N, Mizuno K, Fujino Y, Ogita S. Anti-endometrial IgM autoantibodies in endometriotic patients: a preliminary study. *Clinical and Experimental Obstetrics & Gynecology* 2000;**27**(2):133-7.

Telimaa 1989 *{published data only}*

Telimaa S, Kauppila A, Ronnberg L, Suikkari AM, Seppala M. Elevated serum levels of endometrial secretory protein PP14 in patients with advanced endometriosis. Suppression by treatment with danazol and high-dose medroxyprogesterone acetate. *American Journal of Obstetrics and Gynecology* 1989;**161**(4):866-71.

Tsao 2007 *{published data only}*

Tsao KC, Hong JH, Wu TL, Chang PY, Sun CF, Wu JT. Elevation of CA 19-9 and chromogranin A, in addition to CA 125, are detectable in benign tumors in leiomyomas and endometriosis. *Journal of Clinical Laboratory Analysis* 2007;**21**(3):193-6.

Tuten 2014b *{published data only}*

Tuten A, Kucur M, Imamoglu M, Oncul M, Acikgoz AS, Sofiyeva N, et al. Serum YKL-40 levels are altered in endometriosis. *Gynecological Endocrinology* 2014;**30**(5):381-4.

Venturella 2011 *{published data only}*

Venturella R, D'Alessandro P, Gallo F, Morelli M, Zullo F. CA 125 modifications throughout menstrual cycle and following gnrh-analog administration to diagnose endometriosis as cause of chronic pelvic pain. A prospective controlled study. *Journal of Endometriosis* 2011;**3**(3):151-8.

Vercellini 1992 *{published data only}*

Vercellini P, Sacerdote P, Panerai AE, Manfredi B, Bocciolone L, Crosignani G. Mononuclear cell beta-endorphin concentration in women with and without endometriosis. *Obstetrics and Gynecology* 1992;**79**(5 Pt 1):743-6.

Wang 2007 *{published data only}*

Wang L, Zheng W, Yu JK, Jiang WZ, Mu L, Zhang SZ. Artificial neural networks combined with surface-enhanced laser desorption/ionization mass spectra distinguish endometriosis from healthy population. *Fertility and Sterility* 2007;**88**(6):1700-2.

Wang 2008 *{published data only}*

Wang L, Zheng W, Mu L, Zhang SZ. Identifying biomarkers of endometriosis using serum protein fingerprinting and artificial neural networks. *International Journal of Gynecology and Obstetrics* 2008;**101**(3):253-8.

Wang 2009 *{published data only}*

Wang H, Gorpudolo N, Li Y, Feng D, Wang Z, Zhang Y. Elevated vascular endothelia growth factor-A in the serum and peritoneal fluid of patients with endometriosis. *Huazhong Keji Daxue Xuebao (Yixueban Dewenban) [Journal of Huazhong University of Science and Technology (Medical Sciences)]* 2009;**29**(5):637-41. Chinese.

Wang 2013b *{published data only}*

Wang P, Zhu L, Zhang X. The role of placental protein 14 in the pathogenesis of endometriosis. *Reproductive Sciences* 2013;**20**(12):1465-70.

Watanabe 1990 *{published data only}*

Watanabe J, Johboh T, Hata H, Kuramoto H. Clinical significance of CA19-9 for endometriosis. *Acta Obstetrica et Gynaecologica Japonica* 1990;**42**(2):155-61.

Wild 1985 *{published data only}*

Wild RA, Shivers CA. Antiendometrial antibodies in patients with endometriosis. *American Journal of Reproductive Immunology* 1985;**8**(3):84-6.

Wild 1991b *{published data only}*

Wild RA, Hirisave V, Podczaski ES, Coulam C, Shivers CA, Satyaswaroop PG. Autoantibodies associated with endometriosis: can their detection predict presence of the disease?. *Obstetrics and Gynecology* 1991;**77**(6):927-31.

Wild 1991c *{published data only}*

Wild RA, Medders D, Zhang RJ. F(ab')2 segment is the active component of immunoglobulin G autoantibody generation in patients with endometriosis. *Fertility and Sterility* 1991;**56**(5):900-3.

Wild 1992 *{published data only}*

Wild RA, Shivers CA, Medders D. Detection of antiendometrial antibodies in patients with endometriosis: methodological issues. *Fertility and Sterility* 1992;**58**(3):518-21.

Wilson 1994 *{published data only}*

Wilson TJ, Hertzog PJ, Angus D, Munnery L, Wood EC, Kola I. Decreased natural killer cell activity in endometriosis patients: relationship to disease pathogenesis. *Fertility and Sterility* 1994;**62**(5):1086-8.

Wojcik-Krowiranda 2010 *{published data only}*

Wojcik-Krowiranda K, Litwinska M, Bienkiewicz A. Usefulness of determination of CA125 concentration as a marker for ovarian tumours with special consideration of ovarian endometriotic cysts. *Przeglad Menopauzalny* 2010;**9**:362-5.

Xavier 2006 *{published data only}*

Xavier P, Belo L, Beires J, Rebelo I, Martinez-de-Oliveira J, Lunet N, et al. Serum levels of VEGF and TNF-a and their association with C-reactive protein in patients with endometriosis Loss of ARID1A/BAF250a expression in ovarian endometriosis and clear cell carcinoma. *Archives of Gynecology and Obstetrics* 2006;**273**(7):227-31.

Yang 2013a *{published data only}*

Yang H, Zhu L, Wang S, Lang J, Xu T. Noninvasive diagnosis of moderate to severe endometriosis: the platelet-lymphocyte ratio cannot be a neoadjuvant biomarker for serum cancer antigen 125. Journal of Minimally Invasive Gynecology 2013 Jul 10 [Epub ahead of print].

Yang 2013b *{published data only}*

Yang H, Lang JH, Zhu L, Wang S, Sha GH, Zhang Y. Diagnostic value of the neutrophil-to-lymphocyte ratio and the combination of serum CA-125 for stages III and IV endometriosis. *Zhonghua Yixue Zazhi [Chinese Medical Journal]* 2013;**126**(11):2011-4. Chinese.

Yi 2010 *{published data only}*

Yi YC, Wang SC, Chao CC, Su CL, Lee YL, Chen LY. Evaluation of serum autoantibody levels in the diagnosis of ovarian endometrioma. *Journal of Clinical Laboratory Analysis* 2010;**24**(5):357-62.

Yin 2000 *{published data only}*

Yin H, Wang S, Dai S. Study on interferon gamma and interleukin-4 contents of plasma and cultured monoclear cell supernatants in patients with endometriosis. *Zhonghua Fuchanke Zazhi [Chinese Journal of Obstetrics and Gynecology]* 2000;**35**(6):327-8. Chinese.

Zachariah 2009 *{published data only}*

Zachariah R, Schmid S, Radpour R, Buerki N, Fan AXC, Hahn S, et al. Circulating cell-free DNA as a potential biomarker for minimal and mild endometriosis. *Reproductive Biomedicine Online* 2009;**18**(3):407-11.

Zhang 2006c *{published data only}*

Zhang H, Niu Y, Feng J, Guo H, Ye X, Cui H. Use of proteomic analysis of endometriosis to identify different protein expression in patients with endometriosis versus normal controls. *Fertility and Sterility* 2006;**86**(2):274-82.

Zhang 2009 *{published data only}*

Zhang H, Feng J, Chang XH, Li ZX, Wu XY, Cui H. Effect of surface-enhanced laser desorption/ionization time-of-flight mass spectrometry on identifing biomarkers of endometriosis. *Zhonghua Yixue Zazhi [Chinese Medical Journal]* 2009;**122**(4):373-6. Chinese.

Zhao 2015 *{published data only}*

Zhao Y, Liu YN, Li Y, Tian L, Ye X, Cui H, et al. Identification of biomarkers for endometriosis using clinical proteomics. *Zhonghua Yixue Zazhi [Chinese Medical Journal]* 2015;**128**(4):520-7. Chinese.

Zheng 2011 *{published data only}*

Zheng N, Pan C, Liu W. New serum biomarkers for detection of endometriosis using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Journal of International Medical Research* 2011;**39**(4):1184-92.

Zhu 2007 *{published data only}*

Zhu X, Jiang F. Measurement of CA125 in serum and ovarian cyst fluid of patients with endometriosis and its clinical significance. *Xi'an Jiaotong Daxue Xuebao (Yexueban) [Journal of Xi'an Jiaotong University (Medical Sciences)]* 2007;**28**(6):687-9. Chinese.

Zomer 2013 *{published data only}*

Zomer MT, Ribeiro R, Trippia CH, Cavalcanti TCS, Hayashi RM, Kondo W. Correlation between serum Ca-125 levels and surgical findings in women with symptoms evocative of endometriosis. *Revista Brasileira de Ginecologia e Obstetricia* 2013;**35**(6):262-7.

Zong 2003 *{published data only}*

Zong LL, Li YL, Ha XQ. Determination of HGF concentration in serum and peritoneal fluid in women with endometriosis. *Di Bi Junyi Daxue Xuebao [Journal of the Fourth Military Medical University]* 2003;**23**(8):757-60. Chinese.

References to ongoing studies

JPRN-UMIN000009223 *{published data only}*

Analysis of miRNA in blood for development of diagnostic biomarkers for endometriosisClinicalTrials.gov Identifier: JPRN-UMIN000009223Primary sponsor: Juntendo University Hospital, Department of Obstetrics and Gynecology. Ongoing study February 2013.

NCT01301885 *{published data only}*

ENDOMET - Novel diagnostic tools and treatments for endometriosisClinicalTrials.gov Identifier: NCT01301885Other study name: CA125_VAS_changes. Ongoing study February 2011.

NCT02091557 *{published data only}*

CA-125 and VAS pain score changes to diagnose endometriosisClinicalTrials.gov Identifier: NCT02091557Other study name: CA125_VAS_changes. Ongoing study January 2011.

NCT02337816 *{published and unpublished data}*

Role of metabolomics in the diagnosis of endometriosisClinicalTrials.gov Identifier: NCT02337816Other study name: ENDOMETAB01. Ongoing study December 2014.

Additional references

ACOG 2010

The American College of Obstetricians and Gynecologists. Practice bulletin no. 114: management of endometriosis. *Obstetrics and Gynecology* 2010;**116**(1):223-36.

Adamson 2008

Adamson GD. Endometriosis Fertility Index (EFI): the new validated endometriosis staging system. Art and science of endometriosis. Proceedings of the Tenth World Congress on Endometriosis; 2008 March 11-14; Melbourne (Vic). 2008.

Almeida Filho 2008

Almeida Filho DP, Oliveira LJ, Amaral VF. Accuracy of laparoscopy for assessing patients with endometriosis. *Sao Paulo Medical Journal* 2008;**126**:305-8.

ASRM 1997

American Society for Reproductive Medicine. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. *Fertility and Sterility* 1997;**67**(5):817-21.

Ballard 2008

Ballard KD, Seaman HE, De Vries CS, Wright JT. Can symptomatology help in the diagnosis of endometriosis? Findings from a national case-control study - Part 1. *BJOG: an International Journal of Obstetrics and Gynaecology* 2008;**115**(11):1382-91.

Batt 2003

Batt R, Mitwally MF. Endometriosis from thelarche to midteens: pathogenesis and prognosis, prevention and pedagogy. *Journal of Pediatric and Adolescent Gynecology* 2003;**16**(6):333-47.

Becker 2014

Becker CM, Laufer MR, Stratton P, Hummelshoj l, Missmer SA, Zondervan KT, et al. World Endometriosis Research Foundation Endometriosis Phenome and biobanking harmonization project: I. Surgical phenotype data collection in endometriosis research. *Fertility and Sterility* 2014;**102**(5):1213-22.

Bedaiwy 2004

Bedaiwy M, Falcone T. Laboratory testing for endometriosis. *Clinica Chimica Acta* 2004;**340**(1-2):41-56.

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;**326**(7379):41-4.

Bossuyt 2008

Bossuyt PM, Leeflang MM. Chapter 6: Developing Criteria for Including Studies. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.methods.cochrane.org/sdt/handbook-dta-reviews.

Brosens 2003

Brosens J, Timmerman D, Strazinski-Powitz A, Brosens I. Noninvasive diagnosis of endometriosis: the role of imaging and markers. *Obstetrics and Gynecology Clinics of North America* 2003;**30**(1):95-114.

Chapron 2003a

Chapron C, Fauconnier A, Vieira M, Barakat Dousset HB, Pansini V, Vacher-Lavenu MC, et al. Anatomical distribution of deeply infiltrating endometriosis: surgical implications and proposition for a classification. *Human Reproduction* 2003;**18**:157-61.

Chapron 2003b

Chapron C, Fauconnier A, Dubuisson JB, Barakat H, Vieira M, Bréart G. Deep infiltrating endometriosis: relation between severity of dysmenorrhea and extent of disease. *Human Reproduction* 2003;**18**:760-6.

Chapron 2003c

Chapron C, Cravello L, Chopin N, Kreiker G, Blanc B, Dubuisson JB. Complications during set-up procedures for laparoscopy in gynecology: open laparoscopy does not reduce the risk of major complications. *Acta Obstetrica et Gynecologica Scandinavica* 2003;**82**:1125-9.

D'Hooghe 2001

D'Hooghe TM, Bambra CS, Xiao L, Peixe K, Hill JA. Effect of menstruation and intrapelvic injection of endometrium on inflammatory parameters of peritoneal fluid in the baboon (Papio anubis and Papio cynocephalus). *American Journal of Obstetrics and Gynecology* 2001;**184**(5):917-25.

D'Hooghe 2002

D'Hooghe TM, Debrock S. Endometriosis, retrograde menstruation and peritoneal inflammation in women and in baboons. *Human Reproduction Update* 2002;**8**(1):84-8.

D'Hooghe 2006

D'Hooghe TM, Mihalyi AM, Simsa P, Kyama CK, Peeraer K, De Loecker P, et al. Why we need a non-invasive diagnostic test for minimal to mild endometriosis with a highsensitivity. *Gynecologic and Obstetric Investigation* 2006;**62**(3):136-8.

De Vet 2008

De Vet HCW, Eisinga A, Riphagen II, Aertgeerts B, Pewsner D. Chapter 7: Searching for Studies. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.methods.cochrane.org/sdt/handbook-dta-reviews.

Dmowski 1997

Dmowski WP, Lesniewicz R, Rana N, Pepping P, Noursalehi M. Changing trends in the diagnosis of endometriosis: a comparative study of women with pelvic endometriosis

presenting with chronic pelvic pain or infertility. *Fertility and Sterility* 1997;**67**:238-43.

Dunselman 2014

Dunselman GA, Vermeulen N, Becker C, Calhaz-Jorge C, D'Hooghe T, De Bie B, et al. ESHRE guideline: management of women with endometriosis. *Human Reproduction* 2014;**29**(3):400-12.

Eskenazi 2001

Eskenazi B, Warner M, Bosignore L, Olive D, Samuels S, Vercellini P. Validation study of nonsurgical diagnosis of endometriosis. *Fertility and Sterility* 2001;**76**:929-35.

Fassbender 2015

Fassbender A, Burney RO, Dorien FO, D'Hooghe T, Giudice L. Update on biomarkers for the detection of endometriosis. BioMed Research International 2015 Jul 9 [Epub ahead of print].

Fauconnier 2005

Fauconnier A, Chapron C. Endometriosis and pelvic pain: epidemiological evidence of the relationship and implications. *Human Reproduction Update* 2005;**11**:595-606.

Frishman 2006

Frishman GN, Salak JR. Conservative surgical management of endometriosis in women with pelvic pain. *Journal of Minimally Invasive Gynecology* 2006;**13**:546-58.

Gao 2006

Gao X, Yeh YC, Outley J, Simon J, Botteman M, Spalding J. Health-related quality of life burden of women with endometriosis: a literature review. *Current Medical Research and Opinion* 2006;**22**:1787-97.

Giudice 2004

Giudice LC, Kao LC. Endometriosis. *Lancet* 2004;**364**:1789-99.

Guo 2009

Guo SW. Recurrence of endometriosis and its control. *Human Reproduction Update* 2009;**15**(4):441-61.

Guzick 1997

Guzick DS, Silliman NP, Adamson GD, Buttram VC Jr, Canis M, Malinak LR, et al. Prediction of pregnancy in infertile women based on the American Society for Reproductive Medicines revised classification of endometriosis. *Fertility and Sterility* 1997;**67**(5):822-9.

Halme 1984

Halme J, Hammond MG, Hulka JF, Raj SG, Talbert LM. Retrograde menstruation in healthy women and in patients with endometriosis. *Obstetrics and Gynecology* 1984;**64**(2):151-4.

Higgins 2008

Higgins JPT. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *[International Journal of Epidemiology]. Oxford Journals* 2008;**37**(7):1158-60.

Hull 2008

Hull ML, Escareno CR, Godsland JM, Doig JR, Johnson CM, Phillips SC, et al. Endometrial-peritoneal interactions during endometriotic lesion establishment. *American Journal of Pathology* 2008;**173**(3):700-15.

Johnson 2013

Johnson NP, Hummelshoj L. Consensus on current management of endometriosis. *Human Reproduction* 2013;**28**(6):1552-68.

Johnson 2015

Johnson NP, Hummelshoj L, Adamson GD, Kecstein J, Taylor H, Abrao MS, et al. Consensus on the classification of endometriosis. Human Reproduction 2016 in press.

Kao 2003

Kao LC, Germeyer A, Tulac S, Lobo S, Yang JP, Taylor RN, et al. Expression profiling of endometrium from women with endometriosis reveals candidate genes for disease-based implantation failure and infertility. *Endocrinology* 2003;**144**(7):2870-81.

Kennedy 2005

Kennedy S, Bergqvist A, Chapron C, D'Hooghe T, Dunselman G, Greb R, et al. ESHRE Special Interest Group for Endometriosis. ESHRE guideline for the diagnosis and treatment of endometriosis. *Human Reproduction* 2005;**20**(10):2698-704.

Koninckx 1991

Koninckx PR, Meuleman C, Demeyere S, Lesaffre E, Cornillie FJ. Suggestive evidence that pelvic endometriosis is a progressive disease, whereas deeply infiltrating endometriosis is associated with pelvic pain. *Fertility and Sterility* 1991;**55**(4):759-65.

Ling 1999

Ling F. Randomized controlled trial of depot leuprolide in patients with chronic pelvic pain and clinically suspected endometriosis. *Obstetrics and Gynecology* 1999;**93**:51-8.

Liu 2005

Liu A, Schisterman EF, Mazumdar M, Hu J. Power and sample size calculation of comparative diagnostic accuracy studies with multiple correlated test results. *Biometrical Journal* 2005;**47**(2):140-50.

Marchino 2005

Marchino GL, Gennarelli G, Enria R, Bongioanni F, Lipari G, Massobrio M. Diagnosis of pelvic endometriosis with use of macroscopic versus histologic findings. *Fertility and Sterility* 2005;**84**:12-5.

Martin 2006

Martin DC. Applying STARD criteria to the laparoscopic identification of endometriosis (abstract) . Fertility and Sterility. 2006; Vol. 86 Suppl 2:270.

Matalliotakis 2008

Matalliotakis IM, Vassiliadis S, Goumenou AG, Fragouli Y, Athanassakis I. Suggested markers for the diagnosis of endometriosis. *Current Women's Health Reviews* 2008;**4**:25-8.

Matsuzaki 2006

Matsuzaki S, Canis M, Pouly JL, Rabischong B, Botchorishvili R, Mage G. Relationship between delay of surgical diagnosis and severity of disease in patients with symptomatic deep infiltrating endometriosis. *Fertility and Sterility* 2006;**86**:1314-6.

May 2010

May KE, Conduit-Hulbert SA, Villar J, Kirtley S, Kennedy SH, Becker CM. Peripheral biomarkers of endometriosis: A systematic review. *Human Reproduction Update* 2010;**16**(6):651-74.

McGraw-Hill Dictionary of Medicine 2006

Segen JC. McGraw-Hill Concise Dictionary of Modern Medicine. 2nd Edition. New York: The McGraw-Hill Companies, Inc, 2006.

Medeiros 2009

Medeiros LR, Rosa DD, Bozzetti MC, Fachel JM, Furness S, Garry R, et al. Laparoscopy versus laparotomy for benign ovarian tumour. *Cochrane Database of Systematic Reviews* 2009, Issue 2. [DOI: 10.1002/14651858.CD004751.pub3]

Mol 1998

Mol BW, Bayram N, Lijmer JG, Wiegerinck MA, Bongers MY, Van der Veen F, et al. The performance of CA-125 measurement in the detection of endometriosis: a meta-analysis. *Fertility and Sterility* 1998;**70**:1101-8.

Nyholt 2012

Nyholt DR, Low SK, Anderson CA, Painter JN, Uno S, Morris AP, et al. Genome-wide association meta-analysis identifies new endometriosis risk loci. *Nature Genetics* 2012;**44**(12):1355-9.

Rahimoglu 2014

Rahmioglu N, Fassbender A, Vitonis AF, Tworoger SS, Hummelshoj L, D'Hooghe TM, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking armonization Project: III. Fluid biospecimen collection, processing, and storage in endometriosis research. *Fertility and Sterility* 2014;**102**(5):1233-43.

Redwine 2003

Redwine DB. 'Invisible' microscopic endometriosis: a review. *Gynecologic and Obstetric Investigation* 2003;**55**(2):63-7.

RevMan 2014 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Rogers 2009

Rogers PA, D'Hooghe TM, Fazleabas A, Gargett CE, Giudice LC, Montgomery GW, et al. Priorities for endometriosis research: recommendations from an international consensus workshop. *Reproductive Sciences* 2009;**16**(4):335-46.

Rutjes 2005

Rutjes AWS, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PMM. Case–control and two-gate designs in diagnostic accuracy studies. *Clinical Chemistry* 2005;**51**(8):1335-41.

Sampson 1927

Sampson JA. Peritoneal endometriosis due to menstrual dissemination of endometrial tissue into the peritoneal cavity. *American Journal of Obstetrics and Gynecology* 1927;**14**:442-69.

Simoens 2012

Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, et al. The burden of endometriosis: costs and quality of life of women with endometriosis and treated in referral centres. *Human Reproduction* 2012;**27**(5):1292-9.

Sinaii 2002

Sinaii N, Cleary SD, Ballweg ML, Nieman LK, Stratton P. High rates of autoimmune and endocrine disorders, fibromyalgia, chronic fatigue syndrome and atopic diseases among women with endometriosis: A survey analysis. *Human Reproduction* 2002;**17**(10):2715-24.

SOGC 2010

Society of Obstetricians Gynaecologists of Canada. Endometriosis: diagnosis and management. SOGC clinical practice guideline no. 244. *Journal of Obstetrics and Gynaecology Canada* 2010;**32**:S1-S28.

Somigliana 2006

Somigliana E, Vigano P, Parazzini F, Stoppelli S, Giambattista E, Vercellini P. Association between endometriosis and cancer: A comprehensive review and a critical analysis of clinical and epidemiological evidence. *Gynecologic Oncology* 2006;**101**(2):331-41.

Spaczynski 2003

Spaczynski RZ, Duleba AJ. Diagnosis of endometriosis. *Seminars in Reproductive Medicine* 2003;**21**:193-208.

Stegmann 2008

Stegmann BJ, Sinaii N, Liu S, Segars J, Merino M, Nieman LK, et al. Using location, color, size, and depth to characterize and identify endometriosis lesions in a cohort of 133 women. *Fertility and Sterility* 2008;**89**:1632-6.

The Gale Encyclopedia of Medicine 2011

Olendorf D, Jeryan C, Boyden K (editors). The Gale Encyclopedia of Medicine (5 volume set). 4th Edition. Detroit: The Gale Group, Inc, 2011.

Vercellini 1996

Vercellini P, Trespidi L, De Giorgi O, Cortesi I, Parazzini F, Crosignani GP. Endometriosis and pelvic pain: relation to disease stage and localization. *Fertility and Sterility* 1996;**65**:299-304.

Vigano 2004

Vigano P, Parazzini F, Somigliana E, Vercellini P. Endometriosis: epidemiology and aetiological factors. *Best Practice & Research. Clinical Obstetrics & Gynaecology* 2004;**18**(2):177-200.

Vitonis 2014

Vitonis AF, Vincent K, Rahmioglu N, Fassbender A, Buck Louis G, Hummelshoj L, et al. World Endometriosis Research Foundation Endometriosis Phenome and Biobanking

Harmonization Project: II. Clinical and covariate phenotype data collection in endometriosis research. *Fertility and Sterility* 2014;**102**(5):1223-32.

Whiting 2005

Whiting PF, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology* 2005;**5**:19.

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. the QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

Wykes 2004

Wykes CB, Clark TJ, Khan KS. Accuracy of laparoscopy in the diagnosis of endometriosis: a systematic quantitative review. *BJOG - an International Journal of Obstetrics and Gynaecology* 2004;**111**:1204-12.

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Acien 1989

Study characteristics	
Patient sampling	*Primary objective*: to measure the levels of CA-125 in the serum of normal women and in patients with endometriosis before, during and after treatment with danazol or a luteinising hormone-releasing hormone agonist, to evaluate the influence of these treatments on the levels of CA-125 and the possible relation with reactivation of endometriosis after treatment *Participants*: women with endometriosis confirmed by laparoscopy and a group of regularly menstruating women with a normal pelvis at laparoscopy *Selection criteria*: not specified *Study design*: longitudinal, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis group - infertility in 70.4%, not specified otherwise *Age*: range 22-43 years *Number of participants enrolled*: 68 women (11 postmenopausal women were enrolled and analysed separately - not considered in this review) *Number of participants available for analysis*: 68 women (all in luteal cycle phase) *Setting*: not stated; authors' affiliations: the Royal Free (University) Hospital, London; and School of Medicine, University of Alicante, Spain *Place of study*: not specified, Europe *Period of study*: not stated *Language*: English
Index tests	*Index test*: CA-125

Yang 2004

Yang WC, Chen HW, Au HK, Chang CW, Huang CT, Yen YH, et al. Serum and endometrial markers. *Best Practice & Research. Clinical Obstetrics & Gynaecology* 2004;**18**(2):305-18.

Yeung 2009

Yeung PP Jr, Shwayder J, Pasic RP. Laparoscopic management of endometriosis: comprehensive review of best evidence. *Journal of Minimally Invasive Gynecology* 2009;**16**:269-81.

References to other published versions of this review

Nisenblat 2012

Nisenblat V, Farquhar C, Akoum A, Fraser I, Bossuyt PMM, Hull ML. Non-invasive tests for the diagnosis of endometriosis. *Cochrane Database of Systematic Reviews* 2012, Issue 1. [DOI: 10.1002/14651858.]

Acien 1989 *(Continued)*

	Details of the index test procedure as stated: serum CA-125 was measured with an immunoradiometric assay (Abbot CA-125 RIA); working assay range was 6-500 U/ml. sample processing and experiments not described *Threshold for positive result*: > 30 U/ml, not pre-specified *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: interassay and intra-assay CV 3.5%-6.4%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 54/68 (79%): stage I-II 40, stage III-IV 14; controls n = 14 *Reference standard*: laparoscopy N = 68 (100%) *Description of positive case definition by reference standard as reported*: visual inspection, staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: samples were taken at laparoscopy *Withdrawals*: none
Comparative	
Key conclusions by the authors	Increases in CA-125 values above 30 U/ml were more likely to indicate reactivation of endometriosis than when CA-125 did not increase
Conflict of interest	Not reported
Notes	The reported CA-125 values during and after treatment with Danazol or GnRH analogues are not included in this review Additional control group of postmenopausal women (N = 11) was not considered in calculation of diagnostic estimates

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge	Unclear		

Acien 1989 *(Continued)*

of the results of the reference standard?			
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Agic 2008

Study characteristics	
Patient sampling	*Primary objective*: to investigate the combination of CCR1 mRNA, MCP-1, and CA-125 protein measurements in peripheral blood as a diagnostic test for endometriosis and to study the possible use of these markers in the peripheral blood of patients with adenomyosis *Participants*: patients who underwent laparoscopy for various indications *Selection criteria*: Inclusion criteria: no endocrine therapy for at least 3 months; exclusion criteria: suspected or ascertained diagnosis of malignancy, pregnancy, menopausal age or refusal to participate in the study *Study design*: cross-sectional, two-gate design, prospective collection of samples

Agic 2008 *(Continued)*

Patient characteristics and setting	*Clinical presentation*: endometriosis group - dysmenorrhoea, dyspareunia, chronic pelvic pain and infertility; 12 women with known history of endometriosis; controls - undergoing surgery for subserosal leiomyomata or tubal ligation *Age*: reproductive age *Number of participants enrolled*: 151 women (11 women with adenomyosis were enrolled and analysed separately - not considered in this review) *Number of participants available for analysis*: 151 women (all in follicular cycle phase) *Setting*: Department of O&G, University of Schleswig-Holstein *Place of study*: Luebeck, Germany *Period of study*: not stated *Language*: English
Index tests	*Index test*: CA-125, CCR1, MCP-1 *Details of the index test procedure as stated*: CCR1 expression detected by RT-PCR (SuperScriptTM II RT, SYBR Green MM, normalised to HPRT housekeeping gene); MCP-1 levels detected by using a commercially available ELISA kit (R&D Systems, GmbH, Germany) with assay sensitivity of 5 pg/ml; CA-125 level detected by using a commercially available electro-chemiluminescent immunometric assay (ECLIA, Roche Diagnostics GmbH, Germany) with assay sensitivity of 0.6 IU/ml; all the experiments were repeated x 3 times; the test was considered positive for endometriosis if at least one of the markers was above the threshold; sample processing and experiments described *Threshold for positive result*: CCR1/HPRT > 1.16, MCP-1 > 140 pg/ml, CA-125 > 50 IU/ml- all pre-specified *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: for MCP-1 the intra- and interassay CV was 2.5% and 4.5%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 102/151 (68%): stage I-II 37, stage III-IV 65; controls n = 49 *Reference standard*: laparoscopy N = 151 (100%) + histology *Description of positive case definition by reference standard as reported*: visual inspection, histological diagnosis; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained 24 hours prior to anaesthesia and laparoscopy *Withdrawals*: none
Comparative	
Key conclusions by the authors	The results imply the potential use of CCR1 mRNA, MCP-1, and CA-125 protein measurements for the diagnosis or exclusion of endometriosis
Conflict of interest	Not reported
Notes	The reported estimates for diagnosis of adenomyosis are not presented in this review

Agic 2008 *(Continued)*

The reported diagnostic estimated per severity of endometriosis are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Agic 2008 *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Akoum 1996

Study characteristics	
Patient sampling	*Primary objective*: to evaluate MCP-1 in the peripheral blood of women with and without endometriosis *Participants*: women who underwent laparoscopy for infertility and pelvic pain (endometriosis group) and fertile women who underwent tubal ligation or reanastomosis with normal pelvis (controls) *Selection criteria*: inclusion criteria: no other pelvic disorders; no treatment with any antiinflammatory or hormonal medications at least 3 months before laparoscopy *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis group - infertility - 26 (46%) and pelvic pain; healthy fertile controls *Age*: mean age 31.2 ± 7.2 years (endometriosis group), 33.7 ± 5.6 years (controls) *Number of participants enrolled*: 101 women *Number of participants available for analysis*: 101 women (in follicular or luteal phase of menstrual cycle) *Setting*: university hospital, Saint-Francois d'Assise hospital Universite Laval *Place of study*: Quebec, Canada *Period of study*: not stated *Language*: English
Index tests	*Index test*: MCP-1 *Details of the index test procedure as stated*: MCP-1 concentrations were measured, with ELISA (R & D Systems, Minneapolis); the biologic activity of MCP-1 (monocyte chemotaxis induction) was evaluated by using a Boyden chamber and a human cell line (U937); assay sensitivity limit 50 pg/ml; sample processing and experiments described in detail *Threshold for positive result*: > 100 pg/ml - not pre-specified *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: Intra- and interassay CV < 6%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 57/101 (56%): stage I-II 47, stage III-IV - 10; controls n = 44 *Reference standard*: laparoscopy N = 101 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS system

Akoum 1996 *(Continued)*

	Examiners: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn a few days before laparoscopy *Withdrawals*: none
Comparative	
Key conclusions by the authors	Endometriosis is associated with increased level and activity of MCP-1 in the peripheral blood. The elevation and activation of this cytokine could play a relevant role in the immuno-inflammatory process associated with the disease
Conflict of interest	Not reported; supported by a grant No. MT-12541 from the Medical Research Council, Ottawa, Canada
Notes	The reported data on monocyte chemotactic activity are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Akoum 1996 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes

		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Andreoli 2011

Study characteristics	
Patient sampling	*Primary objective*: to investigate the role of IL-10, -12, -17, and -23 in infertile patients with minimal-mild endometriosis *Participants*: women who underwent laparoscopy for investigation of infertility or for tubal ligation *Selection criteria*: exclusion criteria: presence of autoimmune disease, absence of peritoneal liquid during laparoscopy, coexistence of other causes of infertility, and hormonal medication in the 3 months before surgery *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis group - infertility 100% and pelvic pain (40%), other causes of infertility were excluded by hysterosalpingography, spermiogram, and measurements of serum FSH, PRL, and TSH levels; control group - women requesting tubal ligation *Age*: mean age 32.48 ± 4.99 years (endometriosis group), 33.63 ± 6.51 years (controls) *Number of participants enrolled*: 80 women *Number of participants available for analysis*: 80 women (all in follicular phase of menstrual cycle) *Setting*: university hospital, Hospital de Clınicas de Porto Alegre, Universidade Federal do Rio Grande do Sul *Place of study*: Porto Alegre, Brazil *Period of study*: March 2007 - December 2008 *Language*: English
Index tests	*Index test*: IL-10, IL-12, IL-17, and IL-23

Andreoli 2011 *(Continued)*

	Details of the index test procedure as stated: IL-10, IL-12 (p70), IL-17a, and IL-23 (p19/p40) concentrations were measured, with ELISA Human Ready-SET-Go! commercial kits (eBioscience, San Diego, CA); the sensitivity was 2 pg/ml, 4 pg/ml, 4 pg/ml, and 15 pg/ml, respectively *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 40/80 (50%): all stage I-II; controls n = 40 *Reference standard*: laparoscopy N = 80 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn at laparoscopy *Withdrawals*: none
Comparative	
Key conclusions by the authors	Higher IL-23 levels were encountered in the peritoneal fluid of women with endometriosis, suggesting a possible role of this cytokine in these women's infertility
Conflict of interest	The authors reported no conflict of interests
Notes	The data for markers measured in peritoneal fluid are not presented in this review For all the biomarkers there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		

Andreoli 2011 *(Continued)*

If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Barbati 1994

Study characteristics	
Patient sampling	*Primary objective*: to evaluate CA-125 in peritoneal fluid as an indicator of endometriosis *Participants*: women undergoing laparotomy or diagnostic laparoscopy for infertility or pelvic pain *Selection criteria*: inclusion criteria: no hormonal medications at least 3 months before surgery, mid-follicular cycle phase *Study design*: cross-sectional, single-gate design, prospective enrolment and collection of samples
Patient characteristics and setting	*Clinical presentation*: Inertility or pelvic pain *Age*: range 23-41 years (endometriosis group), 16-55 years (controls) *Number of participants enrolled*: 45 women

Barbati 1994 *(Continued)*

	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 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum levels of CA-125 were measured by immuno-radiometric 'one step' sandwich assay (IRMA CA-125 II K, Sorin Biomedica, Italy); minimal detectable concentration 1.4 U/ml; sample processing and experiments are described in details *Threshold for positive result*: > 35 U/ml, not pre-specified *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: Intra- and interassay CV was 7.5% and 8.7%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 18/45 (40%): stage I-II 12, stage III-IV 6; controls n = 27: normal pelvis - 7, other benign pathologies - 20 *Reference standard*: laparoscopy/laparotomy N = 45 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples collected immediately before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The sensitivity of CA-125 test for endometriosis in peritoneal fluid is greater than in serum. Therefore its measurement could be useful in the detection of early stage of endometriosis, which tends to be overlooked by the CA-125 serum test
Conflict of interest	Not reported; supported by a grant 92.02130.39 ACRO from the National Research Council, Rome, Italy
Notes	The reported data on CA-125 in peritoneal fluid is not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Barbati 1994 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Barbosa 2009

Study characteristics

Barbosa 2009 *(Continued)*

Patient sampling	*Primary objective*: to determine the frequency of endometriosis and the correlation between serum CA-125 levels and the presence of endometriotic lesions in the peritoneum of asymptomatic fertile patients *Participants*: women who underwent laparoscopy for tubal ligation *Selection criteria*: inclusion criteria: reproductive age, no symptoms of endometriosis *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: asymptomatic fertile women requesting tubal ligation *Age*: mean age 33.68 ± 4.63 years, range 21-44 years *Number of participants enrolled*: 80 women *Number of participants available for analysis*: 80 women (all in follicular phase of menstrual cycle) *Setting*: university hospital, family planning outpatient clinic of Faculdade de Medicina do ABC (FMABC) *Place of study*: Santo André, Brazil *Period of study*: not specified *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: Serum CA-125 levels were measured in accordance with the manufacturer's instructions (BYK-Sangtec Diagnostica GmbH, Germany). When the CA-125 values were higher than 35 U/ml, a second measurement was performed to confirm the result; sample handling described *Threshold for positive result*: > 35 U/ml, pre-specified *Examiners*: not specified, unclear if blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 13/80 (16.2%); all stage I-II; controls n = 67 *Reference standard*: laparoscopy N = 80 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: the criterion for histological classification of endometriosis was identification of stromal endometrioid or epithelial elements of Müllerian type, with or without stroma, associated with signs of haemorrhage and fibrosis (peritoneal biopsy from four different sites: left and right ovarian fossae, and left and right sacrouterine ligaments; 320 slides stained with hematoxylin-eosin were studied); the morphological criteria were: stromal disease - only endometrial stroma was found; well-differentiated disease - glands similar to topical endometrium were found; undifferentiated disease - the appearance of the glands was different from topical endometrium; and mixed disease - the appearance of the glands was atypical or undifferentiated; staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn on the first 3 days of cycle prior to surgery *Withdrawals*: none
Comparative	

Barbosa 2009 *(Continued)*

Key conclusions by the authors	The presence of endometriotic lesions in the peritoneum of fertile patients supports the hypothesis that incidental findings of minimal or mild endometriosis may not be of clinical significance, and that the progression of the disease probably occurs as a result of immunological and genetic abnormalities. Serum CA-125 levels did not show any diagnostic significance with regard to detecting the disease
Conflict of interest	The authors reported no conflict of interests; there was no funding for the study
Notes	For CA-125 there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of	Yes		

Barbosa 2009 *(Continued)*

the results of the index tests?			
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Barcz 2002

Study characteristics	
Patient sampling	*Primary objective*: to determine the angiogenic activity and concentrations of IL-8 in peritoneal fluid and sera of patients suffering from endometriosis *Participants*: women who underwent laparoscopy for various indications *Selection criteria*: exclusion criteria for control group: pelvic inflammatory disease, ovarian cysts and peritoneal adhesions; not specified otherwise *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: reproductive age *Number of participants enrolled*: 84 women *Number of participants available for analysis*: 84 women (all in follicular phase of menstrual cycle) *Setting*: university hospital, Department of O&G, the Medical University of Warsaw *Place of study*: Warsaw, Poland *Period of study*: not specified *Language*: English
Index tests	*Index test*: IL-8; angiogenic activity *Details of the index test procedure as stated*: angiogenic activity was determined as the number of newly formed blood vessels, induced by intradermal injection of peritoneal fluid and sera obtained from the patients into at least 3 mice according to the Sidky and Auerbach experimental model; all the newly formed blood vessels were identified and counted using the criteria suggested by Sidky and Auerbach; referenced to the primary source; IL-8 concentrations were determined using ELISA method (R&D SYSTEM); laboratory technique and sample handling described

Barcz 2002 *(Continued)*

	Threshold for positive result: not reported *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 52/84 (62%): stage I-II 22, stage III-IV 30; controls n = 32 *Reference standard*: laparoscopy N = 84 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: endometriosis was diagnosed on the basis of visualised changes and histopathological examination; staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn at laparoscopy *Withdrawals*: for IL-8 no data available for 10 women from the control and for 5 women from endometriosis group, withdrawals not explained
Comparative	
Key conclusions by the authors	Angiogenesis plays an important role in pathogenesis of endometriosis. Although IL-8 takes part in neovascularisation, there are other factors modulating angiogenesis in endometriosis
Conflict of interest	Not reported
Notes	The data for the biomarkers measured in peritoneal fluid are not presented in this review For serum angiogenic activity and IL-8, there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		

Barcz 2002 *(Continued)*

If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Bedaiwy 2002

Study characteristics	
Patient sampling	*Primary objective*: to investigate the ability of a group of serum/peritoneal fluid markers to non-surgically predict endometriosis *Participants*: patients undergoing laparoscopy for pain, infertility, tubal ligation or sterilisation reversal *Selection criteria*: exclusion criteria: blood contaminated peritoneal fluid *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: median age 32.5 years, range 18-44 years *Number of participants enrolled*: 130 women

Bedaiwy 2002 *(Continued)*

	Number of participants available for analysis: 91 women (in follicular or luteal phase the cycle, numbers not reported) *Setting*: tertiary care referral centre, the Cleveland Clinic Foundation *Place of study*: Cleveland, Ohio, USA *Period of study*: 1998-2000 *Language*: English
Index tests	*Index test*: IL-6, IL-ß, IL-12, IL-13,TNF-α *Details of the index test procedure as stated*: serum levels of IL-ß IL-6, IL-12, IL-13 and TNF-α were measured in parallel for each patient by using commercially available, cytokine-specific, ELISA (R&D Systems Inc, Minneapolis, USA); assay sensitivities 1.0, 0.7, 5.0, 32.0 and 4.4 pg/ml, with standard curve ranges of 3.9-250, 3.12-300, 7.8-500, 62.5-4000 and 15.6-1000 pg/ml, respectively; sample preparation described *Threshold for positive result*: IL-6 > 2 pg/ml; > 4 pg/ml; > 7.5 pg/ml - selected during analysis, not prespecified *Examiners*: no information provided; not blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 56/91 (62%): stage I-II 34, stage III-IV 22; controls n = 35 *Reference standard*: laparoscopy N = 91 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rASRM system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: "Blood samples were collected from each patient pre-operatively", from the context - just prior to surgery *Withdrawals*: 39 patients were excluded because of blood-contaminated peritoneal fluid (did not meet inclusion criteria)
Comparative	
Key conclusions by the authors	In summary, serum IL-6 and peritoneal fluid TNF-α may be good markers for endometriosis and permit non-surgical diagnosis; such findings must be verified in larger group of patients and controls before being applied within the clinical situation
Conflict of interest	Not reported; the study was supported by a research grant from MISC of the Cleveland Clinic Foundation (RPC#2156)
Notes	For IL-ß, IL-12, IL-13 there was no statistically significant difference between the groups - no data available for meta-analysis The levels of TNF-α were statistically significantly higher in endometriosis, but there was no data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid are not presented in this review
Methodological quality	

Bedaiwy 2002 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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?	No		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Bedaiwy 2002 *(Continued)*

	High

Bilibio 2014

Study characteristics	
Patient sampling	*Primary objective*: to evaluate serum prolactin and CA-125 levels as biomarkers for the diagnosis of peritoneal endometriosis *Participants*: women who underwent laparoscopy for infertility, pelvic pain or tubal ligation *Selection criteria*: inclusion criteria for endometriosis group: superficial peritoneal implants confirmed by biopsy, regular menstrual cycles, negative transvaginal ultrasonography for endometrioma and deep endometriosis; exclusion criteria: 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 *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis group - infertility, pelvic pain or both; other causes of infertility were excluded by 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 *Language*: English
Index tests	*Index test*: CA-125, prolactin *Details of the index test procedure as stated*: Prolactin was analysed with Roche Diagnostics GmbH, Mannheim, Germany and CA-125 with Roche Diagnostics; sample handling described *Threshold for positive result*: for prolactin > 14.80 ng/ml and > 20 ng/ml, for CA-125 > 19.80 U/l and > 35 U/l; not pre-specified *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: for prolactin the intra- and interassay CV was 2.0% and 1.7%, for CA-125 1.8% and 1.6%
Target condition and reference standard(s)	*Target condition*: peritoneal endometriosis *Prevalence of target condition in the sample*: n = 63/97 (65%): stage I-II 40, stage III-IV 23; controls n = 34 *Reference standard*: laparoscopy n = 97 (100%) + histopathology *Description of positive case definition by reference standard test as reported:* visual inspection, confirmed by histopathology; staging according to the rASRM classification

Bilibio 2014 *(Continued)*

	Examiners: the same surgical staff performed all endoscopic procedures; the surgeons were blinded to the result of index test
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn on day 19–21, prior to surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Serum CA-125 and prolactin levels assessed together, and considering the cut-off for CA-125 (19.9 U/l) and prolactin (14.8 ng/ml), allow the diagnosis of peritoneal endometriosis with acceptable sensitivity and specificity (77 and 88%) and a high negative predictive value (97%)
Conflict of interest	Not reported
Notes	The separate data for different clinical presentations od endometriosis (pain only or infertility only) are not presented in this review The reported diagnostic estimates for the subgroups by severity of endometriosis are not included in the review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**

Bilibio 2014 *(Continued)*

DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Borkowski 2008

Study characteristics	
Patient sampling	*Primary objective*: to evaluate total serum and peritoneal concentrations of vitamin D-binding protein in women with endometriosis, known as an inflammation-associated disease *Participants*: women undergoing surgical visualisation because of pain, infertility or both *Selection criteria*: inclusion criteria: pre-menopausal age, regular cycle (25-32 days) *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility or both *Age*: range 21-50 years *Number of participants enrolled*: 43 women *Number of participants available for analysis*: 43 women (all in follicular phase of the menstrual cycle) *Setting*: Department of O&G, Wrocław Medical University; Laboratory of Reproductive Immunology, Institute of Immunology and Experimental Therapy, Polish Academy of Sciences *Place of study*: Wrocław, Poland

Borkowski 2008 *(Continued)*

	Period of study: not stated *Language*: English
Index tests	*Index test*: vitamin D-binding protein *Details of the index test procedure as stated*: serum vitamin D-binding protein was measured by using ELISA (using goat polyclonal antibody against human Gc globulin; absorbance at 490 nm was read by using a Bio-Tek 340 EL spectrophotometer; data analysed with KC3 software (Bio-Tek Instruments; Winooski, USA); concentration were calculated by interpolation from a six-point logarithmic standard curve); sample processing and experiments described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: intra- and interassay CV < 10% and < 15%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 26/43 (61%): stage I-II 11, stage III-IV 15; controls n = 17 *Reference standard*: laparoscopy N = 43 (100%) + histology *Description of positive case definition by reference standard as reported*: visual inspection, confirmed by histopathology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: venous blood was collected before the induction of anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	Serum and peritoneal DBP concentrations are not affected in women with endometriosis; however, based on the latest published data, it is possible that both the serum and peritoneal concentrations of vitamin D-binding protein may be dependent on Gc genotype, which results in differential modulation of monocyte/macrophage activity
Conflict of interest	Not reported; supported by grant No. 3P05E 077 24 from the Polish Ministry for Scientific Research and Information Technology
Notes	For vitamin D-binding protein there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		

Borkowski 2008 *(Continued)*

Was a 'two-gate' design avoided?	Yes

	Unclear	**Low**
DOMAIN 2: Index Test All tests		
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	
Was a cycle phase considered in interpretation of the result of index test?	Yes	
	High	**Low**
DOMAIN 3: Reference Standard		
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?	Yes	
	Low	**Low**
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Braun 1996

Study characteristics	
Patient sampling	*Primary objective*: to investigate the capacity of peripheral blood monocytes (PBM) from women with endometriosis to secrete tumour necrosis factor-alpha (TNF-α), interleukin (IL) IL-6, IL-8, and IL-10

Braun 1996 *(Continued)*

	Participants: women who underwent laparoscopy for suspected endometriosis or tubal ligation *Selection criteria*: not reported *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: reproductive age *Number of participants enrolled*: 30 women *Number of participants available for analysis*: 30 women (all in luteal phase of menstrual cycle) *Setting*: Institute for the Study and Treatment of Endometriosis, Department of Medicine, Rush Medical College *Place of study*: Chicago, IL, US *Period of study*: not specified *Language*: English
Index tests	*Index test*: TNF-α, IL-6, IL-8, and IL-10 *Details of the index test procedure as stated*: concentrations of cytokines in peripheral blood monocytes we measured by using commercially available ELISA kits (Biosourse International, CA) according to the manufacturer instructions; sensitivity of assays ranged from < 1 pg/ml (TGF-a) to 11 pg/ml (IL-8; sample handling and laboratory methods described *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 20/30 (67%): stages not reported; controls n = 10 *Reference standard*: laparoscopy N = 30 (100%) *Description of positive case definition by reference standard test as reported*: laparoscopic examination and staging *Examiners*: the same surgical staff performed all endoscopic procedures; the surgeons were blinded to the result of index test
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Endometriosis is associated with increased basal and stimulated synthesis and secretion of several different cytokines by PBM. Each of the cytokines found to be affected has the capacity to play a role in the symptomatology or pathogenesis of the disease
Conflict of interest	Not reported; the work was supported in part by Public Health Service grants CA58922, Bethesda Maryland and a grant from Sterling International, New York
Notes	The data for induced monocyte cytokine biosynthesis are not included in the review

Braun 1996 *(Continued)*

For TNF-α, IL-6 and IL-8 there was statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review

For IL-10 there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Braun 1996 *(Continued)*

Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Calienno 2008

Study characteristics	
Patient sampling	*Primary objective*: to evaluate whether endometriotic cells themselves are able to secrete cytokines that may contribute in creating a favourable microenvironment for their implantation and survival in the peritoneal cavity; and to consider levels of inflammatory and chemotactic mediators that can justify a possible immune system involvement *Participants*: women who underwent surgery for suspected endometriosis or benign ovarian cyst *Selection criteria*: not specified *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: mean age 34.76 ± 2.14 years, range 20-50 years *Number of participants enrolled*: 30 women *Number of participants available for analysis*: 30 women (14 in follicular and 16 in luteal phase of menstrual cycle) *Setting*: Department O&G, Hospital San Gerardo di Monza, University of Milan - Bicocca *Place of study*: Milano-Bicocca, Italy *Period of study*: not specified *Language*: Italian
Index tests	*Index test*: MCP-1, IL-8 *Details of the index test procedure as stated*: the concentrations of both markers were measured by using ELISA method; not specified otherwise *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 20/30 (67%): stage I-II 2, stage III-IV 18; controls n = 10 *Reference standard*: laparoscopy/laparotomy N = 30 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection, confirmed by histopathology; staging according to the rASRM classification

Calienno 2008 *(Continued)*

	Examiners: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Increased tissue levels of IL-8 and MCP-1 in patients affected by endometriosis may play an important role in the pathogenesis and development of this disease. Moreover, higher serum concentration of MCP-1 in patients affected by endometriosis may indicate a higher activation of circulating monocytes
Conflict of interest	Not reported
Notes	For MCP-1 there was statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review For IL-8 there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in endometriotic tissue are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Unclear**
DOMAIN 3: Reference Standard			

Calienno 2008 *(Continued)*

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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Chen 1998

Study characteristics	
Patient sampling	*Primary objective*: to review the CA-125 concentration in women with dysmenorrhoea in order to delineate the predicting value for the diagnosis of endometriosis and its severity; to evaluate the significance of CA-125 in monitoring therapy and follow-up *Participants*: patients undergoing laparoscopy for dysmenorrhoea *Selection criteria*: inclusion criterion: luteal cycle phase, not specified otherwise *Study design*: longitudinal prospective single-gate design, consecutive enrolment
Patient characteristics and setting	*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 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 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum CA-125 was determined by immunoradiometric assay ELISA-CA 125 II kit (GIF-SUR-YVETTE CEDEX, France); no other details provided

Chen 1998 *(Continued)*

	Threshold for positive result: > 35 U/ml, pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not stated
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 131/157 (83%); stage I-II 56, stage III-IV 75; controls n = 26 *Reference standard*: laparoscopy N = 157 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection, histological confirmation; staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were taken at admission for laparoscopy *Withdrawals*: 2 patients (1%) excluded from analysis because of fibroid uterus
Comparative	
Key conclusions by the authors	For endometriosis, CA-125 is a valuable adjuvant in the follow-up of recurrence in patients with advanced endometriosis and initially elevated CA-125 levels. It is not an effective screening tool for patients with dysmenorrhoea, or for monitoring therapy. There was no significant correlation between the development of endometriosis and reproductive factors.
Conflict of interest	Not reported
Notes	The reported diagnostic estimates for different stages of endometriosis are not included in this review The reported CA-125 levels at different time points during and after Danazol treatment and relationship CA-125 are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge	Unclear		

Chen 1998 *(Continued)*

of the results of the reference standard?			
If a threshold was used, was it pre-specified?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Cho 2007

Study characteristics	
Patient sampling	*Primary objective*: to evaluate serum and urinary levels of vascular endothelial growth factors TNF-α and sFlt-1 in patients with endometriosis *Participants*: women who underwent laparoscopy or laparotomy for different indications including pelvic masses, pelvic pain, suspicious endometriosis, infertility and diagnostic evaluation *Selection criteria*: inclusion criteria: pre-menopausal age *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility, pelvic mass, other not specified *Age*: mean age 32.65 ± 6.82 years (endometriosis group), 30.96 ± 6.36 years (controls)

Cho 2007 *(Continued)*

	Number of participants enrolled: 43 women *Number of participants available for analysis*: 43 women (in follicular or luteal cycle phase, numbers not specified) *Setting*: Department of O&G, Yongdong Severance Hospital, Yonsei University College of Medicine *Place of study*: Seoul, Korea *Period of study*: not stated *Language*: English
Index tests	*Index test*: VEGF, sFlt-1, TNF-α, CA-125 *Details of the index test procedure as stated*: serum concentrations of VEGF, sFlt-1, and TNF-α, were measured using specific commercial sandwich ELISA kit according to manufacturer protocols (Quantikine; R&D systems Inc, MN, USA); sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 46/70 (66%): stage I-II 15, stage III-IV 31; controls n = 24 *Reference standard*: laparoscopy/laparotomy N = 70 (100%) + histology *Description of positive case definition by reference standard as reported*: visual inspection, confirmed by histopathology in all patients; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	The pathogenesis of minimal-mild endometriosis and moderate-severe endometriosis seems to be different. Increased sFlt-1 levels in serum and urine of minimal-mild disease indicate that sFlt-1 may have an important role in inhibiting angiogenic process of the disease
Conflict of interest	Not reported
Notes	For VEGF and sFlt-1 there was no statistically significant difference between the groups - no data available for meta-analysis For TNF-α and CA-125 there was a statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Cho 2007 *(Continued)*

Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Colacurci 1996a

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the clinical utility of CA-125 in the diagnosis of endometriosis and to compare the sensitivity of the serum and peritoneal test as indicators of disease *Participants*: women undergoing laparoscopy for infertility *Selection criteria*: inclusion criteria: mid-follicular cycle phase *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*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 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: 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 *Threshold for positive result*: > 35 U/ml, pre-specified *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: Intra- and interassay CV was 4.3% and 7.7%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 18/40 (45%): stage I-II 10, stage III-IV 8; controls n = 22: normal pelvic - 12, other benign pathologies - 10 *Reference standard*: laparoscopy N = 40 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	Further investigations are needed to verify the sensitivity of serum and peritoneal CA-125 as diagnostic test for endometriosis using cut-off levels lower for serum and higher for peritoneal fluid, or different assays with high dilution of the samples

Colacurci 1996a *(Continued)*

Conflict of interest	Not reported
Notes	The reported data on CA-125 in peritoneal fluid is not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Colacurci 1996a *(Continued)*

Were all patients included in the analysis?	Yes

	Low

Da Silva 2014

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the presence of myeloperoxidase (MPO), N-acetyl-b-Dglucosaminidase (NAG), tumour necrosis factor alpha (TNF-α) and vascular endothelial growth factor (VEGF) in peripheral and menstrual blood in women with and without endometriosis *Participants*: women undergoing laparoscopy for infertility, pelvic pain or both, or for tubal ligation *Selection criteria*: inclusion criteria: regular menstrual cycles, no use of hormonal nor anti-inflammatory agents in the previous three months and surgical confirmation or exclusion of endometriosis in agreement with the ESHRE guidelines *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis group: infertility, pelvic pain or both; control group - infertility or request for tubal ligation; none of the women had a significant past medical history *Age*: median age 36 years, range 31-48 years *Number of participants enrolled*: 17 women *Number of participants available for analysis*: 17 women, all in follicular cycle phase (day 1-4) *Setting*: University Hospital: Hospital das Clinicas at Universidade Federal de Minas Gerais *Place of study*: Belo Horizonte, Minas Gerais, Brazil *Period of study*: February 2011 - December 2012 *Language*: English
Index tests	*Index test*: NAG, MPO, TNF-α, VEGF *Details of the index test procedure as stated*: Serum NAG and MPO activity were quantified by measuring the levels of the lysosomal enzyme NAG and by assaying MPO activity as previously reported (described and referenced to primary source; values expressed as change in absorbance (OD) at 400 nm and 450 nm, respectively); TNF-α and VEGF levels measured by using commercial specific ELISA kits (Human VEGF (Duoset R&D Systems DY293B range: 31,2-2000 pg/ml) and Human TNF-α (Duoset R&D Systems DY210, MN –USA, range: 15,6-1000 pg/ml); sample handling described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 10/17 (59%): stage II 5, stage IV 2, undetermined stage 3; controls n = 7 *Reference standard*: laparoscopy n = 17 (100%)

Da Silva 2014 *(Continued)*

	Description of positive case definition by reference standard test as reported: diagnosis according to ESHRE guidelines; staging according to rASRM *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: not specified, from the context - perioperative sample collection *Withdrawals*: none
Comparative	
Key conclusions by the authors	These findings point to the existence of an increased local inflammatory activity in women with endometriosis
Conflict of interest	Not reported; the work was supported by a grant from the Brazilian Research Council (CNPq) grant number 474132/2010-2
Notes	For NAG, MPO, TNF-α and VEGF there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in menstrual blood are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			

Da Silva 2014 *(Continued)*

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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Unclear	

Dayangan Sayan 2013

Study characteristics	
Patient sampling	*Primary objective*: to investigate the usefulness of serum IL-8 and the NLR, either by themselves or as adjuncts to CA-125, in the diagnosis of various stages of endometriosis *Participants*: women of reproductive age who were scheduled to undergo laparoscopy or laparotomy because of clinical indications of tubal ligation, benign ovarian cysts, infertility, or pelvic pain *Selection criteria*: inclusion criterion: follicular phase of menstrual cycle; exclusion criteria: hormonal medications for 6 months before surgery, ovarian neoplasia, PID, pregnancy, acute/chronic inflammation, autoimmune disease, refusal to participate, patients with suspected or confirmed leiomyoma or adenomyosis *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: mean age 26.8 ± 6.2 years (endometriosis), 27.4 ± 7.2 years (controls) *Number of participants enrolled*: 110 women *Number of participants available for analysis*: 100 women (all in follicular cycle phase) *Setting*: tertiary referral centre, Zekai Tahir Burak Women's Health Education and Research Hospital *Place of study*: Ankara, Turkey *Period of study*: March 2009 - April 2009 *Language*: English

Dayangan Sayan 2013 *(Continued)*

Index tests	*Index test*: CRP, CA-125, IL-8, Neutrophils, NLR *Details of the index test procedure as stated*: CRP levels were measured using immuno-turbidimetric assay (Hitachi 917/Tina Quant,Roche Diagnostics, Germany), CA-125 levels - using CA-125 II assay (ADVIA Centaur, Siemens, Los Angeles, USA), IL-8 - using IMMULITE 1000 (Siemens); assay sensitivity for CRP 0.003 mg/l, for CA-125 2 U/ml, for IL-8 0.7 pg/ml *Threshold for positive result*: WCC > 6400/ml, CA-125 > 29.9 IU/ml, IL-8 > 24 pg/ml, neutrophils > 4058/ml, NLR > 2.19, Combined marker > 43.1, not pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: Intra- and interassay CVs for CRP 0.2% and 2.5%; for CA-125 4.03%; IL-8 2.5% and 4.5%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 50/87 (58%): stage I-II 18, stage III-IV 32; controls n = 50 *Reference standard*: laparoscopy N = 87 (87%)/laparotomy N = 13 (13%) + histology *Description of positive case definition by reference standard test as reported*: visualisation at surgery with subsequent histological confirmation; staging according to the rAFS classification *Examiners*: trained surgeons who were skilled at detecting and identifying all forms of endometriotic lesions
Flow and timing	*Time interval between index test and reference standard*: blood samples were taken "prior to surgery", implies short time before surgery *Withdrawals*: 10 patients excluded from the study (did not meet inclusion criteria)
Comparative	
Key conclusions by the authors	Neutrophilia accompanied by a relative lymphocytopenia yielded an increased NLR in patients with endometriosis, and the data generated in our study show that a combination of putative inflammatory markers and CA-125 could serve as a multiple-marker screening test for endometriosis
Conflict of interest	Not reported
Notes	The reported diagnostic estimates for different stages of endometriosis are not included in this review For CRP there was no difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		

Dayangan Sayan 2013 *(Continued)*

		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Unclear	

De Placido 1998

Study characteristics

De Placido 1998 *(Continued)*

Patient sampling	*Primary objective*: to verify whether sHLA-I and sICAM-1 serum concentrations are related to the various stages of pelvic endometriosis, which is an immune-related disorder associated with impaired in-vitro NK cell activity *Participants*: women undergoing laparoscopy for various indications *Selection criteria*: not reported *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: indications for surgery: endometriosis group - infertility, pelvic pain, adnexal mass; controls - infertility, tubal ligation, mullerian malformation; none of the subjects was affected by any systemic or pelvic disease other than endometriosis *Age*: mean age 25.6 years, range 22-33 years (endometriosis group), 25.4 years, range 20-31 (controls) *Number of participants enrolled*: 30 women *Number of participants available for analysis*: 30 women (16 in follicular, 14 in luteal cycle phase) *Setting*: University Hospital: Department O&G, Universita degli studi di Napoli *Place of study*: Naples, Italy *Period of study*: not stated *Language*: English
Index tests	*Index test*: sICAM-1, sHLA-I *Details of the index test procedure as stated*: serum levels of sICAM-1 and sHLA-I were measured by using commercial ELISA kits (CD-54 ICAM-1: EIA PAC, Ancell Corp, Bayport, USA and sHLA-STAT Class I: SangStat Medical Corp, Menlo Park, USA); assay sensitivity for sICAM-1 5 ng/l, for sHLA-I CA-125 3 ng/ml *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: Intra- and interassay CVs < 12% for both assays
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 15/30 (50%): stage I-II 5, stage III-IV 10; controls n = 15 *Reference standard*: laparoscopy N = 30 (100%) + histology *Description of positive case definition by reference standard test as reported*: visualisation at surgery with subsequent histological confirmation; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Studies on sHLA-I and sICAM-1 may help to clarify the pathogenic mechanisms of endometriosis, and their serum concentrations may serve as additional markers for the early detection of recurrence of the disease during the monitoring of treatment outcome

De Placido 1998 *(Continued)*

Conflict of interest	Not reported
Notes	For sHLA-I and sICAM-1 there was no difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

De Placido 1998 *(Continued)*

Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Drosdzol-Cop 2012a

Study characteristics	
Patient sampling	*Primary objective*: to determine the role of serum and peritoneal interleukin (IL)-6, tumour necrosis factor (TNF)-α and glycodelin A levels as diagnostic markers of endometriosis in adolescent girls *Participants*: adolescent girls after menarche undergoing laparoscopy for chronic pelvic pain *Selection criteria*: inclusion criteria: chronic pelvic pain, defined as non-cyclic lower abdominal pain, not connected with the menstrual cycle, lasting at least 3 months or cyclic pain ongoing for 6 months, severe enough to cause functional disability or require medical or surgical treatment; exclusion criteria: general, chronic, autoimmune or endocrinological diseases, history of pregnancy or hormonal medications for at least 6 months prior to the study *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: Chronic pelvic pain; age of menarche (12.2 ± 1.4 and 12.8 ± 1.3 years) and percentage of ovulatory menstrual cycles (n = 15, 45.5% and n = 8, 47.1%) were comparable in both groups *Age*: mean age 17.4 ± 1.1 years (endometriosis group) and 16.4 ± 2.0 years (controls) *Number of participants enrolled*: 50 participants *Number of participants available for analysis*: 50 participants (all in follicular cycle phase, day 3-7) *Setting*: University Hospital: Woman's Health Institute, the Medical University of Silesia *Place of study*: Katowice, Poland *Period of study*: not stated *Language*: English
Index tests	*Index test*: IL-6, TNF-α and glycodelin A *Details of the index test procedure as stated*: Serum levels of IL-6, TNF-α and glycodelin A were measured by using commercial ELISA kits according to the manufacturers instructions; the detection limit of IL-6 was 2 pg/ml, of TNF-α was 0.7 pg/ml and of glycodelin A was 6 ng/ml; sample handling described *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: Intra- and interassay CVs were: for IL-6 4.3% and 4.9%, for TNF-α 6.5% and 3.9%, for glycodelin A 8.3% and 4.6%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 33/50 (66%): stage I-II 19, stage III-IV 14; controls n = 17 *Reference standard*: laparoscopy N = 50 (100%) + histology

Drosdzol-Cop 2012a *(Continued)*

	Description of positive case definition by reference standard test as reported: visualisation at surgery with subsequent histological confirmation; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	Studies on sHLA-I and sICAM-1 may help to clarify the pathogenic mechanisms of endometriosis, and their serum concentrations may serve as additional markers for the early detection of recurrence of the disease during the monitoring of treatment outcome
Key conclusions by the authors	At the cut-off value of 3.00 pg/ml, peritoneal TNF-α can be a reliable screening marker for the prediction of endometriosis in adolescents, giving a 14.6-fold higher probability of endometriosis detection in girls with chronic pelvic pain
Conflict of interest	The authors declared no conflict of interests
Notes	For IL-6, TNF-α and glycodelin A there was no difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**

Drosdzol-Cop 2012a *(Continued)*

DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Drosdzol-Cop 2012b

Study characteristics	
Patient sampling	*Primary objective*: to determine serum and peritoneal IL-2, IL-4, and monocyte chemotactic protein-1 levels as diagnostic markers of endometriosis in adolescent girls *Participants*: adolescent girls after menarche undergoing laparoscopy for chronic pelvic pain *Selection criteria*: inclusion criteria: chronic pelvic pain, defined as non-cyclic lower-abdominal pain, not connected with the menstrual cycle, lasting at least 3 months or cyclic pain ongoing for 6 months, severe enough to cause functional disability or require medical or surgical treatment; exclusion criteria: general, chronic, autoimmune or endocrinological diseases, history of pregnancy or hormonal medications for at least 6 months prior to the study *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: chronic pelvic pain; age of menarche (12.2 ± 1.4 and 12.8 ± 1.3 years) and percentage of ovulatory menstrual cycles (n = 15, 45.5% and n = 8, 47.1%) were comparable in both groups *Age*: mean age 17.4 ± 1.1 years (endometriosis group) and 16.4 ± 2.0 years (controls) *Number of participants enrolled*: 50 participants *Number of participants available for analysis*: 50 participants (allin follicular cycle phase, day 3-7)

Drosdzol-Cop 2012b *(Continued)*

	Setting: university Hospital: Woman's Health Institute, the Medical University of Silesia *Place of study*: Katowice, Poland *Period of study*: not stated *Language*: English
Index tests	*Index test*: IL-2, IL-4, and MCP-1 *Details of the index test procedure as stated*: serum levels of IL-2, IL-4 and MCP-1 were measured by using commercial ELISA kits according to the manufacturers instructions; the detection limit of IL-2 was 9.9 pg/ml, IL-4 was 1.2 pg/ml and MCP-1 was 2.3 pg/ml; sample handling described *Threshold for positive result*: IL-4 ≥ 3.00 pg/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: Intra- and interassay CVs were: for IL-2 - 5.2% and 8%, for IL-4 - 3.75% and 5.05%, for MCP-1 - 4.7% and 8.7%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 33/50 (66%): stage I-II 19, stage III-IV 14; controls n = 17 *Reference standard*: laparoscopy N = 50 (100%) + histology *Description of positive case definition by reference standard test as reported*: visualisation at surgery with subsequent histological confirmation; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	The serum IL-4, peritoneal IL-2 and IL-4 provided a good method of discrimination between subjects with endometriosis and controls
Conflict of interest	The authors declared no conflict of interests
Notes	For IL-2 and MCP-1 there was no difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		

Drosdzol-Cop 2012b *(Continued)*

Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Elgafor el Sharkwy 2013

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the diagnostic value of serum measurement of IL-6 combined with the presence of nerve fibres in the functional layer of endometrium for diagnosis of minimal-mild endometriosis

Elgafor el Sharkwy 2013 *(Continued)*

	Participants: women undergoing laparoscopy for evaluation of infertility, pelvic pain or both at the authors' institution *Selection criteria*: inclusion criteria: reproductive age (18-36 years), follicular cycle phase, regular menstrual cycle; exclusion criteria: any current infection (genital or systemic), any medication within 1/12 months prior to laparoscopy, previous surgery for endometriosis, smoking or drinking alcohol *Study design*: cross-sectional, single-gate design, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation* (n/N): dysmenorrhoea - 64/114; dyspareunia - 17/114; dyschezia - 6/114; pelvic pin - 35/114; infertility - 91/114 *Age*: mean age 31 ± 1.1 years (endometriosis group), 29 ± 0.6 years (controls) *Number of participants enrolled*: 114 women *Number of participants available for analysis*: 78 women (only minimal-mild endometriosis included; all in follicular cycle phase) *Setting*: Department of O&G, Zagazig University Hospital *Place of study*: Zagazig, Egypt *Period of study*: December 2010 - April 2012 *Language*: English
Index tests	*Index test*: IL-6 *Details of the index test procedure as stated*: serum IL-6 level using a commercially available ELISA (DRG, Germany); sample processing described *Threshold for positive result*: > 15.4 pg/ml, not pre-specified *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 74/114 (65%): stage I-II 38, stage III-IV 36; controls n = 40 *Reference standard*: laparoscopy n = 114 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rASRM classification *Examiners*: Three experienced gynaecologists in endometriosis
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained 'in the morning, the day before laparoscopy' *Withdrawals*: 36 participants with moderate-severe disease were not included in final analysis
Comparative	
Key conclusions by the authors	Combination of both serum IL-6 and presence of nerve fibres in the endometrium is more reliable method for diagnosis of minimal-mild endometriosis than in single test
Conflict of interest	The authors declared no conflict of interest

Elgafor el Sharkwy 2013 *(Continued)*

Notes	The reported data on endometrial biomarkers and combined endometrial-blood test are not presented in this review Only minimal-mild disease evaluated

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Elgafor el Sharkwy 2013 *(Continued)*

Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
	High

Fairbanks 2009

Study characteristics	
Patient sampling	*Primary objective*: to evaluate IL-12 and IL-18 levels in the serum and peritoneal fluid of women with and without endometriosis *Participants*: women who underwent laparoscopy for clinically suspected endometriosis *Selection criteria*: inclusion criteria: eumenorrhoea, age 18–40 years; exclusion criteria: any autoimmune disease, absence of peritoneal liquid at laparoscopy, the coexistence of any other causes of infertility, and any hormonal medications in the 3 months before surgery *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: severe dysmenorrhoea, deep dyspareunia, chronic pelvic pain, infertility, urinary symptoms (pain, bleeding or both) or cyclic bowel abnormalities (pain, bleeding or both) *Age*: range 18-40 years *Number of participants enrolled*: 105 women *Number of participants available for analysis*: 105 (85 in follicular, 20 in luteal cycle phase) *Setting*: endometriosis referral centre, School of Medicine, University of Sao Paulo *Place of study*: Sao Paulo, Brazil *Period of study*: February 2004 - December 2005 *Language*: English
Index tests	*Index test*: IL-12, IL-18 *Details of the index test procedure as stated*: serum IL-12 and IL-18 levels were measured using Human IL-12 (p70) kits (Human IL-12 p70 Kit, BD Biosciences, San Diego, CA), and ELISA (IL-18 ELISA, IBL, Hamburg, Germany); the measurement of IL-12 and IL-18 levels was performed after all data had been collected; the detection limits for IL-2 kit 4 pg/ml and for IL-18 kit 9.2 pg/ml; sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 72/105 (69%): stage I-II 28, stage III-IV 44; controls n = 33 *Reference standard*: laparoscopy N = 105 (100%) + histology

Fairbanks 2009 *(Continued)*

	Description of positive case definition by reference standard test as reported: visualisation at surgery with subsequent histological confirmation; staging according to the rAFS classification; surgical diagnostic criteria described in details for peritoneal, ovarian and DIE *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	Patients with severe endometriosis have higher IL-12 levels irrespective of IL-18 levels, suggesting that in this disease an alternative pathway is involved in induction of the Th1 immune response
Conflict of interest	Not reported
Notes	For IL-12 and IL-18 there was no difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			

Fairbanks 2009 *(Continued)*

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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Fassbender 2009

Study characteristics	
Patient sampling	*Primary objective*: to test the hypothesis that the plasma concentration of complement factor C3a (anaphylatoxin) can be used as a non-invasive test in the diagnosis of endometriosis *Participants*: women who had undergone laparoscopic surgery for subfertility, pelvic pain or both *Selection criteria*: not specified *Study design*: cross-sectional prospective single-gate design, non-consecutive enrolment
Patient characteristics and setting	*Clinical presentation*: infertility - 160, dysmenorrhoea - 26, hx of hormonal treatment, chronic PID or STI - nil; ethnicity: Caucasian - 136, other - 24 *Age*: median (range) 30 (18–46) years (endometriosis), 33 (20–46) years (controls) *Number of participants enrolled*: 160 women *Number of participants available for analysis*: 160 women (49 in menstrual, 55 in follicular, 56 in luteal cycle phase) *Setting*: Leuven University Fertility Centre *Place of study*: Leuven, Belgium *Period of study*: not stated *Language*: English
Index tests	*Index test*: C3a (anaphylatoxin)

Fassbender 2009 *(Continued)*

	Details of the index test procedure as stated: plasma concentration of C3a-des-Arg was determined with a commercially available immunoassay (Quidel Inc, San Diego, USA); quantification with a standard curve; sensitivity of this experiment was 34 ng/ml; sample handling and laboratory technique described in details *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: Intra-and interassay CV ranged 1.5%-2.8% and 11%-23%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 109/160 (68%); severity: stage I-II 54, stage III-IV 55; controls n = 51 *Reference standard*: laparoscopy N = 160 (100%) + histology *Description of positive case definition by reference standard test as reported*: visualisation at surgery with subsequent histological confirmation; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was collected immediately before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Our data do not confirm our hypothesis that C3a-des-Arg concentration in plasma can be used as a biomarker for the non-invasive diagnosis of endometriosis, but does not rule out the possibility that that measurement of complement activation at the level of the cervix or endometrium may be useful for this purpose
Conflict of interest	Not reported; supported by the Flemish fund for scientific research (FWO) & Leuven University Council (Dienst Onderzoekscoordinatie KU Leuven, Leuven, Belgium)
Notes	For C3a there was no difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			

Fassbender 2009 *(Continued)*

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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Fassbender 2012

Study characteristics	
Patient sampling	*Primary objective*: to test the hypothesis that differential surface-enhanced laser desorption/ionisation time-of-flight mass spectrometry protein or peptide expression in plasma can be used in infertile women with or without pelvic pain to predict the presence of laparoscopically and histologically confirmed endometriosis *Participants*: samples from women who had undergone laparoscopic surgery for subfertility, pelvic pain or both, stored in biobank *Selection criteria*: exclusion criteria: hormonal medications, surgery performed within 6 months before the time of sample collection

Fassbender 2012 *(Continued)*

	Study design: cross-sectional, single-gate design, prospective sample collection retrospective recruitment
Patient characteristics and setting	*Clinical presentation*: infertility - 240, dysmenorrhoea - 177, dyspareunia - 67, CPP - 30, dyschezia - 17, myoma - 16, irregular cycle - 40 *Age*: median age 31 years, range 23-44 years *Number of participants enrolled*: 254 women *Number of participants available for analysis*: 254 women (68 in menstrual, 98 in follicular, 88 in luteal cycle phase) *Setting*: Leuven University Fertility Centre *Place of study*: Leuven, Belgium *Period of study*: 2001-2009 *Language*: English
Index tests	*Index test*: proteome by SELDI-TOF-MS (five peptide and protein peaks, different for each cycle phase) *Details of the index test procedure as stated*: surface-enhanced laser desorption/ionisation coupled to time-of-flight mass spectrometry (plasma depletion by using Proteominer depletion kit, Bio-Rad); sensitivity of this experiment was 34 ng/ml; sample handling and procedure described in details; "training data set" (70%) was used to identify a pattern that discriminates between the presence and absence of disease and to construct the final least squares support vector machine model; "test data set" (30%) evaluated potential biomarkers - the final performance of model was averaged over 100 random splits *Threshold for positive result*: presence of specific protein peaks intensities, not pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: intra- and interassay CV ranged from 1.5% to 2.8% and 11% to 23%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 165/254 (65%): stage I-II 89, stage III-IV 76; controls - 89 *Reference standard*: laparoscopy N = 254 (100%) + histology *Description of positive case definition by reference standard test as reported*: visualisation at surgery with subsequent histological confirmation; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	A non-invasive test using proteomic analysis of plasma samples obtained during the menstrual phase enabled the diagnosis of endometriosis undetectable by ultrasonography with high sensitivity and specificity
Conflict of interest	The authors reported no conflicts of interest; supported by a number of grants
Notes	The diagnostic estimates were calculated separately for each menstrual cycle phase The diagnostic estimates for the validation test set are reported in this review The reported diagnostic estimates for different stages of endometriosis are not included in this review

Fassbender 2012 *(Continued)*

The reported diagnostic estimates for subgroup of ultrasound-negative endometriosis are not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between in-	Yes		

Fassbender 2012 *(Continued)*

dex test and reference standard?	
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Fedele 1989

Study characteristics	
Patient sampling	*Primary objective*: to assess the reliability of serum CA-125 in the detection of endometriosis in a large series of patients with different stages of the disease *Participants*: women undergoing laparoscopy for infertility, pelvic pain or both *Selection criteria*: not stated *Study design*: cross-sectional single-gate design, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: mean 30.9 years (endometriosis), 31.2 years (controls) *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 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum CA-125 was measured by immunoradiometric assay (Sorin Biomedica, Saluggia VC, Italy) *Threshold for positive result*: > 35 U/ml, pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 102/264 (39%): stage I-II 55, stage III-IV 47; controls n = 52 *Reference standard*: laparoscopy N = 264 (100%) + histology *Description of positive case definition by reference standard test as reported*: diagnosis based on endoscopic findings, histologic findings or both; staging according to the rAFS classification

Fedele 1989 *(Continued)*

	Examiners: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn immediately before surgery *Withdrawals*: 110 women did not have index test: "CA-125 was measured only in patients in whom endometriosis was found at laparoscopy and in patients with apparently normal pelvis"
Comparative	
Key conclusions by the authors	The usefulness of serum CA-125 measurements as an initial diagnostic tests is scanty. Because of its elevated specificity this test may be useful in indicating early surgical exploration of the pelvis in cases of infertility, dysmenorrhoea or both, which are associated with elevated CA-125
Conflict of interest	Not reported
Notes	—

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowl-	Yes		

Fedele 1989 *(Continued)*

edge of the results of the index tests?			
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Ferreira 1994

Study characteristics	
Patient sampling	*Primary objective*: to assess correlation between serum CA-125 levels and severity of endometriosis defined by rAFS and to establish diagnostic utility of this test in endometriosis *Participants*: women scheduled for laparoscopy or laparotomy for investigation of infertility *Selection criteria*: exclusion criteria: endocrine abnormalities, systemic disease, abnormal laboratory investigations, uterine fibroids, PID, pelvic pathology other than endometriosis identified at surgery *Study design*: cross-sectional, single-gate design, prospective sample collection
Patient characteristics and setting	*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 *Language*: Portuguese
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: 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 *Threshold for positive result*: > 16 U/ml and > 35 U/ml, pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard

Ferreira 1994 *(Continued)*

	Interobserver variability: Intra- and interobserver CV < 5.3% and < 7.5%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 36/54 (67%): stage I-II 14, stage III-IV 9; controls n = 18 *Reference standard*: laparoscopy/laparotomy N = 54 (100%) + histology *Description of positive case definition by reference standard test as reported*: diagnosis based on endoscopic findings, histologic findings or both; staging according to the rAFS classification; surgical procedure described *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn before surgery *Withdrawals*: 13 women were excluded because they met exclusion criteria
Comparative	
Key conclusions by the authors	In summary, the test is not sensitive enough for discrimination of women with and without endometriosis; observation across several cut-off points revealed that there was a significant lessening of specificity at the expense of sensitivity
Conflict of interest	Not reported
Notes	The reported diagnostic estimates for different stages of endometriosis are not included in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	No		

Ferreira 1994 *(Continued)*

		High	Low
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Ferrero 2005a

Study characteristics	
Patient sampling	*Primary objective*: to examine the presence and expression of vitamin D binding protein (DBP) in the peritoneal fluid (PF) and plasma (PL) of women with endometriosis *Participants*: women scheduled for laparoscopy for various indications *Selection criteria*: inclusion criteria: pre-menopausal (cycle length 21-35 days), no sign of pelvic inflammatory disease, no pregnancy, breastfeeding or abdominal surgery for the last 6 months, have not undergone hysterosalpingography in the 2 months prior to the surgical procedure *Study design*: cross-sectional two-gate design, prospective sample collection
Patient characteristics and setting	*Clinical presentation*:**s**tudy group: infertility - 39.8%, pelvic pain - 28.4%, dysmenorrhoea - 44.3%, dyspareunia - 22.7%, adnexal mass - 52.3%; controls: infertility - 57.5%, pelvic pain - 12.5%, dysmenorrhoea - 7.5%, tubal sterilisation - 30%; n = 17/145 women in study group were on OCP and were analysed separately *Age*: mean age 32.1 ± 5.0 years (endometriosis group), 32.6 ± 6.2 years (controls) *Number of participants enrolled*: 145 women *Number of participants available for analysis*: 145 women (76 in follicular and 69 in luteal menstrual cycle phase)

Ferrero 2005a *(Continued)*

	Setting: university hospital, San Martino Hospital, University of Genoa and St. Bartholomew's Hospital, St Bartholomew's School of Medicine and Dentistry *Place of study*: Genoa, Italy and London, UK *Period of study*: not reported *Language*: English
Index tests	*Index test*: DBP *Details of the index test procedure as stated*: plasma DPB expression was assessed by using 2-D PAGE (referenced to the previously published method): provisional identification was performed by matching with the human plasma 2-D PAGE protein map of ExPASy and subsequently confirmed by western blotting onto PVDF membranes (Hybond-P, Amersham Pharmacia Biotech) that was performed at 30 V for 18 hours using Towbin's transfer buffer; sample handling and laboratory methods described *Threshold for positive result*: not provided *Examiners*: 2 independent investigators who were blinded to the clinical status of the patients *Interobserver variability*: the interassay CV was < 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 105/145 (72%): stage I-II 43, stage III-IV 62; controls - 40 *Reference standard*: laparoscopy N = 145 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection with subsequent histological confirmation in all patients; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	The decreased level of DBPE in the PF but not in PL of women with untreated endometriosis suggests that this molecule may be relevant in the pathogenesis of this disease
Conflict of interest	Not reported
Notes	For DBP there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Ferrero 2005a *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Florio 2007

Study characteristics

Florio 2007 *(Continued)*

Patient sampling	*Primary objective*: to assess the diagnostic performance of urocortin determination in distinguishing endometriomas from other benign ovarian cysts *Participants*: women who underwent laparoscopic excision of ovarian cysts *Selection criteria*: not stated (only severe endometriosis included) *Study design*: cross-sectional single-gate design, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: ovarian cyst - 80 women, chronic pelvic pain - 20 women *Age*: mean age 34.1 ± 7.4 years (endometriosis), 35.2 ± 7.2 years (controls) *Number of participants enrolled*: 80 women *Number of participants available for analysis*: 80 women (menstrual cycle phase not specified) *Setting*: University of Siena academic hospital *Place of study*: Siena, Italy *Period of study*: March 2004 - January 2006 *Language*: English
Index tests	*Index test*: urocortin, CA-125 *Details of the index test procedure as stated*: plasma urocortin levels were measured in a blinded fashion in a single assay according to published methodology (referenced to the original source) with delayed addition of tracer to improve assay sensitivity (˜50 pg/ml); serum CA-125 concentration was assessed by Cobas Core CA 125 enzyme-immunoassay analysis kit (Roche, Basel, Switzerland) with assay sensitivity < 1 U/l; procedure and sample handling described *Threshold for positive result*: Urocortin > 33 pg/ml and 29 pg/ml; CA-125 > 36U/l and 30 U/l, not prespecified *Examiners*: no information provided; blinded to the results of reference standard *Interobserver variability*: Intra- and interassay CV for urocortin < 8%, for CA-125 < 5.6% and < 7.8%
Target condition and reference standard(s)	*Target condition*: endometriosis (ovarian and ovarian + pelvic) *Prevalence of target condition in the sample*: n = 40/80 (50%): all stage III-IV; controls n = 40 *Reference standard*: laparoscopy N = 80 (100%) + histology *Description of positive case definition by reference standard test as reported*: surgical visualisation and histopathology, staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn before anaesthesia for laparoscopy *Withdrawals*: none
Comparative	
Key conclusions by the authors	Immunolocalisation of urocortin and its higher levels in the cystic content than in peritoneal fluid and plasma suggest that it may be secreted by the endometriotic tissue. Urocortin is a sensitive and specific marker for the differential diagnosis of endometrioma compared with other benign ovarian cysts

Florio 2007 *(Continued)*

Conflict of interest	The authors have no potential conflicts of interest to disclose; supported in part by grant # 2004068714-004 from the Italian Ministry of University and Scientific Research (MURST) and the University of Siena
Notes	The reported diagnostic estimates for subgroup of endometrioma with no peritoneal implants are not included in this review For CA-125 - the cohort overlaps with Florio 2009, but a different threshold is presented, hence it is included as separate evaluation

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**

Florio 2007 *(Continued)*

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Florio 2009

Study characteristics	
Patient sampling	*Primary objective*: to quantify the concentration of follistatin and CA-125 in the serum of women with ovarian endometrioma and other benign cysts; to evaluate the follistatin levels in the cystic content and PF of a subset of patients with ovarian endometriotic cyst; to investigate the use of follistatin as a marker in the differential diagnosis of benign ovarian cysts *Participants*: women who underwent laparoscopic excision of benign ovarian cysts detected by ultrasound *Selection criteria*: inclusion criteria: reproductive age, persistent, large (> 5 cm) or complex pelvic mass without evidence of malignancy or pelvic pain not responding to medication; exclusion criteria: use of steroid hormones during the past 3 months, known pituitary, thyroid, renal, liver or adrenal disorders (only severe endometriosis included) *Study design*: cross-sectional single-gate design prospective sample collection
Patient characteristics and setting	*Clinical presentation*: ovarian cyst - 104 women, regular menstrual cycle - 90%, nulliparous - 100%; symptoms and other history not specified *Age*: mean age 34.0 ± 6.0 years (endometrioma), 32.0 ± 4.0 years (controls) *Number of participants enrolled*: 104 women *Number of participants available for analysis*: 104 women (all in follicular phase of menstrual cycle) *Setting*: University of Siena academic hospital *Place of study*: Siena, Italy *Period of study*: September 2004 - August 2006 *Language*: English
Index tests	*Index test*: follistatin, CA-125 *Details of the index test procedure as stated*: follistatin concentrations were measured in duplicates using a commercially available enzyme-linked immunosorbent assay (ELISA) with assay detection limit of 29 pg/ml (range 250 to 16 000 pg/ml); serum CA-125 concentration was assessed by Cobas Core CA 125 enzyme-immunoassay analysis kit (Roche, Basel, Switzerland) with assay sensitivity < 1 U/l; sample handling and laboratory technique for Follistatin described *Threshold for positive result*: urocortin levels: > 33 pg/ml and 29 pg/ml; CA-125 levels > 36 U/l and 30 U/l, not pre-specified

Florio 2009 *(Continued)*

	Examiners: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: Intra- and interassay CV for follistatin < 3.0 and < 9.0%; for CA-125 < 5.6% and < 7.8%
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 52/104 (50%): all stage III-IV 52; controls n = 52 *Reference standard*: laparoscopy N = 104 (100%) + histology *Description of positive case definition by reference standard test as reported*: surgical visualisation and histopathology, staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples collected immediately before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	In conclusion, serum follistatin levels are increased in women with ovarian endometriosis. Follistatin seems to fulfil the requirements of sensitivity, specificity and reproducibility in order to become a useful clinical marker of late stage ovarian endometriosis. Further studies, including a blind validation in a cohort series, will be required to support the clinical use of follistatin in the diagnosis of endometriosis.
Conflict of interest	Not reported; the work was supported by grants from the Italian Ministry of University and Scientific Research (MURST) and the University of Siena
Notes	Originally, this was a a two-gate design study, which also includes healthy controls (N = 27) and women with non-ovarian endometriosis (N = 11), these groups seem to be separately enrolled and the data for these groups or for the whole cohort are not available - not included in the review The reported diagnostic estimates for 'Endometrioma versus no ovarian cyst' are not included in this review, because number of participants and analysed subgroups are unclear For CA-125 - the cohort overlaps with Florio 2007, but different threshold is presented, hence included as separate evaluation

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			

Florio 2009 *(Continued)*

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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Foda 2012

Study characteristics

Foda 2012 *(Continued)*

Patient sampling	*Primary objective*: to determine the clinical usefulness of IL-6, TNF-α, CA-125, Hs-CRP and VEGF levels in infertile women with pelvic pain as markers of the early stages of peritoneal endometriosis during which imaging is not effective *Participants*: infertile women complaining of chronic pelvic pain undergoing laparoscopy *Selection criteria*: inclusion criteria: reproductive age, regular menstrual cycles, do not smoke or drink alcohol; exclusion criteria: age > 35 years, any current infection (genital or systemic), any medication within 1 month prior to laparoscopy, minimal amount or bloody peritoneal fluid, patients using IUD *Study design*: cross-sectional single-gate design prospective sample collection
Patient characteristics and setting	*Clinical presentation*: infertility - 95 women; dysmenorrhoea - 37 women; dyspareunia - 15; dyschezia - 9; pelvic/abdominal pain - 43; menorrhagia - 22; urinary symptoms - 10 *Age*: range 18-35 years *Number of participants enrolled*: 95 women *Number of participants available for analysis*: 95 women (all in follicular phase of menstrual cycle, days 5–10) *Setting*: Department of O&G, Mansoura University Hospital *Place of study*: Mansoura, Egypt *Period of study*: January 2009 - May 2010 *Language*: English
Index tests	*Index test*: IL-6, CA-125, TNF-α, Hs-CRP VEGF *Details of the index test procedure as stated*: IL-6 and TNF-α levels were estimated by using a commercially available enzyme-linked immunosorbent assay (ELISA, DRG, Germany); CA-125 and Hs-CRP were measured by automated electro-chemiluminescent immunoassay instrument (Elecsys 2010, Roche, Germany); VEGF was determined by a competitive enzyme immunoassay technique using Accucyte human VEGF kit; lower detection limit of IL-6, CA-125, TNF-α, Hs-CRP & VEGF kits were 2 pg/ml, < 1 IU/ml, 2.2 pg/ml, 65 ng/ml and 5 pg/ml respectively; sample collection and storage described *Threshold for positive result*: IL-6 > 12.2 pg/ml; CA-125 > 17.6 IU/ml; TNF-α > 12.45 pg/ml; Hs-CRP >438 µg/ml; VEGF > 236 pg/ml; the thresholds were not pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: interassay CV for IL-6, CA-125, TNF-α, Hs-CRP & VEGF kits were < 4%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 65/95 (68%); stage I-II 37, stage III-IV 28; controls n = 30 *Reference standard*: laparoscopy N = 95 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was collected the day before laparoscopy *Withdrawals*: none
Comparative	

Foda 2012 *(Continued)*

Key conclusions by the authors	Serum IL-6 and TNF-α levels can be used to discriminate between patients with or without endometriosis. Also, minimal-mild endometriosis patients display higher serum IL-6 and TNF-α level than moderate-severe endometriosis or the control cases; this sheds light on markers of the early stages of the disease. CA-125, VEGF and Hs-CRP appear to be advantageous only for the diagnosis of severe endometriosis and positively correlate with the stage of the disease; very low levels might serve as a marker for an absence of endometriosis.
Conflict of interest	The authors have no potential conflicts of interest to disclose
Notes	The reported cost analysis: cost of the markers per case was about EGP 110, much less than the costs of the hospital stay and diagnostic laparoscopy

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowl-	Yes		

Foda 2012 *(Continued)*

edge of the results of the index tests?			
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Franchi 1993

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the value and potential use of CA-125 determinations in the diagnosis and management of endometriosis *Participants*: patients of reproductive age undergoing laparotomy or laparoscopy for pelvic mass *Selection criteria*: not provided *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic mass, not specified *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 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum CA-125 levels assessed by radioimmunoassay; sample processing and laboratory technique not described *Threshold for positive result*: > 35 IU/ml, pre-specified

Franchi 1993 *(Continued)*

	Examiners: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 37/120 (31%): stage I-II 13, stage III-IV 24; controls - 9 *Reference standard*: laparoscopy/laparotomy N = 120 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was collected "immediately before surgery" *Withdrawals*: 74 women were excluded from analysis (only patients with endometriosis and patients with normal pelvis were included)
Comparative	
Key conclusions by the authors	Serum CA-125 levels correlated significantly with disease severity, but the low sensitivity of the test precludes its use as a screening procedure for endometriosis
Conflict of interest	Not reported
Notes	The reported diagnostic estimates per degree of severity of endometriosis are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it prespecified?	Yes		

Franchi 1993 *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Gagne 2003a

Study characteristics	
Patient sampling	*Primary objective*: to determine whether the proportion of several leukocyte subsets is modulated in the endometrium of patients with endometriosis and, if so, whether it can be used for diagnostic purposes *Participants*: women who were scheduled to undergo laparoscopy or laparotomy at 1 of the 8 clinical institutions in the Montreal area *Selection criteria*: inclusion criteria: 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. *Study design*: multicentre study of two-gate design, prospective recruitment, random sample of patients (participation rate 94%)
Patient characteristics and setting	*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 mean age of 37.3 ± 6.4 years *Number of participants enrolled*: 368 women

Gagne 2003a *(Continued)*

	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 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: 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 described in details. The bootstrap method validation was performed by drawing 200 replicate samples with replacement from the original data set *Threshold for positive result*: CA-125 > 12.8 U/ml and > 35 U/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: Inter- and intra-assay variations < 5%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 173/368 (47%): stage I-II 78%, stage III-IV 22%; controls n = 195 *Reference standard*: laparoscopy/laparotomy N = 368 (100%) *Description of positive case definition by reference standard test as reported*: cases were defined by the presence of endometriotic lesions confirmed at the time of surgical examination; staging according to the ASRM system *Examiners*: gynaecologists collaborating in the study were trained surgeons experienced with the management of endometriosis who were skilled in detecting and identifying all forms of endometriotic lesions
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	The predictive model represents a novel diagnostic tool to identify women with a high likelihood of suffering from endometriosis
Conflict of interest	All the authors except RM are (or were) employees of PROCREA BioSciences; supported by the Industrial Research Assistance Program (IRAP) from NSERC grant #15453Q and internal resources at PROCREA BioSciences
Notes	The reported diagnostic estimates of the predictive model based on the combination of blood and endometrial test with clinical and demographic data are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Gagne 2003a *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Gagne 2003a *(Continued)*

	Low

Gagne 2003b

Study characteristics	
Patient sampling	*Primary objective*: to determine whether high levels of VEGF could also be found in the serum of patients with endometriosis *Participants*: women who were scheduled to undergo laparoscopy or laparotomy at 1 of the 8 clinical institutions in the Montreal area *Selection criteria*: inclusion criteria: pre-menopausal age, no past pregnancy, luteal phase of the menstrual cycle (based on the last period and further confirmed by histology), regular cycles (21-35 days), no acute salpingitis, no hormonal treatment or IUD in previous 3 months. *Study design*: multicentre study of two-gate design, prospective recruitment, random sample of patients (participation rate > 90%)
Patient characteristics and setting	*Clinical presentation*: infertility (11% controls, 28% cases); pain (19% controls, 34% cases); tubal ligation (60% controls, 26% cases); hysterectomy (18% controls, 35% cases); diagnostic laparoscopy (22% controls, 39% cases); history of acute infections (30% controls, 34% cases); smoking (63% controls, 53% cases) *Age*: sampling from a population with mean age of 37.3 ± 6.4 years *Number of participants enrolled*: 277 women *Number of participants available for analysis*: 277 women (all in luteal cycle phase) *Setting*: biotech firm - MetrioGene BioSciences (a subsidiary of PROCREA BioSciences) *Place of study*: Montreal, Canada *Period of study*: July 1997 - May 2001 *Language*: English
Index tests	*Index test*: VEGF *Details of the index test procedure as stated*: serum VEGF levels were measured using a commercially available ELISA kit (R&D Systems, Minneapolis, MN) according to the manufacturer's instruction; assay sensitivity <9.0 pg/ml; sample handling and laboratory procedure described in details *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: Inter- and intra-assay CV <10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 131/277 (47%): stages I-IV, numbers not specified; controls n = 146 *Reference standard*: laparoscopy/laparotomy N = 277 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the ASRM classification

Gagne 2003b *(Continued)*

	Examiners: gynaecologists collaborating in the study were trained surgeons experienced with the management of endometriosis who were skilled in detecting and identifying all forms of endometriotic lesions
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	Although VEGF seems to play a pivotal role locally in the implantation and development of endometriotic lesions, the disease is not associated with a significant modulation in the levels of circulating VEGF
Conflict of interest	Not reported (the authors' affiliation is MetrioGene BioSciences, a biotech firm)
Notes	For VEGF there was no difference between the groups - no data available for meta-analysis

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Gagne 2003b *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Gazvani 1998

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the role of IL-8 in the pathogenesis of endometriosis in relation to the stage of disease *Participants*: patients undergoing laparoscopic surgery for benign gynaecological indications *Selection criteria*: not specified *Study design*: cross-sectional, two-gate design, prospective collection of samples, consecutive patients
Patient characteristics and setting	*Clinical presentation*: indications for surgery: abdominal pain (n = 21), sterilisation (n = 11), infertility (n = 18); none of the patients had been on medication at least 1 month prior to the laparoscopy and none was on any long-acting drugs *Age*: mean age 28 ± 8.1 years (endometriosis group) and 29 ± 6.9 years (controls) *Number of participants enrolled*: 50 women *Number of participants available for analysis*: 47 (23 in follicular, 24 in luteal cycle phase) *Setting*: not specified, the authors' affiliations are 2 university hospitals: Liverpool Women's Hospital, University of Liverpool and Department of O&G, University of Aberdeen *Place of study*: Aberdeen and Liverpool, UK *Period of study*: not provided *Language*: English
Index tests	*Index test*: IL-8 *Details of the index test procedure as stated*: IL-8 levels were measured using an enzyme-linked immunosorbent assay (CYTokit Red; CYTimmune Sciences, USA) according to the manufacturer's instructions; sample processing described

Gazvani 1998 *(Continued)*

	Threshold for positive result: not provided *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 25/105 (24%): stage I-II 14, stage III-IV 11; controls n = 22 *Reference standard*: laparoscopy N = 105 (100%) + histology *Description of positive case definition by reference standard test as reported*: visualisation at surgery: the condition of tubes, ovaries, pouch of Douglas, and bowels were inspected; staging according to the AFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: not specified, but from the context, samples were obtained at surgery *Withdrawals*: 3 patients were excluded before analysis because of inadequate peritoneal fluid sample
Comparative	
Key conclusions by the authors	Peripheral blood concentrations did not correlate with peritoneal fluid concentrations of IL-8 or the presence of endometriosis. IL-8 (in PF) is an important factor that may contribute to the pathogenesis of endometriosis possibly by promoting neovascularisation
Conflict of interest	Not reported
Notes	For IL-8 there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		

Gazvani 1998 *(Continued)*

If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Unclear	

Glitz 2009

Study characteristics	
Patient sampling	*Primary objective*: to evaluate IL-18 levels in the serum and peritoneal fluid of infertile women with minimal-mild endometriosis in order to determine association of IL-18 with infertility *Participants*: women with minimal or mild endometriosis submitted to laparoscopy to investigate infertility (endometriosis group) and patients who underwent laparoscopy for tubal ligation (controls) *Selection criteria*: inclusion criteria: first menstrual phase, no hormonal medications for at least 3 months prior to surgery *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified; all controls were fertile and none had a significant past medical history *Age*: mean age 31.51 ± 4.54 years (endometriosis group) and 34.23 ± 3.56 years (controls)

Glitz 2009 *(Continued)*

	Number of participants enrolled: 78 women *Number of participants available for analysis*: 78 women (in follicular phase of menstrual cycle) *Setting*: Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul *Place of study*: Porto Alegre, Brazil *Period of study*: March 2006 - December 2007 *Language*: English
Index tests	*Index test*: IL-18 *Details of the index test procedure as stated*: serum IL-18 levels were measured using the Human IL-18 ImmunoAssay ELISA kit (MBL Co.Ltd, Japan); assay sensitivity 12.5 pg/ml, minimal estimated detection 12.5 ± 6.25 pg/ml *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 56/78 (72%): all stage I-II; controls n = 22 *Reference standard*: laparoscopy N = 78 (100%) *Description of positive case definition by reference standard test as reported*: Endometriosis was diagnosed by visualisation at surgery; staging according to the rAFS classification *Examiners*: the same investigator performed all endoscopic procedures
Flow and timing	*Time interval between index test and reference standard*: blood was collected at the time of the laparoscopy *Withdrawals*: none
Comparative	
Key conclusions by the authors	Women with minimal-mild endometriosis did not show any alteration in the concentration of IL-18 in serum or peritoneal fluid
Conflict of interest	Not reported; supported by CNPq, Fundo de Incentivo à Pesquisa (FIPE) do Hospital de Clínicas de Porto Alegre, CAPES and FAPERGS
Notes	For IL-18 there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Glitz 2009 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Gogacz 2014

Study characteristics

Gogacz 2014 *(Continued)*

Patient sampling	*Primary objective*: to investigate the presence of T regulatory cells (Tregs) in the peripheral blood (PB) and peritoneal fluid (PF) in females with endometriosis *Participants*: women who underwent laparoscopy for suspected endometriosis or infertility investigation *Selection criteria*: not specified *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis group - not specified; controls - unexplained infertility; all participants had regular menstrual cycles *Age*: mean age 33.58 ± 4.74 years (endometriosis group) and 31.2 ± 5.9 years (controls) *Number of participants enrolled*: 42 women *Number of participants available for analysis*: 42 women (in follicular phase of menstrual cycle, days 9-12) *Setting*: University hospital: Department of Gynaecology, Medical University of Lublin *Place of study*: Lublin, Poland *Period of study*: not stated *Language*: English
Index tests	*Index test*: Tregs, WBC, lymphocytes *Details of the index test procedure as stated*: Tregs in peripheral blood were assessed by analysing expression of CD4 and CD25 cell surface antigens, and intracellular FOXP3 antigen using a BD FACSCalibur flow cytometer (BD Biosciences, San Jose, USA); the percentage of CD4+ CD25+ FOXP3+ Tregs in the CD4+ T lymphocyte subpopulation was determined using the Human Treg Flow™ kit (FOXP3 Alexa Fluor® 488/CD4 PE-Cy5/CD25 PE) from BioLegend (San Diego, USA); WBC and lymphocyte counts were determined by using a peroxidase method with ADVIA 2120 system (Siemens) *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 22/42 (53%): stage I-II 15, stage III-IV 7; controls n = 20 *Reference standard*: laparoscopy N = 42 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was collected at the time of the laparoscopy *Withdrawals*: none
Comparative	

Gogacz 2014 *(Continued)*

Key conclusions by the authors	The local host-defence mechanism is deficient in patients with endometriosis, thus endometriosis should not be treated as an autoimmune condition
Conflict of interest	Not reported
Notes	For Tregs, WBC, lymphocytes there was no statistically significant difference between the groups - no data available for meta-analysis For CA-125 there was statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**

Gogacz 2014 *(Continued)*

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Goluda 1998

Study characteristics	
Patient sampling	*Primary objective*: to establish the concentration of the adhesion molecules (ICAM-1 and E-Selectin) in the sera and peritoneal fluids of women with endometriosis in comparison to the control group *Participants*: women who underwent laparoscopy for infertility and pelvic pain *Selection criteria*: inclusion criteria: luteal phase of menstrual cycle (only minimal-mild endometriosis included) *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility, pelvic pain *Age*: range 26-40 years (endometriosis group) and 20-42 years (controls) *Number of participants enrolled*: 20 women *Number of participants available for analysis*: 20 women (all in in luteal phase of menstrual cycle) *Setting*: 2nd Department & Gynaecological Clinic of Medical Academy in Wroclaw *Place of study*: Wroclaw, Poland *Period of study*: March 2006 - December 2007 *Language*: English
Index tests	*Index test*: ICAM-1 and E-Selectin *Details of the index test procedure as stated*: the levels of sICAM-1 and sE-selectins were measured using ELISA (R&D wg) according to the manufacturers protocol *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis

Goluda 1998 *(Continued)*

	Prevalence of target condition in the sample: n = 11/20 (55%), all stage I-II; controls n = 9 *Reference standard*: laparoscopy N = 20 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	We did not find any significant differences between the two examined groups, although further studies should be carried out
Conflict of interest	Not reported
Notes	For ICAM-1 and E-Selectin there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it prespecified?	No		
Was a cycle phase considered in interpretation of the result of index test?			
		High	**Low**
DOMAIN 3: Reference Standard			

Goluda 1998 *(Continued)*

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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Gorai 1993

Study characteristics	
Patient sampling	*Primary objective*: to investigate endometrial antigens involved in the autoimmunity of endometriosis *Participants*: women who underwent laparoscopy or laparotomy *Selection criteria*: not presented *Study design*: cross-sectional, unclear if single or two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not presented; none of the study subjects were on oral contraceptives or other hormones such as danazol or GnRH agonists *Age*: range 20-46 years *Number of participants enrolled*: 36 women *Number of participants available for analysis*: 36 women (phase of menstrual cycle not specified) *Setting*: University Hospital: Department O&G, Yokohama City University School of Medicine *Place of study*: Yokohama, Japan *Period of study*: not specified *Language*: English
Index tests	*Index test*: anti-endometrial antibodies

Gorai 1993 *(Continued)*

	Details of the index test procedure as stated: the expression of anti-endometrial antibodies was tested by using Western Blot analysis (endometrial antigens were prepared from endometrium of 6 fertile women without endometriosis collected at hysterectomy according to Coulam and Ryan method; anti-human immunoglobulin, biotinylated whole antibody from sheep was used to detect antibodies bound to the endometrial antigens); sample handling and laboratory technique described in detail *Threshold for positive result*: positive test was defined when distinct dark bands were seen on the blot for at least one antibody; threshold not pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 18/36 (50%): stage I-II 4, stage III-IV 14; controls n = 18 *Reference standard*: laparoscopy/laparotomy N = 36 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Autoantibodies reactive against endometrial antigens are present in patients with endometriosis
Conflict of interest	Not reported; the work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan
Notes	For anti-endometrial antibodies with MW of 28, 38, 64 kDa there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Unclear		
		Unclear	**Unclear**

Gorai 1993 *(Continued)*

DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Guerriero 1996a

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the accuracy of CA-19.9 plasma levels (with or without CA-125 levels) combined with transvaginal ultrasonography in the differential diagnosis of endometriosis *Participants*: women undergoing laparoscopy or laparotomy for persistent adnexal mass at the authors' institution *Selection criteria*: inclusion criteria: pre-menopausal, non-pregnant (only moderate-severe endometriosis included)

Guerriero 1996a *(Continued)*

	Study design: cross-sectional, single-gate design, prospective recruitment and collection of samples, consecutive series
Patient characteristics and setting	*Clinical presentation*: pelvic mass - 100%, infertility - 53% *Age*: mean age 33.3 ± 9.6 years *Number of participants enrolled*: 118 women *Number of participants available for analysis*: 118 women (only moderate-severe endometriosis included; all in follicular cycle phase) *Setting*: Department of O&G, University of Cagliari *Place of study*: Cagliari, Italy *Period of study*: November 1994 - November 1995 *Language*: English
Index tests	*Index test*: CA-19.9, CA-19.9 + CA-125 *Details of the index test procedure as stated*: serum CA-125 levels assessed by immunoradiometric assay (CIS Bio International, Gif sur Yvette, France), limit of detection 0.5 U/ml; serum CA-19.9 levels assessed by immunoradiometric assay (CIS Bio International, Gif sur Yvette, France), limit of detection 1.5 U/ml; sample processing and laboratory technique not described *Threshold for positive result*: CA-125: ≥ 25 U/ml, pre-specified; CA-19.9 ≥12 U/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: Intra- and interassay CV for CA-125 3.9% and 4.2%; for CA-19.9 4.6% and 5.3%
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 39/118 (33%): all stage III-IV; controls n = 79 *Reference standard*: laparoscopy n = 99/laparotomy n = 19 (N = 118, 100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection with careful assessment of the ovaries, followed by histopathological diagnosis; surgical staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was collected on the day of surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Transvaginal ultrasonography used alone is the most cost-effective method in the preoperative differential diagnosis of endometrioma
Conflict of interest	Not reported
Notes	The reported diagnostic estimates for combination of blood test with ultrasound are not presented in this review The diagnostic estimates were available only for combination of CA-125 with CA-19.9 and for either 1 of the 2 positive markers

Guerriero 1996a *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Guerriero 1996a *(Continued)*

	Low

Guerriero 1996b

Study characteristics	
Patient sampling	*Primary objective*: to assess the role of transvaginal ultrasonography combined with CA-125 plasma levels in the diagnosis of endometrioma *Participants*: women undergoing laparoscopy or laparotomy for persistent adnexal mass at the authors' institution *Selection criteria*: inclusion criteria: pre-menopausal, non-pregnant (only moderate-severe endometriosis included) *Study design*: cross-sectional, single-gate design, prospective recruitment and collection of samples, consecutive series
Patient characteristics and setting	*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 *Period of study*: November 1993 - October 1994 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: 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 *Threshold for positive result*: 3 pre-selected cut-offs: ≥ 20 U/ml, ≥ 25 U/ml, ≥ 35 U/ml *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: Intra- and interassay CV 3.9% and 4.2%
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 29/101 (29%): all stage III-IV; controls n = 72 *Reference standard*: laparoscopy/laparotomy + histology *Description of positive case definition by reference standard test as reported*: visual inspection with careful assessment of the ovaries, followed by histopathological diagnosis; visual inspection confirmed on histopathology; histological criteria reported; surgical procedure described; surgical staging according to the rAFS classification *Examiners*: no information provided

Guerriero 1996b *(Continued)*

Flow and timing	*Time interval between index test and reference standard*: blood was collected on the day of surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Transvaginal ultrasonography used alone has a better predictive capacity in differentiating endometrioma from other adnexal masses than combined methods
Conflict of interest	Not reported
Notes	The reported diagnostic estimates for combination of blood test with ultrasound are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		

Guerriero 1996b *(Continued)*

		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Gurgan 1990

Study characteristics	
Patient sampling	*Primary objective*: to evaluate possible value of peritoneal fluid CA-125 levels as a more sensitive marker of minimal (stage I) endometriosis when compared to serum levels measured simultaneously *Participants*: women undergoing laparoscopy as part of infertility work-up or tubal sterilisation *Selection criteria*: exclusion criteria: patients with more advanced endometriosis (> stage I) or other pathological findings *Study design*: cross-sectional study of two-gate design, prospective recruitment
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: mean age 30.1 ± 2.6 years (endometriosis), 27.9 ± 2.6 years (controls) *Number of participants enrolled*: 38 women *Number of participants available for analysis*: 38 women (all in mid-secretory phase of menstrual cycle) *Setting*: Department of Obstetrics and Gynaecology, Faculty of Medicine, University of Hacettepe *Place of study*: Sihiye-Ankara, Turkey *Period of study*: October 1988 - June 1989 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: CA-125 in serum and PF was measured in duplicates using an immunoradiometric assay assay (ELISA CA-125, Compagnie ORIS Industrie, France); assay sensitivity 2.4 U/ml; sample handling and laboratory procedure described *Threshold for positive result*: > 16 U/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: Inter- and intra-assay CV 5.7%-8.1% and 2%-10%

Gurgan 1990 *(Continued)*

Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 17/38 (45%) all stage I; controls n = 21 *Reference standard*: laparoscopy N = 38 (100%) *Description of positive case definition by reference standard test as reported*: classification according to the ASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	CA-125 levels have been found to be mildly, but not significantly elevated in sera of patients with minimal endometriosis; laparoscopic evaluation remains the most reliable method of diagnosis of minimal endometriosis
Conflict of interest	Not reported
Notes	The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			

Gurgan 1990 *(Continued)*

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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Gurgan 1999

Study characteristics	
Patient sampling	*Primary objective*: to examine whether IGF-I, IGF-II and IGFBP 3 in serum and peritoneal fluid correlate with the presence and severity of endometriosis *Participants*: patients undergoing laparoscopy for various indications *Selection criteria*: exclusion criteria: any other pelvic pathology (myoma uteri, ovarian mass or adhesions not secondary to endometriosis), blood-contaminated PF sample, other medical problems and/or using any medication for at least the last six months before laparoscopy *Study design*: cross-sectional, two-gate design, prospective collection of samples, consecutive patients
Patient characteristics and setting	*Clinical presentation*: indications for surgery: infertility, pelvic pain and tubal sterilisation *Age*: mean age 30.8 ± 5.4 years (stage I-II endometriosis), 32 ± 4.2 years (stage III-IV endometriosis), 31.7 ±6.7 years (controls) *Number of participants enrolled*: 44 women *Number of participants available for analysis*: 44 (21 in follicular, 23 in luteal cycle phase) *Setting*: O&G Department, Hacettepe University Hospital *Place of study*: Sihiye-Ankara, Turkey *Period of study*: not stated *Language*: English
Index tests	*Index test*: IGF-I, IGF-II and IGFBP 3

Gurgan 1999 *(Continued)*

	Details of the index test procedure as stated: serum levels of IGF-I and II and IGFBP 3 were measured by using immunoradiometric assay kits (Diagnostic System Laboratories,Texas); assay sensitivities were 0.8 ng/ml, 0.13 ng/ml and 0.5 ng/ml, respectively; sample handling described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: the intra- and interassay CV of IGF-I and II and IGFBP 3 assays were 3.4%, 4.3%, 1.8% and 8.2%, 9.5%, 1.9%, respectively
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 29/44 (66%): stage I-II 15, stage III-IV 14; controls n = 15 *Reference standard*: laparoscopy N = 44 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rASRM classification *Examiners*: all the procedures were performed by a single operator (the first author)
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	IGF-I is most probably associated with late-stage endometriosis and may be an important mediator in progression to late-stage disease. IGF-I may also act as a local factor in persistence of endometriotic implants in mild cases
Conflict of interest	The authors declared no conflict of interest
Notes	For IGF-II and IGFBP3 there was no difference between the groups - no data available for meta-analysis For IGF-I there was statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**Low**

Gurgan 1999 *(Continued)*

DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?			
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Hallamaa 2012

Study characteristics	
Patient sampling	*Primary objective*: to investigate whether serum HE4 concentration varies within the normal menstrual cycle and whether common gynaecological hormonal treatments have an effect on HE4 values *Participants*: patients undergoing laparoscopy for suspected endometriosis or tubal ligation *Selection criteria*: exclusion criteria: suspicion of malignancy, pregnancy or infection *Study design*: cross-sectional, two-gate design, prospective collection of samples

Hallamaa 2012 *(Continued)*

Patient characteristics and setting	*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 *Language*: English
Index tests	*Index test*: HE4, CA-125 *Details of the index test procedure as stated*: serum HE4 and CA-125 concentrations were analysed by ELISA analysis (Fujirebio Diagnostics inc, Malvern, PA, USA) according to the manufacturer's instructions *Threshold for positive result*: For HE4 not provided, for CA-125 > 35 U/l, not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 123/175 (70%): stages I-IV, the number of participants per each stage not reported; controls n = 52 *Reference standard*: laparoscopy N = 175 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected 24 h before surgery *Withdrawals*: 5 women were excluded when endometrial biopsy was non-conclusive regarding cycle phase
Comparative	
Key conclusions by the authors	HE4 measurement in healthy pre-menopausal women as well as in women with endometriosis can be carried out at any phase of the menstrual cycle, and irrespective of hormonal medication, extending the benefits of HE4 use in clinical practice
Conflict of interest	One of the authors received lecture honoraria from several pharmaceutical companies; other authors declared no conflict of interest; the study was supported by the Finnish Funding Agency for Technology and Innovation (projects 40343/05 and 599/05); Hormos Medical Ltd, Finland (subsidiary of QuatRx Pharmaceutical, USA); Biotop Oy, Finland; Genolyze Oy, Finland
Notes	For HE4 there was no difference between the groups - no data available for meta-analysis
Methodological quality	

Hallamaa 2012 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Hallamaa 2012 *(Continued)*

	Low

Hapangama 2008

Study characteristics	
Patient sampling	*Primary objective*: to assess endometrial expression of the human telomerase enzyme and telomere length (TL) *Participants*: patients undergoing laparoscopy for suspected endometriosis or tubal ligation *Selection criteria*: inclusion criteria: pre-menopausal women (18–46 years), regular menstrual cycle (25–31 day), no hormonal treatments *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis - not specified, controls - healthy fertile women requesting tubal ligation *Age*: mean age 37 ± 5 years (endometriosis group) and 38 ± 5 years (controls) *Number of participants enrolled*: 56 women *Number of participants available for analysis*: 56 (all in luteal menstrual cycle phase) *Setting*: School of Reproductive and Developmental Medicine, University of Liverpool, Liverpool Women's Hospital *Place of study*: Liverpool, UK *Period of study*: not reported *Language*: English
Index tests	*Index test*: TL, progesterone, E2 *Details of the index test procedure as stated*: peripheral blood TL expression was assessed by using RT-PCT (extracted from peripheral mononuclear cells, reaction by SYBR green chemistry, measured on iCycler RT PCR system (Bio-Rad Laboratories, Hercules, USA), expressed in base pairs); sample handling and laboratory method described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 29/56 (52%): stage I-II 14, stage III-IV 15; controls n = 27 *Reference standard*: laparoscopy N = 56 (100%) *Description of positive case definition by reference standard test as reported*: surgical diagnosis; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before surgery (personal communication with the author)

Hapangama 2008 *(Continued)*

	Withdrawals: the data for TL was not available for 6 participants (9%); reason not explained
Comparative	
Key conclusions by the authors	We speculate that aberrant endometrial expression of telomerase mediates alterations in cell fate that enhance proliferation, contributing to the pathogenesis of endometriosis
Conflict of interest	Not reported; the work was supported by a RDF grant from the University of Liverpool and RCOG millennium grant
Notes	For TL, progesterone and E2 there was no difference between the groups - no data available for meta-analysis The data for markers measured in eutopic endometrium are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		

Hapangama 2008 *(Continued)*

		Unclear	Low
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Harada 2002

Study characteristics	
Patient sampling	*Primary objective*: to investigate the clinical value of the serum CA-19.9 level in comparison with the serum CA-125 level for diagnosing endometriosis *Participants*: patients who underwent laparotomy or laparoscopy with the preoperative diagnosis of infertility, myoma uteri, adenomyosis or endometriosis (cases) and patients who underwent laparoscopy for infertility investigation (controls) *Selection criteria*: exclusion criteria: patients with malignant tumours or inflammatory disease *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: mean age 35.4 ± 6.7 years, range 21-52 years *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 *Language*: English
Index tests	*Index test*: CA-125, CA-19.9 *Details of the index test procedure as stated*: serum CA-19.9 and CA-125 levels were measured by enzyme immunoassay (TFB Co,Tokyo, Japan) and were expressed in arbitrary units based on a primary reference standard *Threshold for positive result*: CA-19.9 > 37.0 U/ml, CA-125 > 35.0 U/ml, pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not stated

Harada 2002 *(Continued)*

Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 101/123 (82%); stage I-II 38, stage III-IV 63; controls n = 22 *Reference standard*: laparoscopy/laparotomy N = 123 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected from all before the operation *Withdrawals*: none
Comparative	
Key conclusions by the authors	The mean serum CA-19.9 levels in patients at all stages of endometriosis were significantly higher than those in patients without endometriosis and significantly correlated with the rASRM classification scores. CA-19.9 levels and serum CA-125 levels may prove to be valuable tools for predicting the severity of endometriosis as diagnosed by laparoscopy
Conflict of interest	Not reported; the study was supported by a Science Research Grant (11671599) from the Ministry of Education, Culture, Sports, Science and Technology of Japan
Notes	The reported data enabled to calculate diagnostic estimates for the subgroups by severity of endometriosis - not included in the review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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?	Yes		

Harada 2002 *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Hassa 2009

Study characteristics	
Patient sampling	*Primary objective*: to investigate the changes in Th1 and Th2 immune responses, characterised by a change in the levels of IL-2, IL-4, IL-10 and IFN-γ, and determinations of T helper, T suppressor, NK, and B cells in peripheral blood and peritoneal fluid of different stages of endometriosis *Participants*: patients who underwent laparoscopy for pain or infertility (cases) and for tubal ligation (controls) *Selection criteria*: exclusion criteria: any medical treatment employed prior to laparoscopy that may interfere with the results *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: controls had no history of infertility and no pelvic pathology during surgical inspection *Age*: mean 30.9 ± 5.6; 29.9 ± 6.7 years (endometriosis stage I-II; III-IV), 30.1 ± 6.7 years (controls) *Number of participants enrolled*: 97 women *Number of participants available for analysis*: 97 women (all in follicular phase of menstrual cycle)

Hassa 2009 *(Continued)*

	Setting: O&G Department, Eskisehir Osmangazi University School of Medicine *Place of study*: Eskisehir, Turkey *Period of study*: 2003–2005 *Language*: English
Index tests	*Index test*: IL-2, IL-4, IL-10, IFN-γ, and lymphocytes: Th, Ts, AL and NK *Details of the index test procedure as stated*: cytokines were measured by using ELISA assay (Cellular Communication Investigations, Beckman Coulter, USA); lymphocytes were assessed by using cluster determinant-3 (CD-3), CD4, CD8, CD25, CD28, CD45, CD16, CD23, Abs against early T cell activation antigens such as CD45RA/CD45RO, CD-69 and late activation antigens such as HLA-DR; sensitivity limits of the kits were 5 pg/ml, 5 pg/ml, 5 pg/ml, and 0.08 pg/ml for IL-2, IL-4, IL-10, IFN-γ; sample handling and technique described *Threshold for positive result*: not provided *Examiners*: experienced technicians blind to the status of cases at laboratory conducted the detection of both cytokine and immune cell levels *Interobserver variability*: Intra- and interassay CVs were < 10% for all assays
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 60/97 (62%): stage I-II 42, stage III-IV 18; controls n = 37 *Reference standard*: laparoscopy N = 97 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed histopathologically; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: not clearly stated, but from the context, blood was collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The result of this study did not show any significant difference in peripheral blood and peritoneal fluid cytokine and lymphocyte subgroups between normal women and those with early and late stage endometriosis
Conflict of interest	All the authors had a conflict of interest (financial or otherwise)
Notes	For IL-2, IL-4, IL-10, IFN-γ, Th, Ts, AL, NK there was not statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Hassa 2009 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Unclear	

Hornstein 1995

Study characteristics

Hornstein 1995 *(Continued)*

Patient sampling	*Primary objective*: to compare the serum CA-125 concentrations determined by assays in women with and without endometriosis, and to determine if the new assay improves the clinical utility of CA-125 in the diagnosis of endometriosis *Participants*: patients with the preoperative diagnosis of endometriosis, pelvic pain, or infertility recruited from 2 fertility units *Selection criteria*: not specified *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*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 *Place of study*: Boston, MA, USA and Camden, NJ, USA *Period of study*: not stated *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: 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 *Threshold for positive result*: CA-125 > 35.0 U/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: the intra- and interassay CVs were 8.3% and 12.1% for the older assay and 5.2% and 7.5% for the new CA-125 assay
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 74/123 (60%); stage I-II 54, stage III-IV 20; controls n = 49 *Reference standard*: laparoscopy N = 123 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rAFS classification *Examiners*: no information provided; the operating surgeon was not sure of the patients' CA-125 concentration at the time of surgery
Flow and timing	*Time interval between index test and reference standard*: blood drawn one menstrual cycle preceding laparoscopy *Withdrawals*: none
Comparative	
Key conclusions by the authors	The sensitivity and specificity were slightly improved using the new CA-125 assay; however, this assay did not dramatically improve detection of endometriosis

Hornstein 1995 *(Continued)*

Conflict of interest	Not reported; the work was supported in part by a grant from Centocor, Inc, Malvern, PA, USA
Notes	Only the diagnostic estimates for a new generation assay were included in this review because they were the closest to the currently used technique The reported diagnostic estimates for stage III-IV endometriosis are not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			

Hornstein 1995 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Inagaki 2003

Study characteristics	
Patient sampling	*Primary objective*: to assess whether the IgG anti-laminin-1 auto-Abs in infertile patients are associated with reproductive disorders, particularly during pre- and peri-implantation stages *Participants*: infertile patients who underwent laparoscopy or laparotomy as part of their infertility investigation *Selection criteria*: not specified *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility *Age*: mean age 33.7 years, range: 26-45 years *Number of participants enrolled*: 68 women *Number of participants available for analysis*: 68 women (menstrual cycle phase not specified) *Setting*: Okayama University Hospital and at Nagoya City University Hospital *Place of study*: Okayama and Nagoya, Japan *Period of study*: not stated *Language*: English
Index tests	*Index test*: IgG anti-laminin-1 auto-Abs *Details of the index test procedure as stated*: detection of IgG anti-laminin-1 Abs was performed using ELISA (referenced to the original source); laboratory technique described in details *Threshold for positive result*: 1.0 U/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: the inter and intra-assay CV < 3.1% and 6.9%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 42/68 (62%); stage I-II 14, stage III-IV 28; controls n= 26 *Reference standard*: laparoscopy/laparotomy N = 68 (100%) + histology

Inagaki 2003 *(Continued)*

	Description of positive case definition by reference standard test as reported: visual inspection confirmed by histopathology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: not provided, but context suggests peri-operative sampling *Withdrawals*: none
Comparative	
Key conclusions by the authors	The assessment of IgG anti-laminin auto-Abs might prove useful for the diagnosis and medical treatment of endometriosis
Conflict of interest	Not reported
Notes	The presented data enabled calculation of the diagnostic estimates according to severity of endometriosis - not included in this review We did not consider a group of separately recruited healthy controls (N = 39) that did not have surgery and were not included in the calculations of the diagnostic estimates

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			

Inagaki 2003 *(Continued)*

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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Unclear	

Iwasaki 1993

Study characteristics	
Patient sampling	*Primary objective*: to evaluate cell-mediated immunity in endometriosis *Participants*: women who underwent laparoscopy or laparotomy for infertility or benign adnexal mass *Selection criteria*: not specified *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility or adnexal mass *Age*: mean age 34.8 ± 6.9 years (endometriosis group), 32.3 ± 3.8 years (controls) *Number of participants enrolled*: 45 women *Number of participants available for analysis*: 45 women (all in mid-follicular menstrual cycle phase) *Setting*: Department of O&G, School of Medicine, Keio University *Place of study*: Keio, Japan *Period of study*: not stated *Language*: English
Index tests	*Index test*: lymphocyte subsets and NK activity *Details of the index test procedure as stated*: subsets of lymphocytes in peripheral blood were analysed with flow cytometry FACS scan by using several combinations of monoclonal Abs (Bec-

Iwasaki 1993 *(Continued)*

	ton Dickinson, CA); NK cytotoxicity was assessed in K562 cell line; sample handling and laboratory methods described *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 19/45 (42%); stage I-II 16, stage III-IV 3; controls n= 26 *Reference standard*: laparoscopy/laparotomy N = 45 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	An alteration in cell-mediated immunity may be among the pathogenetic, or developing, factors in endometriosis
Conflict of interest	Not reported
Notes	For suppressor-T cells, cytotoxic-T cells, activated-T cells and NK activity, there was a statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review For T-lymphocytes, B-lymphocytes, inducer-T cells, helper-T cells, non-MHC restricted T cells and NK cells, there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			

Iwasaki 1993 *(Continued)*

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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Jee 2008

Study characteristics	
Patient sampling	*Primary objective*: to assess whether sCD163 and IL-6 could be used as serum markers for discriminating ovarian endometriomas from other benign ovarian masses *Participants*: women who had adnexal cystic tumours and underwent adnexal surgery either via laparoscopy or laparotomy *Selection criteria*: exclusion criteria: patients who had ≥ 2 pathologic diagnoses, recent history of any inflammatory disease (only moderate-severe endometriosis included) *Study design*: cross-sectional single-gate, prospective collection of samples

Jee 2008 *(Continued)*

Patient characteristics and setting	*Clinical presentation*: dysmenorrhoea - 54.5% of women with endometrioma; not specified otherwise *Age*: reproductive age (values presented for each type of ovarian neoplasm) *Number of participants enrolled*: 95 women *Number of participants available for analysis*: 95 women (menstrual cycle phase not specified) *Setting*: Seoul National University Bundang Hospital *Place of study*: Seoul, Korea *Period of study*: July 2003 - November 2004 *Language*: English
Index tests	*Index test*: sCD163 and IL-6 *Details of the index test procedure as stated*: serum levels of sCD163 and IL-6 were determined with a commercial ELISA kit (Soluble CD163 ELISA; Cedarlane Laboratories, Canada) and IL-6 ELISA kit (DuoSet ELISA Development System; R&D System Inc, USA) according to the manufacturer's instructions; assay sensitivity for sCD163 is 0.15 ng/ml, for IL-6, 0.7 pg/ml; sample processing and laboratory techniques described in details *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: intra- and interassay CV sCD163 < 5%; for IL-6, 2.5% and 4.5%
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 44/95 (46%), all stage III-IV; controls n = 51 *Reference standard*: laparoscopy/laparotomy n = 95 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rAFS classification; histopathology of the specimens was proven by pathologists *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected on the day before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Serum levels of sCD163 as well as IL-6 are not useful markers for ovarian endometriomas.
Conflict of interest	Not reported
Notes	For sCD163 and IL-6 there was no difference between the groups - no data available for meta-analysis For CA-125 there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review

Jee 2008 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Jee 2008 *(Continued)*

	Low

Jia 2013

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the feasibility of using plasma microRNAs as a non-invasive diagnostic test for the detection of endometriosis *Participants*: women who underwent laparoscopy for various indications, including pelvic masses, pelvic pain, infertility and uterine leiomyoma *Selection criteria*: exclusion criteria: postmenopausal status, previous hormonal use within 3 months, adenomyosis or malignancy (only moderate-severe endometriosis included) *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: indications for surgery: pelvic pain, infertility, pelvic mass, uterine fibroids *Age*: mean age: 34.1 ± 5.03, range: 25–44 years (endometriosis); 32.1 ± 6.95 years, range 22–45 years (controls) *Number of participants enrolled*: 46 women *Number of participants available for analysis*: 40 women (31 in follicular and 9 in luteal cycle phase) *Setting*: Department of O&G, Peking Union Medical College Hospital *Place of study*: Beijing, PR China *Period of study*: January 2012 - May 2012 *Language*: English
Index tests	*Index test*: miR-17-5p, miR-20a and miR-22 *Details of the index test procedure as stated*: plasma miRNA expression by RT-PCR (normalised to miR-16 levels and calculated using the $2^{-\Delta\Delta Ct}$ method); sample processing and laboratory technique described in details *Threshold for positive result*: miR-17-5p: 0.9057, miR-20a: 0.6879, miR-22: 0.5647; not pre-specified *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 20/40 (50%): all stage III-IV; controls n = 20 *Reference standard*: laparoscopy N = 46 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection with a thorough inspection of the abdominopelvic cavity to detect any typical or atypical endometriotic lesion; all possible lesions were excised and sent for pathological examination; staging according to the rASRM system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was collected "immediately before administration of anaesthesia"

Jia 2013 *(Continued)*

	Withdrawals: 6 samples (3 endometriosis and 3 controls) were used for preliminary screening experiment and were not included in the final single-plex analysis
Comparative	
Key conclusions by the authors	Plasma miR-17-5p, miR-20a and miR-22 are down-regulated in women with endometriosis, which raises the potential clinical utility of plasma microRNA profiling in endometriosis diagnosis
Conflict of interest	The authors declared no conflict of interest; supported by grants from the National Natural Science Foundation of China (81170548) and Key Project for Clinical Faculty Foundation, Ministry of Health, China
Notes	The reported data for combination of miRs was insufficient to construct 2 x 2 tables and hence are not resented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted	Yes		

Jia 2013 *(Continued)*

without knowledge of the results of the index tests?			
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Joshi 1986

Study characteristics	
Patient sampling	*Primary objective*: to determine whether a concentration of PEP or other specific proteins in serum or PF is altered in endometriosis and, if so, whether this alteration is associated with development of an antibody response *Participants*: untreated pre-menopausal women who underwent diagnostic laparoscopy for infertility, dysmenorrhoea or tubal ligation *Selection criteria*: not specified *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: Pre-menopausal, not specified *Number of participants enrolled*: 55 women *Number of participants available for analysis*: 55 women (35 in proliferative, 20 in secretory cycle phase) *Setting*: not stated; the authors' affiliations: Department of O&G Albany Medical College and Baylor College of Medicine *Place of study*: Albany, NY and Houston, TX, USA *Period of study*: not stated *Language*: English
Index tests	*Index test*: PEP and total proteins in follicular and luteal phase of menstrual cycle *Details of the index test procedure as stated*: PEP in serum was assessed with a specific RIA; protein profiles were examined by polyamide gel electrophoresis (SDS PAGE); specific assays were developed to detect and quantify anti-PEP and anti-EG; sample processing and laboratory techniques described in details *Threshold for positive result*: not provided

Joshi 1986 *(Continued)*

	Examiners: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 36/55 (66%): stage I-II 21, stage III-IV 15; controls n = 19 *Reference standard*: laparoscopy N = 55 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: not specified, from context - blood samples were collected short time before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Levels of PEP were not different in serum from women with moderate-severe or mild endometriosis or from disease-free cycling controls
Conflict of interest	Not reported
Notes	For PEP there was no statistically significant difference between the groups - no data available for meta-analysis The data for total proteins were reported only for peritoneal fluid - not assessed in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		

Joshi 1986 *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Unclear	

Kalu 2007

Study characteristics	
Patient sampling	*Primary objective*: to compare the concentration of the cytokines: IL-6, IL-8, IL-1β, VEGF, TNF-α, MCP-1, RANTES, PDGF, sFas, sFasL in both biological fluids (PF and serum) in women with and without endometriosis *Participants*: women undergoing laparoscopy for unexplained infertility *Selection criteria*: exclusion criteria: active PID, hydrosalpinges, any autoimmune disease, hormonal treatment or hysterosalpingography in the 2 months preceding laparoscopy, pregnancy in the last 6 months (only minimal-mild endometriosis included) *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility *Age*: mean 31.0 ± 6.5 years (endometriosis group) and 30.5 ± 6 years (controls) *Number of participants enrolled*: 57 women *Number of participants available for analysis*: 40 or 35 women - number of participants varied for different assays (all in luteal phase of menstrual cycle)

Kalu 2007 *(Continued)*

	Setting: Assisted Conception Unit, St Helier University Hospital *Place of study*: Carshalton, Surrey, UK *Period of study*: not stated *Language*: English
Index tests	*Index test*: IL-6, IL-8, IL-1β, TNF-α, RANTES, PDGF, VEGF, MCP-1, sFasL, sFas *Details of the index test procedure as stated*: PDGF, sFas, RANTES, MCP-1 were determined in duplicate by quantitative sandwich EIA using commercial Quantikine kits (R&D systems, USA); the sensitivity was 15 pg/ml, 20 pg/ml and 8 pg/ml, 4.7 pg/ml, respectively. FasL was determined in duplicate by EIA kits (Diaclone, France); sensitivity < 12.5 pg/ml. IL-6, IL-8, IL-1β and TNF-α were determined using an 'IMMULITE' analyser (DPC, USA); sensitivity was 5.0 pg/ml, 2 pg/ml, 1.5 pg/ml, 1.7 pg/ml, respectively. VEGF was determined by the Neogen (Lexington, USA) immunoassay, limit of detection was 18.6 pg/ml; sample handling described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 26/57 (46%), all stage I-II; controls n = 31 *Reference standard*: laparoscopy N = 57 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection: positive diagnosis was defined as red endometriotic lesions - red vesicles, red flame-like lesions or gland-like lesions; staged according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained before anaesthesia *Withdrawals*: samples of 8-11 participants from control group and 9-11 participants from endometriosis group were not available (number of missing samples varied for each assay due to limitation of sample quantity)
Comparative	
Key conclusions by the authors	The elevated levels of MCP-1, IL-6, and IL-8 in peritoneal fluid but not serum may indicate the importance of local macrophage activating factors in the pathogenesis of endometriosis
Conflict of interest	Not reported
Notes	For IL-6, IL-8, IL-1β, TNF-α, RANTES, PDGF and sFas, there was no statistically significant difference between the groups - no data available for meta-analysis For VEGF, MCP-1, sFasL the data was not available (insufficient sample to assay)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Kalu 2007 *(Continued)*

Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	Yes		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		

Kalu 2007 *(Continued)*

	High

Khan 2006

Study characteristics	
Patient sampling	*Primary objective*: to examine the peritoneal fluid (PF) and serum concentrations of hepatocyte growth factor (HGF) in different r-ASRM staging and morphologic appearances of endometriosis in an attempt to determine whether HGF can be clinically useful to predict the activity of pelvic endometriosis *Participants*: women undergoing laparoscopy for infertility, pelvic pain or benign ovarian mass *Selection criteria*: exclusion criteria: controls - fibroid uterus, PID *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility, dysmenorrhoea or pelvic pain (endometriosis), benign ovarian cyst (controls); none of the participants had been on hormonal medication in the 3/12 months prior to surgery; all women had regular menstrual cycles (28/32 days) *Age*: range 15-43 years (endometriosis group) and 17-39 years (controls) *Number of participants enrolled*: 194 women *Number of participants available for analysis*: 58 women (21 in follicular and 37 in luteal phase of menstrual cycle) *Setting*: Department of O&G, Nagasaki University School of Medicine, Nagasaki Municipal Hospital *Place of study*: Nagasaki, Japan *Period of study*: not stated *Language*: English
Index tests	*Index test*: HGF *Details of the index test procedure as stated*: serum concentrations of HGF were measured using a commercially available ELISA kit (Quantikine, R&D system, Minneapolis, MN); the limit of detection was 40.0 pg/ml *Threshold for positive result*: not provided *Examiners*: no information provided; assay was performed in blind fashion *Interobserver variability*: the intra- and interassay CV were < 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 37/57 (65%): stage I-II 19, stage III-IV 18; controls n = 21 *Reference standard*: laparoscopy N = 57 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed on histopathology: peritoneal lesions of endometriosis were diagnosed according to published criteria (referenced to the primary source) and categorised as red, black, and white lesions; peritoneal lesions and chocolate cysts were measured; staged according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained at surgery

Khan 2006 *(Continued)*

	Withdrawals: 136 participants did not consent for blood collection and were not included in the study
Comparative	
Key conclusions by the authors	Women with early or advanced endometriosis as measured by rASRM scoring system are not associated with an increase in either serum or PF concentrations of HGF. Rather HGF levels in serum and PF were significantly increased in women harbouring blood-filled red peritoneal lesions and may be clinically useful to predict the activity of pelvic endometriosis
Conflict of interest	Not reported
Notes	For HGF there was no statistically significant difference between the groups - no data available for meta-analysis The data for HGF measured in peritoneal fluid are not presented in this review The data for HGF expression stratified by type of endometriotic lesions, severity of endometriosis or cycle phase are not included in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly	Yes		

Khan 2006 *(Continued)*

classify the target condition?			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Khan 2012

Study characteristics	
Patient sampling	*Primary objective*: to measure PGE2 levels in different body fluids; namely MF, PF and sera derived from women with and without endometriosis and to investigate effect of PGE2 on the replication of *E. coli* in a bacteria culture and on growth of PBLs derived from women with and without endometriosis *Participants*: women undergoing laparoscopy for infertility, pelvic pain or benign ovarian mass *Selection criteria*: exclusion criteria: induced menstrual cycles *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility, dysmenorrhoea or pelvic pain (endometriosis), benign ovarian cyst (controls); none of the participants had been on hormonal medication in the 3/12 months prior to surgery; all women had regular menstrual cycles (28/32 days) *Age*: mean age 30.2 ± 3.5 years, range 20-42 years (endometriosis group); 28.4 ± 3.9 years, range 18-32 years (controls) *Number of participants enrolled*: 86 women *Number of participants available for analysis*: 86 women (30 in proliferative, 47 in secretory and 9 in menstrual cycle phase) *Setting*: Department of O&G, Nagasaki University School of Medicine, Saiseikai Nagasaki Hospital *Place of study*: Nagasaki, Japan *Period of study*: not stated

Khan 2012 *(Continued)*

	Language: English
Index tests	*Index test*: PGE2 *Details of the index test procedure as stated*: serum concentrations of PGE2 were measured using ELISA (Quantikine, R&D system, Minneapolis, MN); the limit of detection was 8.25 pg/ml *Threshold for positive result*: not provided *Examiners*: no information provided; assay was performed in blind fashion *Interobserver variability*: the intra- and interassay CV were < 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 58/86 (67%): stage I-II 35, stage III-IV 23; controls n = 28 *Reference standard*: laparoscopy n = 86 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed on histopathology; staged according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: not specified, the context suggests that the samples were collected short time before surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	PGE2 promotes bacterial growth in women with endometriosis
Conflict of interest	The authors declared no conflict of interests; the work was supported by grants-in-aid for Scientific Research (grant no. 16591671 and 18591837) from the Ministry of Education, Sports, Culture, Science and Technology of Japan
Notes	For PGE2 there was no statistically significant difference between the groups - no data available for meta-analysis The data for HGF measured in menstrual blood or peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		

Khan 2012 *(Continued)*

		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Unclear	

Khan 2013

Study characteristics	
Patient sampling	*Primary objective*: to measure measure the HSP70 levels in sera, menstrual and peritoneal fluid collected from women with and without endometriosis, to examine the role of LPS in the production of HSP70 by eutopic endometrium and to investigate the effects of LPS and HSP70 on the production of cytokines by peritoneal macrophages in endometriosis *Participants*: women undergoing laparoscopy for infertility, pelvic pain or benign ovarian mass

Khan 2013 *(Continued)*

	Selection criteria: not specified *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility, dysmenorrhoea or pelvic pain (endometriosis), benign ovarian cyst (controls); none of the participants had been on hormonal medication in the 3/12 months prior to surgery; all women had regular menstrual cycles (28/32 days) *Age*: mean age 29.8 ± 4.6 years, range 20-42 years (endometriosis group); 28.6 ± 3.8 years, range 18-32 years (controls) *Number of participants enrolled*: 63 women (16 in proliferative, 31 in secretory and 12 in menstrual cycle phase) *Number of participants available for analysis*: 50 women *Setting*: Department of O&G, Nagasaki University School of Medicine, Saiseikai Nagasaki Hospital *Place of study*: Nagasaki, Japan *Period of study*: not stated *Language*: English
Index tests	*Index test*: HSP70 *Details of the index test procedure as stated*: serum concentrations of HSP70 were measured using a commercially available ELISA (StressXpressTM, EKS-700; Stressgen Victoria, Canada) according to the manufacturer's instructions; the limit of detection was 200 pg/ml *Threshold for positive result*: not provided *Examiners*: no information provided; assay was performed in blind fashion *Interobserver variability*: the intra- and interassay CV were < 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 43/63 (68%): stage I-II 28, stage III-IV 15; controls n = 20 *Reference standard*: laparoscopy N = 63 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed on histopathology; staged according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before and at surgery *Withdrawals*: 13 participants (21%) were not included in the analysis, presumably blood samples were not available
Comparative	
Key conclusions by the authors	A crosstalk between local inflammation and tissue stress reaction in the pelvic environment may be involved in TLR4-mediated growth of endometriotic cells
Conflict of interest	The authors declared no conflict of interests; the work was supported by grants-in-aid for Scientific Research (grant no. 16591671 and 18591837) from the Ministry of Education, Sports, Culture, Science and Technology of Japan

Khan 2013 *(Continued)*

Notes	For HSP70 there was no statistically significant difference between the groups - no data available for meta-analysis The data for HGF measured in menstrual blood, peritoneal fluid and eutopic endometrium are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index	Yes		

Khan 2013 *(Continued)*

test and reference standard?	
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	No
	High

Khanaki 2012

Study characteristics	
Patient sampling	*Primary objective*: to compare serum phospholipid fatty acid profile in endometriosis patients with controls, and to explore the correlation of this profile with the severity of the disease *Participants*: women undergoing laparoscopy or laparotomy for various indications *Selection criteria*: exclusion criteria: anti-inflammatory drugs during 3/12 months before surgery, any diseases (endometritis, gastrointestinal or urological disease with pelvic pain, liver or endocrine autoimmune disease, previous endometriosis or neoplastic disorders and chronic PID) *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified; surgical diagnosis in controls: uterine myoma, dermoid cyst, serous cyst, paraovarian cyst or mucinous cyst; all women had regular menstrual cycles *Age*: mean age 30.57 ± 5.04 years (endometriosis group) and 30.57 ± 5.71 years (controls) *Number of participants enrolled*: 138 women *Number of participants available for analysis*: 138 women (in proliferative or secretory cycle phase) *Setting*: university Hospital: Alzahra Hospital, Tabriz University of Medical Sciences and Sarem Hospital *Place of study*: Tabriz, Iran and Tehran, Iran *Period of study*: not stated *Language*: English
Index tests	*Index test*: phospholipid fatty acids *Details of the index test procedure as stated*: serum phospholipid fatty acids were purified of the total phospholipids by using TLC technique and measured using a gas chromatograph (Buck Scientific model 610, USA); the relative amount of each fatty acid was stated as the percentage of total area on chromatograms *Threshold for positive result*: not provided *Examiners*: no information provided; assay was performed in blind fashion *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis

Khanaki 2012 *(Continued)*

	Prevalence of target condition in the sample: n = 64/138 (46%): stage I-II 46, stage III-IV 18; controls n = 74 *Reference standard*: laparoscopy/laparotomy N = 138 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed on histopathology; staged according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Levels of fatty acids in serum total phospholipids do not seem to be a marker for endometriosis, but the EPA to AA ratio was a relevant factor indicating severity of illness
Conflict of interest	The authors declared no conflict of interests
Notes	For most of the total phospholipid fatty acids (N = 16) there was no statistically significant difference between the groups - no data available for meta-analysis For 18:0 (stearic acid) there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		

Khanaki 2012 *(Continued)*

		High	Low
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Kianpour 2012

Study characteristics	
Patient sampling	*Primary objective*: to investigate CRP levels as a marker of inflammatory process in serum and peritoneal fluid of patients with endometriosis *Participants*: patients subjected to laparoscopy for the evaluation of infertility or pelvic pain *Selection criteria*: exclusion criteria: patients with hypertension, coronary arterial diseases, diabetes, renal diseases, active pelvic inflammatory disease or polycystic ovarian syndrome *Study design*: cross-sectional, single-gate design, prospective collection of samples, non-consecutive enrolment
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility *Age*: mean age 28.9 years, range: 19-44 years (endometriosis group), 30.2 years, range: 24-42 years (controls) *Number of participants enrolled*: 179 women *Number of participants available for analysis*: 179 women (166 in follicular, 13 in luteal cycle phase) *Setting*: Isfahan Fertility and Infertility Center, Isfahan University *Place of study*: Isfahan, Iran

Kianpour 2012 *(Continued)*

	Period of study: 2009-2011 *Language*: English
Index tests	*Index test*: CRP *Details of the index test procedure as stated*: serum concentrations of CRP were measured using enzyme immunoassay kit (Monobind Inc, CA, USA); absorbance at 450 nm was determined by plate reader; sample processing and experiment described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 90/179 (50%): stages not specified; controls n = 89 *Reference standard*: laparoscopy N = 179 (100%) *Description of positive case definition by reference standard as reported*: not reported *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	Measurement of CRP in patients' serum or plasma cannot be used to diagnose endometriosis. It is further recommended that a combination of different markers might be helpful in this regard that could be studied in future
Conflict of interest	The authors declared no conflict of interests
Notes	For CRP there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	**Unclear**

Kianpour 2012 *(Continued)*

DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Kianpour 2013

Study characteristics	
Patient sampling	*Primary objective*: to determine the serum and PF levels of VEGF in endometriosis patients, and to compare with normal subjects *Participants*: patients subjected to laparoscopy for the evaluation of infertility or pelvic pain *Selection criteria*: exclusion criteria: patients with hypertension, coronary arterial diseases, diabetes, renal diseases, active pelvic inflammatory disease or polycystic ovarian syndrome *Study design*: cross-sectional, single-gate design, prospective collection of samples, non-consecutive enrolment

Kianpour 2013 *(Continued)*

Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility *Age*: mean age 28.9 years, range: 19-44 years (endometriosis group), 30.2 years, range: 24-42 years (controls) *Number of participants enrolled*: 179 women *Number of participants available for analysis*: 179 women (166 in follicular, 13 in luteal cycle phase) *Setting*: Isfahan Fertility and Infertility Center, Isfahan University *Place of study*: Isfahan, Iran *Period of study*: 2009-2011 *Language*: English
Index tests	*Index test*: VEGF *Details of the index test procedure as stated*: serum concentrations of VEGF were measured using ELISA kit (Immuno-Biological Laboratory Co, Japan); absorbance at 450 nm was determined by plate reader; concentration was determined using standard curve; sample processing and experiment described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 90/179 (50%): stages not specified; controls n = 89 *Reference standard*: laparoscopy N = 179 (100%) *Description of positive case definition by reference standard as reported*: not reported *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	According to our findings, endometriosis is not associated with change in the level of circulating VEGF
Conflict of interest	Not reported
Notes	For VEGF there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Kianpour 2013 *(Continued)*

Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it prespecified?	No		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Kim 2008

Study characteristics

Kim 2008 *(Continued)*

Patient sampling	*Primary objective*: to investigate the associations between endometriosis and the G(-2518)A polymorphism of monocyte chemotactic protein-1 (MCP-1), and serum and peritoneal fluid MCP-1 levels in Korean women *Participants*: women who underwent laparoscopy for investigation of pelvic pain, ovarian mass, or infertility *Selection criteria*: not reported *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain, ovarian mass, or infertility; no patient had received any medication associated with endometriosis or had any history of pelvic surgery; all women had regular menstrual cycles *Age*: range 20-40 years *Number of participants enrolled*: 206 women *Number of participants available for analysis*: 170 women (all in follicular cycle phase) *Setting*: Department of O&G, College of Medicine, Seoul National University *Place of study*: Seoul, Korea *Period of study*: not reported *Language*: English
Index tests	*Index test*: MCP-1 *Details of the index test procedure as stated*: serum concentrations of MCP-1 measured by using a Quantikine (M) enzyme-linked immunosorbent assay kit (R&D, Minneapolis, USA), according to the manufacturer's instructions; the kit sensitivity was 5 pg/ml; sample processing and laboratory technique described *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: the intra- and interassay CV 4.7% and 5.8%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 94/170 (55%): stage I-II 55, stage III-IV 39; controls n = 76 *Reference standard*: laparoscopy, N = 170 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection followed by histologic examination; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn immediately after surgery *Withdrawals*: 36 participants (17%) were not included in the analysis, the reason for exclusion not explained
Comparative	

Kim 2008 *(Continued)*

Key conclusions by the authors	Serum and peritoneal fluid MCP-1 levels and the G (-2518)A MCP-1 polymorphism were found not to be associated with endometriosis in Korean women
Conflict of interest	Not reported; the work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2005-041-E00224)
Notes	The reported data for MCP-1 polymorphism are not presented in this review For MCP-1 there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		

Kim 2008 *(Continued)*

		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Kitawaki 2005

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the diagnostic significance of CA-125 for endometriosis without ovarian endometriomas *Participants*: patients who underwent laparoscopy or laparotomy and were diagnosed with endometriosis, adenomyosis, leiomyomas, or a normal pelvis *Selection criteria*: inclusion criteria: reproductive age, cyclic menstruation patterns; exclusion criteria: endocrine therapy, including GnRH agonists, danazol, or combination oestrogen–progestin therapy for at least 6 months before enrolment; patients diagnosed with other uterine neoplasms, ovarian neoplasms, pelvic inflammation, or pregnancy *Study design*: cross-sectional, unclear if two- or single-gate design, prospective collection of samples, consecutive series
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: reproductive age, not specified *Number of participants enrolled*: 775 women *Number of participants available for analysis*: 775 women (in follicular or in luteal cycle phase, number of women in each phase is not reported) *Setting*: O&G Department, Kyoto Prefectural University of Medicine *Place of study*: Kyoto, Japan *Period of study*: January 1999 - December 2003 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum concentrations of CA-125 measured by an immunoradiometric assay kit (Centocor, Malvern, USA) and expressed in arbitrary units based on a primary standard; sample processing and laboratory technique described in details *Threshold for positive result*: > 20U/ml, > 26 U/ml, > 30 U/ml; not pre-specified

Kitawaki 2005 *(Continued)*

	Examiners: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: the intra- and interassay CV 5.3% and 3.4%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 433/775 (57%): stage I-II 141, stage III-IV 292; controls n = 342: normal pelvis - 101, other pelvic pathologies - 241 *Reference standard*: laparoscopy/laparotomy N = 775 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection followed by histologic examination; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: "Blood samples were drawn before surgery on days other than those during menstruation" suggests shortly before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	In the diagnosis of endometriosis without endometriomas, combined use of two cut-off values for CA-125, 20 and 30 U/ml, provides improved diagnostic performance. However, the accuracy of using only CA-125 testing for diagnosis is still limited. Serum CA-125 testing can be done during initial screenings of women with possible endometriosis
Conflict of interest	Not reported; supported in part by Grants-in-Aid for Scientific Research (15591772, 15790903 and 16790965) from the Ministry of Education, Culture, Sports, Science and Technology, Japan
Notes	The reported data for CA-125 in diagnosing endometriosis without endometriomas is not presented in this review The diagnostic estimates were calculated for the all the women with versus all the women without endometriosis (regardless of presence of other pelvic pathologies), based on the raw data provided by the authors The diagnostic estimates for the widely used cut-off > 35 U/ml was also provided for the data set, even though this cut-off was not originally assessed by the authors

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Unclear		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			

Kitawaki 2005 *(Continued)*

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		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Kocbek 2013

Study characteristics	
Patient sampling	*Primary objective*: to evaluate serum and peritoneal fluid glycodelin-A concentrations in women with ovarian endometriosis *Participants*: women undergoing surgery for various indications at the authors' institution

Kocbek 2013 *(Continued)*

	Selection criteria: not specified *Study design*: observational, two-gate design, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: endometriosis: not specified, 24/57 were on OCP; controls: indication for surgery was benign ovarian cysts or tubal ligation, 16/42 were using OCP *Age*: mean age 32.9 ± 5.6 years (endometriosis group), 38.4 ± 5.8 years (controls) *Number of participants enrolled*: 99 women *Number of participants available for analysis*: 99 women (57 in follicular and 42 in luteal cycle phase) *Setting*: Faculty of Medicine, University of Ljubljana *Place of study*: Ljubljana, Slovenia *Period of study*: not stated *Language*: English
Index tests	*Index test*: glycodelin-A *Details of the index test procedure as stated*: serum glycodelin level were determined by using ELISA commercial kit (Bio-Serv Dispolab, Switzerland); sample handling described, referenced to the source describing the laboratory technique *Threshold for positive result*: > 2.07 ng/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 57/99 (58%): stage I-II, 12; stage III-IV, 45; controls n = 42 *Reference standard*: surgery (type of surgery not stated) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection and histopathology, staging according to the rAFS classification *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: the samples were collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Our data show significantly increased glycodelin-A concentrations in serum and PF in women suffering from ovarian endometriosis. Our results suggest that glycodelin-A is a potentially useful biomarker for ovarian endometriosis, most likely in combination with other molecules
Conflict of interest	The authors report no declarations of interest; funded by the Slovenian Human Resources Development and Scholarship and a J3-9448 grant from the Slovenian Research Agency
Notes	The reported diagnostic estimated for serum glycodelin-A were adjusted for age and BMI The reported diagnostic estimates for peritoneal glycodelin-A are not presented in this review

Kocbek 2013 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Kocbek 2013 *(Continued)*

	Low

Kocbek 2014a

Study characteristics	
Patient sampling	*Primary objective*: to evaluate PLA2G2A mRNA and protein levels in tissue samples (endometriomas and normal endometrium) and in serum and peritoneal fluid of ovarian endometriosis patients and control women *Participants*: women undergoing surgery for various indications at the authors' institution *Selection criteria*: not specified *Study design*: observational, two-gate design, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: endometriosis group: ovarian endometriosis; controls: women with benign ovarian cysts and women who were undergoing tubal sterilisation; 14 patients with endometriosis and 4 control women took NSAID or other analgesics in the last week before blood collection *Age*: mean age 32.9 ± 6.2 years (endometriosis group), 39.5 ± 3.8 years (controls) *Number of participants enrolled*: 116 women (68 in follicular and 43 in luteal cycle phase; for 5 women information on cycle phase was not available) *Number of participants available for analysis*: 91 women *Setting*: Faculty of Medicine, University of Ljubljana *Place of study*: Ljubljana, Slovenia *Period of study*: 2008-2011 *Language*: English
Index tests	*Index test*: PLA2G2A *Details of the index test procedure as stated*: serum PLA2G2A levels were determined by using commercially available ELISA kits (Cat. #585000; Cayman Chemicals, PA); the limit of detection was 15 pg/ml, and the linear range was 0–1000 pg/ml; sample handling described, referenced to the source describing the laboratory technique *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 70/116 (60%): stages I-II 18, stage III-IV 48; not available 4; controls n = 46 *Reference standard*: laparoscopy, N = 116 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection and histopathology, staging according to the rAFS classification *Examiners*: not stated

Kocbek 2014a *(Continued)*

Flow and timing	*Time interval between index test and reference standard*: not specified; from the context appears that the samples were collected at surgery *Withdrawals*: In 25 women (22%) blood samples were not collected
Comparative	
Key conclusions by the authors	PLA2G2A is implicated in the pathophysiology of ovarian endometriosis, but that it cannot be used as a diagnostic biomarker
Conflict of interest	The authors report no declarations of interest; the study was supported by a Slovenian Human Resource Scholarship and a J3-4135 grant from the Slovenian Research Agency
Notes	The reported data for PLA2G2A in peritoneal fluid and endometrium are not presented in this review For PLA2G2A there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Kocbek 2014a *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Kocbek 2014b

Study characteristics	
Patient sampling	*Primary objective*: to evaluate biglycan expression at the protein level in tissue, serum and peritoneal fluid (PF) from ovarian endometriosis patients, patients with benign ovarian cysts and healthy women *Participants*: women undergoing surgery for various indications at the authors' institution *Selection criteria*: not specified *Study design*: observational, two-gate design, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: endometriosis group: ovarian endometriosis; controls: benign ovarian cyst (n=10) and tubal sterilisation (n=30) *Age*: Reproductive age *Number of participants enrolled*: 96 women *Number of participants available for analysis*: 96 women (in proliferative or secretory cycle phase) *Setting*: Faculty of Medicine, University of Ljubljana *Place of study*: Ljubljana, Slovenia *Period of study*: 2008-2011 *Language*: English
Index tests	*Index test*: biglycan *Details of the index test procedure as stated*: serum Biglycan level was measured by ELISA using rabbit (Sigma-Aldrich HPA003157) and goat (R&D Systems, MN, USA) anti-biglycan polyclonal antibodies and the recombinant biglycan protein (R&D Systems 2667-ICM-050); assay sensitivity was 10 pg/ml with a linear detection range 10 pg/ml - 100 ng/ml; sample handling described

Kocbek 2014b *(Continued)*

	Threshold for positive result: not reported *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 56/96 (58%): stages I-IV; controls n = 40 *Reference standard*: laparoscopy, n = 96 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection and histopathology *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: not specified; from the context appears that the samples were collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Biglycan appears to be involved in ovarian pathologies and probably has different roles in benign cysts as compared to ovarian endometriomas
Conflict of interest	The authors report no declarations of interest; the study was supported by a Slovene Human Resource Scholarship 2011 and a J3-4135 grant from the Slovenian Research Agency
Notes	The reported data for biglycan in peritoneal fluid and endometrium are not presented in this review For biglycan there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		

Kocbek 2014b *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Unclear	

Koninckx 1996

Study characteristics	
Patient sampling	*Primary objective*: to evaluate clinical examination during menstruation and plasma CA-125 concentration to diagnose endometriosis *Participants*: women scheduled for laparoscopy for suspected endometriosis *Selection criteria*: exclusion criteria: hormonal treatment or medical treatment for endometriosis in the 3 months preceding laparoscopy, refusal a clinical examination during menstruation (only DIE considered) *Study design*: cross-sectional single-gate, prospective recruitment and collection of samples, consecutive series
Patient characteristics and setting	*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

Koninckx 1996 *(Continued)*

	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 *Language*: English
Index tests	*Index test*: CA-125 in mid-follicular phase *Details of the index test procedure as stated*: CA-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 *Threshold for positive result*: > 35 U/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: intra- and interassay variation < 5% and < 8%
Target condition and reference standard(s)	*Target condition*: deep infiltrating endometriosis and ovarian endometrioma *Prevalence of target condition in the sample*: n = 38/55 (69%): stage I-II 29, stage III-IV 9; deep endometriosis 13, endometrioma 9, deep endometriosis + severe cul-de-sac adhesions + endometrioma 24; controls n = 17 *Reference standard*: laparoscopy N = 55 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection, deep endometriosis classified as type I and type II, reference to the source with diagnostic criteria and described; staging according to the rAFS classification . *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: the samples were collected up to 4 months before surgery (personal communication with the author) *Withdrawals*: in 6 women (11%) the surgery was cancelled for various reasons
Comparative	
Key conclusions by the authors	Clinical examination during menstruation can reliably diagnose deep infiltrating endometriosis, cystic ovarian endometriosis or cul-de-sac adhesions. This test, preferentially combined with a follicular phase CA-125 assay, should be used to decide whether a preparation for bowel surgery should be given
Conflict of interest	Not reported
Notes	The reported diagnostic estimates for clinical examination or for a combination of clinical examination with blood test are not presented in this review The presented diagnostic estimates are for DIE, ovarian endometrioma and severe cul-de-sac adhesions; the authors also report separate diagnostic estimates for each of these conditions - not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Koninckx 1996 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Kubatova 2013

Study characteristics	
Patient sampling	*Primary objective*: to define markers that can be used in the diagnosis and follow-up of patients with endometriosis by determining serum CA-125, transforming growth factor beta1 (TGF-β1), interleukin 6 (IL-6), and IL-12 levels *Participants*: women who underwent laparoscopy for suspected endometriosis or tubal ligation *Selection criteria*: exclusion criteria: myoma uteri, dermoid cysts, ovarian cystic structures > 3 cm other than endometrioma, pelvic inflammatory disease, any malignancy, oral contraceptives, GnRH analogues, progestin, danazol or any other hormonal therapy *Study design*: observational, two-gate design, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: dysmenorrhoea: 42/61 (endometriosis), 2/12 (controls); chronic pelvic pain: 29/61 (endometriosis), 2/12 (controls); dyspareunia: 22/61 (endometriosis); infertility: 20/61 (endometriosis), 4/12 (controls); none of the patients had taken anti-inflammatory medications or had been diagnosed with an inflammation or infection in previous 6/12 months before the study *Age*: range 18-40 years *Number of participants enrolled*: 73 women *Number of participants available for analysis*: 73 women (all in follicular cycle phase) *Setting*: Department of Obstetrics and Gynaecology, Gazi University School of Medicine, *Place of study*: Ankara, Turkey *Period of study*: not reported *Language*: English
Index tests	*Index test*: CA-125, TGF-β1, IL-6, IL-12 *Details of the index test procedure as stated*: serum CA-125 levels were measured by chemiluminescence using IMMULITE 2000 hormone analyser (Diagnostic Products Corporation, CA, USA); serum TGF-β1, IL-6, and IL-12 levels were measured by using ELISA kits (Biosource International, USA); sample processing described *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if were blinded to the results of reference standard *Interobserver variability*: the intra- and interassay CV were < 10% for all assays
Target condition and reference standard(s)	*Target condition*: endometriosis (peritoneal and ovarian) *Prevalence of target condition in the sample*: n = 61/73 (84%): stage I-II 14, stage III-IV 47; controls n = 12 *Reference standard*: laparoscopy N = 73 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection followed by histologic examination; same protocol was used in diagnostic phase of surgery: inspection of pelvic and peritoneal organs, peritoneal washings and staging according to the rASRM classification *Examiners*: all the procedures were performed by the same team of 2 experienced laparoscopists
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn at surgery *Withdrawals*: none

Kubatova 2013 *(Continued)*

Comparative	
Key conclusions by the authors	TGF-β1 and IL-6 measurements might be a promising alternative in adjunct to CA-125 for the non-invasive diagnosis of endometrioma. However CA-125, TGF-β1, IL-6 and IL-12 seem not to have the diagnostic value in the diagnosis of early stage endometriosis. Of all the serum markers studied, only TGF-β1 was found to be correlated with the stage of endometriosis
Conflict of interest	The authors declared no conflict of interests
Notes	The data for association between the biomarkers levels and type of endometriosis or clinical findings are not presented in this review For IL-12 there was no statistically significant difference between the groups - no data available for meta-analysis For CA-125, TGF-β1, IL-6 there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly	Yes		

Kubatova 2013 *(Continued)*

classify the target condition?			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Kuessel 2014

Study characteristics	
Patient sampling	*Primary objective*: to determine the serum and PF levels of VEGF in endometriosis patients and to compare with normal subjects *Participants*: women undergoing diagnostic or therapeutic laparoscopy because of suspected endometriosis, pelvic pain of unknown origin, benign adnexal masses or leiomyoma uteri *Selection criteria*: inclusion criteria: pre-menopausal age (18-50 years), written informed consent; exclusion criteria: known infectious or chronic autoimmune diseases *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility, pelvic mass and other not specified; hormonal therapy during 3/12 months before surgery - 8/44 in endometriosis and 4/32 in control group *Age*: mean age 33.9 ± 7.8 years (endometriosis group), 36.8 ± 7.4 years (controls) *Number of participants enrolled*: 76 women *Number of participants available for analysis*: 76 women (49 in follicular, 27 in luteal cycle phase) *Setting*: Department of O&G, Medical University of Vienna *Place of study*: Vienna, Austria *Period of study*: not provided *Language*: English

Kuessel 2014 *(Continued)*

Index tests	*Index test*: CK19 *Details of the index test procedure as stated*: serum concentrations of CK19 were measured using a sandwich ELISA TM-Cyfra21-1 (DRG Instruments GmbH, Germany); the lower limit of quantification was 1.5 ng/ml, defined as the lowest step in a dilution series of the standard where CV was still < 30%; sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 44/76 (58%): stages not specified; controls n = 32 *Reference standard*: laparoscopy N = 76 (100%) + histopathology *Description of positive case definition by reference standard as reported*: visual inspection and diagnosis was proven histologically *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected during surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	In this study, the promising data reported in the recent literature about CK19 serving as a sufficient biomarker for endometriosis could not be verified when tested in a larger sample size. Further studies are warranted to explore the usefulness of CK19 in the diagnosis of endometriosis
Conflict of interest	Not reported; supported by Bayer Pharma AG
Notes	For CK19 there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			

Kuessel 2014 *(Continued)*

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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Kurdoglu 2009

Study characteristics	
Patient sampling	*Primary objective*: to clarify the value of serum CA-19.9 in the clinical evaluation of endometriosis *Participants*: women undergoing laparoscopy or laparotomy or various indications at the authors' institution *Selection criteria*: exclusion criteria: 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 *Study design*: cross-sectional two-gate, prospective recruitment and collection of samples

Kurdoglu 2009 *(Continued)*

Patient characteristics and setting	*Clinical presentation*: indications for surgery: 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 *Language*: English
Index tests	*Index test*: CA-19.9, CA-125 in serum *Details of the index test procedure as stated*: not reported *Threshold for positive result*: CA-125 > 35.0 U/ml; CA-19.9 > 37.0 U/ml - pre-specified *Examiners*: not stated *Interobserver variability*: not stated
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 101/127 (80%): stage I-II 26, stage III-IV 75; pelvic endometriosis - 86, ovarian endometrioma - 15; controls n = 26 *Reference standard*: laparoscopy/laparotomy N = 127 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection and histological examination of all excised surgical material; staging according to the rASRM classification *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before surgery *Withdrawals*: 52 patients from control group were excluded (48 for benign ovarian mass and 4 for salpingitis)
Comparative	
Key conclusions by the authors	Both CA-125 and CA-19.9 had high sensitivity with relatively low specificity in the detection of endometriosis. However, the predictive values of CA-125 and CA-19.9 seem high only to predict severe (stages III and IV) disease.
Conflict of interest	Not reported; supported by Gazi University, Unit of Scientific Research Projects, Turkey, grant number 01/2003-42
Notes	The presented data enabled calculation of the diagnostic estimates for different stages of endometriosis - not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Kurdoglu 2009 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Unclear**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Lambrinoudaki 2009

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the hypothesis of increased systemic oxidative stress in patients with endometriosis *Participants*: women of reproductive undergoing laparoscopy for unexplained infertility, pelvic pain, adnexal mass, or tubal ligation *Selection criteria*: exclusion criteria: treatment with antioxidants or anti-inflammatory or hormonal preparations for at least 6 months before laparoscopy; elevated CRP level or WBC or basal body temperature > 37 C° on admission *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility, pelvic pain, adnexal mass *Age*: mean 33.1 ± 6.0 years (endometriosis group) and 34.9 ± 9.2 years (controls) *Number of participants enrolled*: 90 women *Number of participants available for analysis*: 66 women (phase of menstrual cycle not specified) *Setting*: Department of O&G, Aretaieion Hospital, University of Athens *Place of study*: Athens, Greece *Period of study*: January 2006 - November 2006 *Language*: English
Index tests	*Index test*: oxidative stress proteins: HSP70, HSP70', TRX, IMA *Details of the index test procedure as stated*: serum levels of HSP70, HSP70', TRX, IMA were determined in ELISA commercial kits (Hsp 70 ELISA Kit,Stressgen Bioreagents, Canada), (Hsp 70b' ELISA Kit, Stressgen), (TRX ELISA Kit; Redox Bioscience Inc, Japan),(Albumin Cobalt Binding test; Inverness Medical Professional Diagnostics, CO); the sensitivity of HSP70, HSP70', TRX, IMA assays was 0.5 ng/ml, 0.06 ng/ml, 0.25 ng/ml, 28.00 U/ml; sample handling described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: For HSP70 and HSP70' intra- and interassay CV < 10%; for TRX intra- and interassay, CV was 8.3% and 12.2%; for IMA intra- and interassay, CV was 1.7% and 3.5%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 45/66 (68%), stage I-II 13, stage III-IV 32; controls n = 21 *Reference standard*: laparoscopy N = 66 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection (a thorough search for endometriotic foci in each patient); staging according to the rAFS classification *Examiners*: all laparoscopic procedures were performed by the same surgeon, who was blinded to the indication of laparoscopy
Flow and timing	*Time interval between index test and reference standard*: blood sampling was performed 48 h before surgery *Withdrawals*: 24 of recruited participants were not eligible and were excluded from the study

Lambrinoudaki 2009 *(Continued)*

Comparative	
Key conclusions by the authors	Women with endometriosis have evidence of increased systemic oxidative stress expressed by higher levels of HSP70b'. The stage of the disease is not associated with circulating HSP70b'
Conflict of interest	Not reported
Notes	For HSP70, IMA, TRX there was no statistically significant difference between the groups - no data available for meta-analysis For HSP70b' there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		

Lambrinoudaki 2009 *(Continued)*

		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Lamp 2012

Study characteristics	
Patient sampling	*Primary objective*: to evaluate associations between survivin promoter polymorphisms and the risk of endometriosis, as well as to compare the immunoreactivity to survivin in sera of patients with and without endometriosis *Participants*: women undergoing laparoscopy for infertility, pelvic pain and suspected endometriosis *Selection criteria*: not specified *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis group - not specified; controls: infertility (n = 35) or pelvic pain (n = 12) *Age*: mean 30.9 ± 6.5 years (endometriosis group) and 30.0 ± 6.1 years (controls) *Number of participants enrolled*: 196 women *Number of participants available for analysis*: 145 women (phase of menstrual cycle not specified) *Setting*: Department of O&G, University of Tartu *Place of study*: Tartu, Estonia *Period of study*: not reported *Language*: English
Index tests	*Index test*: anti-survivin antibodies *Details of the index test procedure as stated*: serum anti-survivin antibodies were detected with a specific ELISA kit (Uscn Life Science Inc, Wuhan, China) according to the manufacturer's protocol *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not reported

Lamp 2012 *(Continued)*

Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 98/145 (68%): stage I-II 55, stage III-IV 43; controls n = 47 *Reference standard*: laparoscopy N = 145 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: Surgically and histologically confirmed endometriosis; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained before surgery *Withdrawals*: 51 of recruited participants with endometriosis were not included in anti-survivin antibody testing, reason not explained
Comparative	
Key conclusions by the authors	Survivin promoter polymorphisms are not associated with susceptibility to endometriosis in the Estonian population, and though the expression of survivin is increased in endometriotic lesions, autoimmune reactivity against it is similar in women with and without the disease
Conflict of interest	Not reported; the work was funded by the European Union Regional Development Fund and by Enterprise Estonia, Grant no. EU30200, by the Estonian Science Foundation (grants 6573 and 6585) and by the Estonian Ministry of Education and Research (core grants SF0180044s09 and SF0180035s08)
Notes	For anti-survivin antibodies there was no statistically significant difference between the groups - no data available for meta-analysis The data for survivin promoter region polymorphisms are not included in this review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowl-	Unclear		

Lamp 2012 *(Continued)*

edge of the results of the reference standard?			
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Lanzone 1991

Study characteristics	
Patient sampling	*Primary objective*: to evaluate CA-125 in serum and peritoneal fluid of women with various stages of endometriosis and in the control subjects *Participants*: women undergoing laparoscopy for infertility or pelvic pain during luteal phase of the cycle *Selection criteria*: exclusion criteria: peritoneal fluid positive for mycoplasma and chlamydia *Study design*: longitudinal single-gate, prospective recruitment and collection of samples, consecutive series
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility or both

Lanzone 1991 *(Continued)*

	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 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 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum CA-125 levels measured with radioimmunoassay (CIS Diagnostici); all samples from the same patient were assayed at the same time *Threshold for positive result*: CA-125 > 35.0 U/ml - pre-specified *Examiners*: not stated *Interobserver variability*: the inter- and intra-assay CV were 8% and 15%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 81/270 (30%): stage I-II 31, stage III-IV 50; controls n = 38 *Reference standard*: laparoscopy N = 270 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rAFS classification *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before surgery *Withdrawals*: 151 participants were excluded (reason not explained)
Comparative	
Key conclusions by the authors	The measurement of serum CA-125 does not appear to be useful for the diagnosis and management of endometriosis.Therefore, at present, laparoscopy should be considered the most specific and sensitive method of detecting and following the disease
Conflict of interest	Not reported
Notes	The reported estimates for peritoneal fluid and the estimates following medical treatment for endometriosis are not presented in this review The reported diagnostic estimates per stages of severity of endometriosis are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Lanzone 1991 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Li 2005

Study characteristics

Li 2005 *(Continued)*

Patient sampling	*Primary objective*: to investigate the function of T-lymphocyte subsets in patients with endometriosis *Participants*: women with endometriosis confirmed by laparoscopy and a group of women who underwent tubal ligation or anastomosis with a normal pelvis at laparoscopy *Selection criteria*: exclusion criteria: autoimmune diseases, allergic diseases and acute inflammation, no steroid treatment 3 months prior to surgery *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: mean age 35 ± 7 years (endometriosis group), 38 ± 4 years (controls) *Number of participants enrolled*: 50 participants (10 women with fibroid uterus in whom endometrial samples were assessed comprised separate control group and were not included in this review) *Number of participants available for analysis*: 50 participants (cycle phase not reported) *Setting*: Department of O&G, Qingdao Eighth People's Hospital *Place of study*: Qingdao, China *Period of study*: September 2001 - September 2002 *Language*: Chinese
Index tests	*Index test*: IL-2 and IL-6 *Details of the index test procedure as stated*: serum IL-2 and IL-6 were measured with ELISA kits (LIFEFEY BioMeditech Corporation USA), working assay range or minimal detection limit are not included in the paper; sample handling described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: no information provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 30/50 (60%): stage I-II 9, stage III-IV 21; controls n = 20 *Reference standard*: laparoscopy N = 50 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected on the day of surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The levels of IL-6 in the serum and peritoneal fluid of patients with endometriosis are increased, implying that IL-6 might play a role in the pathophysiology of endometriosis. The ratio of IL-2/IL-6 in the serum and peritoneal fluid was decreased in patients with endometriosis compared with the control group, suggesting shift of Th1 cell toward Th2 cell in patients with endometriosis. Stronger expression of IL-2 and IL-6 in the ectopic endometrial tissues may contribute to the disturbed immune regulation in patients with endometriosis.

Li 2005 *(Continued)*

Conflict of interest	Not reported
Notes	The data for markers measured in peritoneal fluid and endometrium are not reported in this review For IL-2 there was no statistically significant difference between the groups - no data available for meta-analysis The levels of IL-6 were statistically significantly higher in endometriosis, but there were no data to construct 2 x 2 tables - not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			

Li 2005 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Lima 2006

Study characteristics	
Patient sampling	*Primary objective*: to determine FSH, LH, E2, progesterone, and Hi concentrations in serum, PF and FF of women with and without endometriosis *Participants*: women undergoing laparoscopy for infertility and/or pelvic pain (cases) and tubal sterilisation (controls) *Selection criteria*: inclusion criteria: secretory cycle phase, no medical treatment for at least three months preceding surgery, absence of other gynaecological diseases, absence of pelvic pain, age between 18 and 40 years *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility (cases); asymptomatic fertile women requesting sterilisation (controls) *Age*: mean age 33.9 ± 7.8 years (endometriosis group), 36.8 ± 7.4 years (controls) *Number of participants enrolled*: 49 women *Number of participants available for analysis*: 49 women (all in luteal cycle phase) *Setting*: Department of O&G, Hospital das Clinicas, Faculty of Medicine of Ribeirão Preto, University of São Paulo *Place of study*: São Paulo, Brazil *Period of study*: 2002-2004 *Language*: Portuguese
Index tests	*Index test*: FSH , LH, E2, progesterone *Details of the index test procedure as stated*: serum concentrations of FSH, LH, E2 and progesterone were measured using a commercial kit (DPC Imm Sys, California) by chemiluminescence; sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: Intra- and interassay CV were for FSH 7.9% and 6.5%, for LH 8.8% and 11.3%, for E2 8.4% and 9.3%, for P 5.8% and 10.3%

Lima 2006 *(Continued)*

Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 28/49 (57%): stage I-II 18, stage III-IV 10; controls n = 21 *Reference standard*: laparoscopy n = 49 (100%) *Description of positive case definition by reference standard as reported*: visual inspection; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	Ovary dysfunction in women with endometriosis, with reduction on E, P and Hi concentrations, which may contribute to the subfertility often associated with the disease
Conflict of interest	Not reported
Notes	For LH and FSH there was no statistically significant difference between the groups - no data available for meta-analysis For E2 and progesterone there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		

Lima 2006 *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	Yes

	High	**Low**
DOMAIN 3: Reference Standard		
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?	Yes	
	Unclear	**Low**
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
	Low	

Lin 2005

Study characteristics	
Patient sampling	*Primary objective*: to investigate the role of interleukin-16 (IL-16) in the pathogenesis of endometriosis *Participants*: women with suspected endometriosis who underwent laparoscopy *Selection criteria*: exclusion criteria: autoimmune diseases, no steroid treatment or immunosuppressant treatment 6 months prior to surgery *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: mean age 37 ± 10.3 years (endometriosis group), 36.8 ± 12.1 years (controls) *Number of participants enrolled*: 44 participants *Number of participants available for analysis*: 44 participants (cycle phase not reported) *Setting*: Department of O&G, College of Medicine, Zhejiang University *Place of study*: Hangzhou, China

Lin 2005 *(Continued)*

	Period of study: September 2001 - June 2002 *Language*: Chinese
Index tests	*Index test*: IL-16 *Details of the index test procedure as stated*: serum IL-16 was measured with enzyme-linked immunosorbent assay (ELISA) (human IL-16 BMS 248, Bender Medsystems, Vienna, Austria); no working ranges were reported; sample handling described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: CV < 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 22/44 (50%): stage I-II 8, stage III-IV 14; controls n = 22 *Reference standard*: laparoscopy/laparotomy N = 44 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected on the day of surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Reduced levels of IL-16 in peritoneal fluid and serum of women with advanced stage endometriosis may imply a role of IL-16 in the development and progression of endometriosis.
Conflict of interest	Not reported
Notes	The data for markers measured in peritoneal fluid are not reported in this review For IL-16 there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**

Lin 2005 *(Continued)*

DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Liu 2009

Study characteristics	
Patient sampling	*Primary objective*: to establish the diagnostic model for endometriosis by screening the plasma biomarkers of endometriosis using surface enhanced laser desorption/ionisation time of flight mass spectrometry (SELDI-TOF-MS) coupled with bioinformatic *Participants*: women undergoing laparoscopy for infertility, pelvic pain or benign ovarian mass *Selection criteria*: not reported *Study design*: cross-sectional, single-gate, prospective sample collection

Liu 2009 *(Continued)*

Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility or adnexal mass *Age*: mean age 33.6 ±4.7 years (training set), 34.2 ± 3.6 years (test set) (endometriosis group); 32.5 ± 3.2 years (training set), 33.0 ± 2.8 years (test set) (controls) *Number of participants enrolled*: 102 participants (71 women - training set; 31 women - test set) *Number of participants available for analysis*: 102 participants (cycle phase not reported) *Setting*: Department of O&G, Peking Union Medical Colledge Hospital *Place of study*: Beijing, China *Period of study*: January 2007 - October 2007 *Language*: Chinese
Index tests	*Index test*: proteome by SELDI-TOF-MS (3 protein peaks with the molecular weight of 3,956.00 Da, 11,710.00 Da and 6,986.00 Da) *Details of the index test procedure as stated*: surface-enhanced laser desorption/ionisation coupled to time-of-flight mass spectrometry (detection analysis of protein chips was done with ProteinChip Biotechnology System mass spectrometer (PBS-II, Ciphergen Co, America); bioinformatic analysis by using ProteinChip Software 3.1.1 and Biomarker Pattern Software; CART model for training set (70%) with double blind validation on test set (30%); sample handling and procedure described in details *Threshold for positive result*: presence of specific protein peaks intensities, not pre-specified *Examiners*: not stated, blinded to the clinical outcomes *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 52/102 (51%): stage I-II 23, stage III-IV 29; controls n = 50 *Reference standard*: laparoscopy N = 102 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to rAFS classification *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	SELDI-TOF-MS is a new approach for screening markers of endometriosis. Its clinical value deserves further investigation
Conflict of interest	Not reported
Notes	The diagnostic estimates for the validation test set are reported in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Liu 2009 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Mabrouk 2012

Study characteristics	
Patient sampling	*Primary objective*: to investigate the univariable and multivariable performances of the mRNA levels of MMP-3, MMP-9, VEGF and survivin in peripheral blood and the serum levels of CA-125, CA-19.9 to diagnose or exclude the endometriosis and to differentiate between deep infiltrating and ovarian endometriosis *Participants*: women of reproductive age undergoing laparoscopy for suspected endometriosis or non-malignant conditions (myoma, tubal ligation, and ovarian biopsy) *Selection criteria*: exclusion criteria: suspected or ascertained diagnosis of systemic pathologies (malignancies, autoimmune diseases, liver diseases) or pregnancy *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: range 26-40 years *Number of participants enrolled*: 60 women *Number of participants available for analysis*: 60 women (all in the follicular phase of the menstrual cycle) *Setting*: the Minimally Invasive Gynecological Surgery Unit, S. Orsola-Malpighi Hospital, University of Bologna *Place of study*: Bologna, Italy *Period of study*: February 2007 - May 2008 *Language*: English
Index tests	*Index test*: MMP-3 mRNA, MMP-9 mRNA, VEGF mRNA, survivin mRNA, CA-125, CA19-9 *Details of the index test procedure as stated*: detection of serum CA-125 and CA-19.9 was performed using a commercially available chemiluminescent immunometric assay (Roche Diagnostics GmbH, Germany) by using the Elecsys Analyzer; sensitivity for both assays was 0.6 IU/ml. All other biomarkers in peripheral blood were detected by qRT-PCR with gene-specific primers on the ABI PRISM 7900 Sequence Detection System (PE Applied Biosystems); laboratory techniques and sample processing described in details *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 40/60 (67%) (DIE and ovarian endometrioma); controls n = 20 *Reference standard*: laparoscopy N = 60 (100%) *Description of positive case definition by reference standard test as reported*: diagnosis of endometriosis was surgical and histological *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained from the patients immediately before laparoscopy

Mabrouk 2012 *(Continued)*

	Withdrawals: none
Comparative	
Key conclusions by the authors	A combination of serum and molecular markers could allow a better diagnosis of endometriosis
Conflict of interest	The authors declare that they have no conflict of interest
Notes	For VEGF and MMP9 there was no difference between the groups - no data available for meta-analysis For MMP3 there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		

Mabrouk 2012 *(Continued)*

		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Maeda 2002a

Study characteristics	
Patient sampling	*Primary objective*: to investigate host immunologic response to endometriosis in terms of intercellular adhesion molecule (ICAM)-1 expression by macrophages and killer cell inhibitory receptor (KIR) expression by natural killer (NK) cells *Participants*: women undergoing laparoscopy for various indications *Selection criteria*: exclusion criteria: history of pregnancy or history of treatment with GnRH analogues within 3 years, complications from apparent pelvic inflammatory disease *Study design*: cross-sectional, two-gate, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: endometriosis group - not specified; controls: benign ovarian cysts - 12, uterine myoma - 7, infertility - 4, paraovarian cysts - 2, carcinoma in situ of uterine cervix - 1 *Age*: mean age 32.8 ± 7.5 years (endometriosis group), 35.0 ± 8.9 years (controls) *Number of participants enrolled*: 54 participants *Number of participants available for analysis*: 54 participants (all in early follicular cycle phase) *Setting*: Department of O&G, Kochi Medical School *Place of study*: Kochi, Japan *Period of study*: April 1999 - August 2000 *Language*: English
Index tests	*Index test*: PBMC (CD3, CD4, CD8, CD19, CD16, CD14), ICAM-1, KIR2DL1^{+}NK, KIR2DL2^{+}NK *Details of the index test procedure as stated*:**p**eripheral blood mononuclear cells were measured by flow cytometry using specific mononuclear antibodies (FITC-labelled anti-CD3, anti-CD4 mAb and PE-labelled anti-CD8 mAb as T cell markers, PE-labelled anti-CD19 mAb as B cell marker, FITC-labeled anti-CD16 mAb as NK cell and FITC-labelled anti-CD14 mAb as monocyte/macrophage marker, PE-labeled anti-CD54 (ICAM-1) mAb as marker for monocyte/macrophage activation, and PE-labelled anti-CD158a, anti-CD158b, and CD94 as markers for KIRs (all from Beckman-Coulter, Fullerton, CA); laboratory technique described *Threshold for positive result*: not reported

Maeda 2002a *(Continued)*

	Examiners: not information provided, unclear if were blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 28/54 (52%): stage I-II 11, stage III-IV 17; controls n = 26 *Reference standard*: laparoscopy, N = 54 (100%) *Description of positive case definition by reference standard test as reported*: staging according to rAFS classification *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Properties of macrophages and NK cells in women with endometriosis promote immunotolerance to implanted tissue in the peritoneal environment. Increased KIR(+)NK cells in peripheral blood may represent a risk factor for endometriosis
Conflict of interest	Not reported
Notes	For PBMC (CD3, CD4, CD8, CD19, CD16, CD14) there was no statistically significant difference between the groups - no data available for meta-analysis For KIR2DL2^{+}NK the data reported in larger overlapping study (Maeda 2002b) For ICAM-1 and KIR2DL1^{+}NK there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results	Unclear		

Maeda 2002a *(Continued)*

of the reference standard?			
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Maeda 2002b

Study characteristics	
Patient sampling	*Primary objective*: to investigate the host immunologic response to endometriosis in terms of killer inhibitory receptor (KIR) expression by natural killer (NK) cells *Participants*: women undergoing laparoscopy for various indications *Selection criteria*: exclusion criteria: history of pregnancy or history of treatment with GnRH analogues within 3 years, complications from apparent pelvic inflammatory disease *Study design*: cross-sectional, two-gate, p

Maeda 2002b *(Continued)*

	rospective sample collection
Patient characteristics and setting	*Clinical presentation*: endometriosis group - not specified; controls: benign ovarian cysts - 15, uterine myoma - 14, infertility - 7, paraovarian cysts - 2, carcinoma in situ of uterine cervix - 2 *Age*: mean age 32.0 ± 7.2 years (endometriosis group), 35.0 ± 9.2 years (controls) *Number of participants enrolled*: 82 participants *Number of participants available for analysis*: 82 participants (cycle phase not reported) *Setting*: Department of O&G, Kochi Medical School *Place of study*: Kochi, Japan *Period of study*: April 1999 - January 2001 *Language*: English
Index tests	*Index test*: KIR2DL1$^+$NK, KIR2DL2$^+$NK, CD94$^+$NK cells *Details of the index test procedure as stated*: NK cells were measured by flow cytometry using specific mononuclear antibodies (FITC-labeled anti-CD16 mAb as NK cell, PE-labelled anti-CD158a and anti-CD158b as markers for KIR subfamilies KIR2DL1 and KIR2DL2 expressed on NK cells and CD94 as lectin-like receptor marker on NK cells (all from Beckman-Coulter, Fullerton, CA); laboratory technique described *Threshold for positive result*: not reported *Examiners*: not information provided, unclear if were blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 42/82 (51%): stage I-II 12, stage III-IV 30; controls n = 40 *Reference standard*: laparoscopy, N = 82 (100%) *Description of positive case definition by reference standard test as reported*: staging according to rAFS classification *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	The proportion of KIR2DL1(+)NK cells was increased in peritoneal fluid and peripheral blood in women with endometriosis; this difference is probably related to NK cell suppression in endometriosis. This increase in KIR2DL1 expression by NK cells may represent a risk factor in the pathogenesis of endometriosis
Conflict of interest	Not reported
Notes	For KIR2DL2$^+$NK and CD94$^+$NK there was no statistically significant difference between the groups - no data available for meta-analysis For KIR2DL1$^+$NK there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review

Maeda 2002b *(Continued)*

	The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Unclear		

Maeda 2002b *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Maiorana 2007

Study characteristics	
Patient sampling	*Primary objective*: to evaluate if serum CA-125 levels correlate with rAFS and whether serum CA-125 measurement should be performed in the routine work-up of dysmenorrhoea and dyspareunia *Participants*: 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) *Selection criteria*: exclusion criteria: patients with malignant tumours or inflammatory disease *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*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 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: 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 *Threshold for positive result*: > 35 U/ml, pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 69/86 (79%): stage I-II 14, stage III-IV 55; controls n = 17 *Reference standard*: laparoscopy N = 86 (100%) *Description of positive case definition by reference standard test as reported*: surgical diagnosis, rASRM classification

Maiorana 2007 *(Continued)*

	Examiners: no information provided
Flow and timing	*Time interval between index test and reference standard*: not specified, but statement 'preoperative blood sample' implies short time before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	CA-125 levels are related to endometriosis and rAFS score in the evaluated patient series; no correlation was found between CA-125 and pelvic pain with endometriosis
Conflict of interest	Not reported
Notes	The presented diagnostic estimates according to severity of endometriosis are not included in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		

Maiorana 2007 *(Continued)*

		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Markham 1997a

Study characteristics	
Patient sampling	*Primary objective*: to analyse PF and peripheral blood for concentration of both RANTES and TNF-α in a group of women with and without endometriosis *Participants*: patients undergoing routine gynaecological treatment in hospital for non-malignant conditions *Selection criteria*: not specified *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: reproductive age (personal communication with the authors) *Number of participants enrolled*: 32 women *Number of participants available for analysis*: 32 women (cycle phase not specified) *Setting*: Department of O&G, Queen Elizabeth II Research Institute for Mothers and Infants, University of Sydney *Place of study*: Sydney, Australia *Period of study*: not specified *Language*: English
Index tests	*Index test*: RANTES, TNF-α *Details of the index test procedure as stated*: serum concentrations of RANTES were measured using a commercial sandwich ELISA (R&D Systems, USA) with assay sensitivity 2.5 pg/ml; TNF-α levels were measured by "in house amplified ELISA sandwich" assay with sensitivity of 1.0 pg/ml; sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: Inter- and intra-assay CV for RANTES < 6%, for TNF-α < 9%

Markham 1997a *(Continued)*

Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 23/32 (72%): stage I-II 11, stage III-IV 12; controls n = 9 *Reference standard*: laparoscopy N = 32 (100%) *Description of positive case definition by reference standard as reported*: staging according to the rAFS score *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	We found that RANTES concentrations in blood or peritoneal fluid are unlikely to be helpful as a potential marker for endometriosis
Conflict of interest	Not reported
Notes	For RANTES there was no statistically significant difference between the groups - no data available for meta-analysis For TNF-α there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		

Markham 1997a *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Martinez 2007

Study characteristics	
Patient sampling	*Primary objective*: to evaluate whether serum IL-6 levels could serve as a marker of the early stages of endometriosis and to determine the value of CA-125 as a diagnostic marker *Participants*: women undergoing laparoscopy for various indications at the authors' institution *Selection criteria*: inclusion criteria: reproductive age and regular menstrual cycles; exclusion criteria: administration of any medication over the previous 2 years, acute inflammatory diseases or neoplasms, 2 or more concomitant findings at laparoscopy *Study design*: cross-sectional two-gate, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation*: indications for laparoscopy were pelvic pain (n = 5), 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

Martinez 2007 *(Continued)*

	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 *Language*: English
Index tests	*Index test*: CA-125 and IL-6 *Details of the index test procedure as stated*: 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 IL-6 measured by immunoassay (Quantikine, R&D Systems Inc, MN, USA), minimum detectable value of 0.7 pg/ml. Serum CA-125 level performed using a commercially available chemiluminescent microparticle immunoassay (ARCHITECT CA-125 II Abbott Diagnositics, Spain) with assay sensitivity of < 1.0 IU/ml *Threshold for positive result*: IL-6 > 25.75 pg/ml U/ml, CA-125 > 35 IU/ml - not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: Inter- and intra-assay CV for IL-6 6.4 and 4.2%, for CA-125 ≤ 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 47/119 (40%): stage I-II 11, stage III-IV 36; controls n = 72 *Reference standard*: laparoscopy N = 119 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples collected up to 3 months before surgery *Withdrawals*: 9 women were excluded before the analysis as did not meet inclusion criteria (4 refused surgery, 2 had adhesions related to PID, 3 had fibroid uterus + endometriosis)
Comparative	
Key conclusions by the authors	Serum IL-6 is a reliable, non-invasive marker of minimal and mild endometriosis. Combined with clinical data, this will allow doctors to detect which women are at risk of having early stages of the disease
Conflict of interest	Not reported
Notes	The diagnostic estimates for IL-6 were reported only for minimal-mild endometriosis and for CA-125 reported only for moderate-severe endometriosis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Martinez 2007 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Matalliotakis 2003a

Study characteristics

Matalliotakis 2003a *(Continued)*

Patient sampling	*Primary objective*: to investigate the soluble levels of the angiogenic factors VEGF, EGF-R, GM-CSF, IGF-1, IFN-γ in women with and without endometriosis and to investigate whether administration of danazol and leuprorelin depot to patients with endometriosis regulates their expression *Participants*: women selected from a cohort of 387 women undergoing laparoscopy at the authors' institution *Selection criteria*: inclusion criteria: pre-menopausal, not-pregnant *Study design*: longitudinal, single-gate, prospective collection of samples, selected group from larger cohort
Patient characteristics and setting	*Clinical presentation*: indications for surgery - infertility and suspected endometriosis; infertility work-up (ovulation, cervical mucus, tubal patency and semen analysis) were normal in all women *Age*: mean 28.2 ± 5.6 years (endometriosis group) and 29.3 ± 5.8 years (controls) *Number of participants enrolled*: 48 women *Number of participants available for analysis*: 48 women (phase of menstrual cycle not specified) *Setting*: Department of O&G, the University Hospital of Crete *Place of study*: Crete, Greece *Period of study*: 1991-1999 *Language*: English
Index tests	*Index test*: angiogenic factors VEGF, EGF-R, GM-CSF, IGF-1, IFN-γ *Details of the index test procedure as stated*: serum levels of GM-CSF and IFN-γ were measured with commercial kits (Endogen, MA); by using using ELISA method as specified by the suppliers at test and reference wavelengths of 450 and 550 nm, respectively. Serum levels of IGF-1, VEGF, EGF-R were measured with the affinity-purified goat polyclonal IGF-1 (G-17, sc-1422, Santa Cruz Biotechnology, CA) and the mouse monoclonals for VEGF (Ab-3; JH121, NeoMarkers, CA) and EGFR (Ab-4; clone F4, NeoMarkers, CA); sample handling and laboratory technique described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 28/48 (58%): stage I-II 17, stage III-IV 11; controls - 20 *Reference standard*: laparoscopy N = 48 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were taken before laparoscopy *Withdrawals*: 24 of recruited participants were not eligible and were excluded from the study
Comparative	

Matalliotakis 2003a *(Continued)*

Key conclusions by the authors	EGF-R, GM-CSF, IFN-γ and IGF-1 are being released at high rates in both healthy and endometriotic subjects indicating that they do not actively participate in the disease but not excluding, however, other regulatory roles. VEGF may be associated with the disease process.
Conflict of interest	Not reported
Notes	For EGF-R, GM-CSF, IGF-1 and IFN-γ there was no statistically significant difference between the groups - no data available for meta-analysis For VEGF there was statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review The reported data for the biomarkers following medical treatment are not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Matalliotakis 2003a *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Matalliotakis 2004

Study characteristics	
Patient sampling	*Primary objective*: to investigate the effects of danazol and leuprorelin acetate on CA-125 levels during treatment for endometriosis *Participants*: women who underwent laparoscopy for pelvic pain, infertility or both *Selection criteria*: not specified *Study design*: longitudinal, single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility or both *Age*: mean 28.6 ± 5.2 years (endometriosis group) and 29.4 ± 5.3 years (controls) *Number of participants enrolled*: 100 women *Number of participants available for analysis*: 100 women (phase of menstrual cycle not specified) *Setting*: Department of O&G, the University Hospital of Crete *Place of study*: Crete, Greece *Period of study*: 1991-1999 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum levels of CA-125 were measured by radioimmunoassay with commercial kits (CIS Biointernational, France); kit sensitivity was 1.0 U/ml; sample handling and laboratory technique described

Matalliotakis 2004 *(Continued)*

	Threshold for positive result: > 33 U/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: the intra- and interassay CV were 4.9% and 5.9%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 50/100 (50%): stage I-II 29, stage III-IV 21; controls - 50 *Reference standard*: laparoscopy n = 100 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rAFS system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were taken before laparoscopy *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Danazol and leuprorelin acetate are equally effective in the treatment of endometriosis. Moreover, the results support the view that the determination of CA-125 levels may assist in evaluating progress of endometriosis treatment
Conflict of interest	Not reported
Notes	The reported data for the biomarkers following medical treatment are not included in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it prespecified?	No		

Matalliotakis 2004 *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Matveeva 1990

Study characteristics	
Patient sampling	*Primary objective*: to investigate inhibitory and activation motif expression of killer immunoglobulin-like receptor (KIR) by natural killer (NK) cells, which may be pathogenetically involved in endometriosis *Participants*: women undergoing laparoscopy for various indications *Selection criteria*: exclusion criteria: history of pregnancy or history of treatment with GnRH analogues within previous year, complications from apparent pelvic inflammatory disease *Study design*: cross-sectional, two-gate, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: infertility; all women had regular ovulatory menstrual cycles *Age*: mean age 30.6 years, range 26-35 years *Number of participants enrolled*: 119 participants *Number of participants available for analysis*: 119/74 participants (in follicular or luteal cycle phase), different number of samples for different tests *Setting*: National research centre of mother and child health, Ministry of Health

Matveeva 1990 *(Continued)*

	Place of study: Moscow, Russia *Period of study*: not reported *Language*: Russian
Index tests	*Index test*: PBMC (CD3, CD4, CD8, CD2), IgA, IgM, IgG *Details of the index test procedure as stated*: PBMC were measured by flow cytometry using FACScan (Becton Dickinson, USA); serum immunoglobulins were determined by using Manchini method; laboratory technique described *Threshold for positive result*: not reported *Examiners*: no information provided, unclear if were blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 62/119 (52%): all stage I-II; controls n = 57 *Reference standard*: laparoscopy, n= 119 (100%) *Description of positive case definition by reference standard test as reported*: staging according to rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: data were not reported for up to 45 participants for some of the index tests, reason not explained
Comparative	
Key conclusions by the authors	The tested cells did not markedly differ from those in control fertile patients. Serum concentrations of immunoglobulin M were increased in women with endometriosis. Immunoglobulin concentrations widely varied in the peritoneal fluid, with a statistically significant elevation of IgA and IgM in women with tubal-peritoneal infertility
Conflict of interest	Not reported
Notes	For PBMC (CD3, CD4, CD8, CD2), IgA and IgG there was no statistically significant difference between the groups - no data available for meta-analysis For IgM there was statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid are not presented in this review The data for a group of healthy women (n = 10) who did not have laparoscopy are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Matveeva 1990 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		Low	

Mier-Cabrera 2011

Study characteristics	
Patient sampling	*Primary objective*: to assess immunological variables, T-cell apoptosis and oxidative stress markers in the peripheral blood and peritoneal fluid of women with and without endometriosis *Participants*: women undergoing laparoscopy for infertility or for tubal ligation *Selection criteria*: exclusion criteria: PID, autoimmune disease, endocrine metabolic disease; use of antioxidant medication in the last year, mononuclear peritoneal cell viability < 80% and a final reconstituted peritoneal cell number < 2 x 106 cells/ml *Study design*: cross-sectional, two-gate, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: endometriosis group - infertility, had never received any hormonal treatment; controls: healthy women requesting tubal ligation, had not taken contraceptive hormones in the last 3/12 months *Age*: mean age 32.7 ± 2.5 years (endometriosis group), 33.8 ± 5.4 years (controls) *Number of participants enrolled*: 62 participants *Number of participants available for analysis*: 62 participants (all in peri-ovulatory cycle phase) *Setting*: National Institute of Perinatology *Place of study*: Mexico City, Mexico *Period of study*: not reported *Language*: English
Index tests	*Index test*: intracellular cytokines (CD4+/IFN-γ, CD4+/IL-2, CD8+/IFN-γ, CD8+/IL-2), apoptotic cells, and oxidant markers (malondialdehyde and ascorbic acid) *Details of the index test procedure as stated*: lymphocyte subsets were measured by flow cytometry; degree of apoptosis in T lymphocytes was analysed using a FACS Calibur instrument (BD Biosciences, San Jose, CA, USA) equipped with CellQuest 3.3 software; concentrations of thiobarbituric acid reactive substances were determined according to the method developed by Ohkawa et al; cytokines were measured by using Bio-Plex human cytokine assay (Bio-Plex, Hercules, USA); sample handling and laboratory technique described in details *Threshold for positive result*: not reported *Examiners*: not information provided, unclear if were blinded to the result of reference standard *Interobserver variability*: Intra- and interassays CV for malondialdehyde were 3.5% and 7.5%, for ascorbic acid were 5.0% and 8.0%, for cytokines were 2.0%–7.0% and 3.5%–12.0%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 32/62 (52%): all stage I-II; controls n = 30 *Reference standard*: laparoscopy, N= 62 (100%) *Description of positive case definition by reference standard test as reported*: staging according to rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported

Mier-Cabrera 2011 *(Continued)*

Comparative	
Key conclusions by the authors	The alterations observed in women with endometriosis were associated with a diminished peritoneal T helper type 1 immune response. Pro-inflammatory, chemotactic, angiogenic and oxidative stress markers were altered in the peritoneal milieu of women with endometriosis
Conflict of interest	Not reported; the work was supported by Consejo Nacional de Ciencia y Tecnología: Grant SALUD-2002-C-01-7615/A-1
Notes	For intracellular cytokines (CD4+/IL-2, CD8+/IFN-γ, CD8+/IL-2), apoptotic cells, and oxidant markers (malondialdehyde and ascorbic acid) there was no statistically significant difference between the groups - no data available for meta-analysis For CD4+/IFN-γ there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review For lymphocyte subsets (CD3, CD19, CD4, CD8, CD16+56) there was no statistically significant difference between the groups, but there was insufficient data to confirm negative findings - not included in this review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**

Mier-Cabrera 2011 *(Continued)*

DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Mihalyi 2010

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the combined performance of 6 potential plasma biomarkers in the diagnosis of endometriosis *Participants*: women who underwent laparoscopy for subfertility with or without pain at the authors' institution - identified through electronic database of the bio bank samples *Selection criteria*: exclusion criteria: samples collected from women who were on hormonal medication or had other pelvic inflammatory disease or general diseases at the time of collection, surgery within 6 months prior to the time of collection *Study design*: cross-sectional single-gate, prospective collection of samples, retrospective selection of cases
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility or both *Age*: reproductive age *Number of participants enrolled*: 294 women *Number of participants available for analysis*: 294 women (59 in menstrual, 119 in follicular, 116 in luteal cycle phase) *Setting*: Department of O&G, University Hospital Gasthuisberg

Mihalyi 2010 *(Continued)*

	Place of study: Leuven, Belgium *Period of study*: not specified; samples collected since 1999 *Language*: English
Index tests	*Index test*: IL-6, IL-8, TNF-α, hsCRP, CA-125, CA-19.9 *Details of the index test procedure as stated*: plasma concentrations of IL-6, IL-8 and TNF-α were determined by using commercially available ELISA kits (BD Biosciences, Erembodegem,Belgium) according to the manufacturer's instructions. Plasma concentrations of CA-125, CA-19.9 and hsCRP levels were measured by automated assays on a Roche Modular P or Modular E170 instruments (Roche, Vilvoorde, Belgium) at the central laboratories of the university Hospitals Leuven (Gasthuisberg, Leuven). The predictive model was built by using a multivariate analysis (stepwise logistic regression with and without LSSVM analysis *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 201/294 (68%): stage I-II 132, stage III-IV 69; controls n = 93 *Reference standard*: laparoscopy n = 294 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection with histological confirmation for most of the samples; rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	Advanced statistical analysis of a panel of 6 selected plasma biomarkers on samples obtained during the secretory phase or during menstruation allows the diagnosis of both minimal–mild and moderate–severe endometriosis with high sensitivity and clinically acceptable specificity
Conflict of interest	Not reported; supported by a TBM (Toegepast Biomedisch Onderzoek met Primair Maatschappelijke Finaliteit) grant from the Institute for Innovative Science and Technology IWT (Innovatie door Wetenschap en technologie) in Flanders, Belgium
Notes	The reported diagnostic estimates according to severity of endometriosis are not presented in this review The diagnostic estimates for each individual marker were reported only for luteal cycle phase and were the result of univariate logistic regression model The diagnostic estimates for the combination of biomarkers were reported for the overall group and for each cycle phase and were the results of multivariate logistic regression and LS-SVM models

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Mihalyi 2010 *(Continued)*

Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Mihalyi 2010 *(Continued)*

	Low

Mohamed 2013

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the role of serum level of VEGF-A in comparison to CA-125 in the diagnosis and detection of recurrence of patients, with advanced endometriosis after conservative laparoscopic surgery *Participants*: women referred for laparoscopy for unexplained primary infertility, chronic pelvic pain or both *Selection criteria*: inclusion criteria: regular menses, follicular cycle phase; only patients with advanced disease selected; exclusion criteria: 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 *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*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 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 *Language*: English
Index tests	*Index test*: CA-125 and VEGF-A *Details of the index test procedure as stated*: serum VEGF was measured by Human VEGF Quantikine ELISA Kit (DVE00, R&D Systems, Minneapolis, MN) and 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) *Threshold for positive result*: CA-125 > 35 µg/ml, VEGF-A > 680 pg/ml; not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 30/60 (50%), all stage III-IV; controls n = 30 *Reference standard*: laparoscopy + histology N = 60 (100%) *Description of positive case definition by reference standard test as reported*: surgical diagnosis - reference to the source on morphologic criteria; confirmed by histopathology; staging according to the rASRM classification

Mohamed 2013 *(Continued)*

	Examiners: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The use of VEGF-A for diagnosis of advanced endometriosis at cut-off 680 pg/ml and for follow-up is better than CA-125
Conflict of interest	The authors reported no conflict of interest
Notes	—

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without	Yes		

Mohamed 2013 *(Continued)*

knowledge of the results of the index tests?			
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Molo 1994

Study characteristics	
Patient sampling	*Primary objective*: to evaluate CA-125 and CA-72 prior to diagnostic laparoscopy in women with infertility *Participants*: consecutive patients undergoing laparoscopy for infertility investigation *Selection criteria*: not specified *Study design*: cross-sectional single-gate, prospective recruitment and collection of samples, consecutive series
Patient characteristics and setting	*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 *Language*: English
Index tests	*Index test*: CA-125, CA-72 *Details of the index test procedure as stated*: plasma concentrations of CA-125 and CA-72 were measured by radioimmunoassay (Contocor Inc, Malvern, PA) *Threshold for positive result*: CA-125 > 35 U/ml, CA-72 > 4 U/ml, pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard

Molo 1994 *(Continued)*

	Interobserver variability: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 19/35 (54%): stages not specified; controls n = 16 *Reference standard*: laparoscopy N = 35 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection (endometriosis defined as classic powder burn lesion, areas of hypervascularity, petechial lesions, clear lesions and pseudoperitoneal pockets; suspicious areas confirmed by histopathology; staging according to rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples collected 1 week before scheduled laparoscopy *Withdrawals*: none
Comparative	
Key conclusions by the authors	There was no advantage in using CA-125 and CA-72 preoperatively to determine the likelihood of pelvic endometriosis. There is no evidence that these tumour-associated antigens are helpful in the routine work-up of the female infertility patient
Conflict of interest	Not reported
Notes	—

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it prespecified?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		

Molo 1994 *(Continued)*

		Unclear	**Low**
DOMAIN 3: Reference Standard			
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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Morin 2005

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the concentrations of MIF in the peripheral blood of normal women and patients with endometriosis *Participants*: women undergoing laparoscopy for infertility, pelvic pain, or tubal ligation. *Selection criteria*: inclusion criteria: no other pelvic pathology and no treatment with any anti-inflammatory or hormone medication at least 3 months before laparoscopy *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis group: pain - 42/55, infertility - 34/55; controls - fertile women requesting tubal ligation or reanastomosis *Age*: mean age 33.6 ± 4.7 years (endometriosis) and 36.7 ± 6.2 years (controls) *Number of participants enrolled*: 93 women *Number of participants available for analysis*: 93 women (47 in follicular and 45 in luteal cycle phase) *Setting*: University hospital, Saint-Francois d'Assise hospital Universite Laval *Place of study*: Quebec, Canada *Period of study*: not specified

Morin 2005 *(Continued)*

	Language: English
Index tests	*Index test*: MIF *Details of the index test procedure as stated*: serum concentrations of MIF measured by ELISA samples run in duplicate; concentrations extrapolated from a standard curve using recombinant human MIF; sample handling and laboratory method described *Threshold for positive result*: > 0.57 ng/ml, not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: the inter- and intra-assay CV 2.9% and 3.8%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 55/93 (54%): stage I-II 36, stage III-IV 19; controls n = 38 *Reference standard*: laparoscopy N = 93 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were drawn a few days before laparoscopy *Withdrawals*: none
Comparative	
Key conclusions by the authors	This study showed a marked increase in MIF concentrations in the peripheral blood of women with endometriosis and a relationship with disease progress, and suggests that MIF may be involved in endometriosis-related pain and infertility
Conflict of interest	Not reported; supported by grant MOP-37921 from The Canadian Institutes for Health Research. AA is a Chercheur-Boursier National of the Fonds de la Recherche en Santé du Québec (FRSQ).
Notes	The presented data enabled calculation of diagnostic estimated per severity of endometriosis - not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**

Morin 2005 *(Continued)*

DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Muscatello 1992

Study characteristics	
Patient sampling	*Primary objective*: to verify the clinical usefulness of CA-125, TAG-72 and CA-15.3 in the diagnosis of endometriosis either by themselves, or when combined *Participants*: women who underwent laparoscopy for infertility, pelvic pain or both at the authors' institution *Selection criteria*: not specified

Muscatello 1992 *(Continued)*

	Study design: cross-sectional single-gate, prospective collection of samples, non-consecutive series
Patient characteristics and setting	*Clinical presentation*: infertility, pelvic pain or both *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 *Language*: English
Index tests	*Index test*: CA-123, CA-15.3 and TAG-72 *Details of the index test procedure as stated*: serum concentrations of CA-125 and CA-15.3 measured by using a commercially available radioimmunoassay (CIS Diagnostici); serum levels of TAG-72 assessed by using a solid-phase double-determinant radio immunometric assay (Centocor); all assays were performed in duplicate; concentration assessed with a standard curve; sample handling described *Threshold for positive result*: CA-125 > 35 U/ml; CA-15.3 > 30 U/ml; TAG-72 > 6 U/ml; all pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: the intra-and interassay CV 8% and 15% for CA-125
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 81/119 (68%): stage I-II 31, stage III-IV 50; controls n = 38 *Reference standard*: laparoscopy N = 119 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were taken immediately before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Measurement of serum CA-15.3 and TAG- 72 in addition to CA-125 does not provide any advantage for the diagnosis of endometriosis
Conflict of interest	Not provided
Notes	The presented data enabled calculation of diagnostic estimated per severity of endometriosis - not presented in this review

Muscatello 1992 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Odukoya 1996

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the serum concentration of sCD23 and the serum endometrial IgG antibody in patients with endometriosis to determine if B cell activation occur in these patients *Participants*: fertile patients with chronic pelvic pain who underwent laparoscopy and were diagnosed with endometriosis (endometriosis group) and fertile pain-free patients who at laparoscopic tubal sterilisation were found to have normal pelvis (controls) *Selection criteria*: not specified *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified; all patients have regular menstrual cycle (22-35 days) *Age*: mean age 33.5 ± 5.7, range 21-45 years (endometriosis); 32.6 ± 6.8, range 25-45 years (controls) *Number of participants enrolled*: 97 women *Number of participants available for analysis*: 97 women (55 follicular and 42 luteal phase) *Setting*: University Department, Jessop Hospital for Women *Place of study*: Sheffield, UK *Period of study*: not stated *Language*: English
Index tests	*Index test*: sCD23 (soluble CD23) and endometrial IgG auto-Ab *Details of the index test procedure as stated*: serum sCD23 concentration was estimated by chemilumescent ELISA; endometrial Ab was measured with ELISA, laboratory techniques described in details *Threshold for positive result*: positivity defined as absorbance value of the ELISA > than the plate control mean ± SD (male and postmenopausal serum); threshold pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Reference standard*: laparoscopy + histology, N = 97 (100%) *Prevalence of target condition in the sample*: n = 57/97 (59%): stage I-II 40, stage III-IV 17; controls n = 40 *Reference standard*: laparoscopy + histology N = 97 (100%) *Description of positive case definition by reference standard test as reported*: surgical diagnosis confirmed by histology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood was taken at the time of laparoscopy *Withdrawals*: none

Odukoya 1996 *(Continued)*

Comparative	
Key conclusions by the authors	These data suggest the existence of B cell activation in patients with endometriosis with a significant correlation between endometrial antibodies and sCD23. Mild endometriosis appears to be immunologically more active than the severe form. The value of sCD23 in the management of endometriosis needs further evaluation
Conflict of interest	Not reported; financial assistance from Lederle Laboratories, Gosport, England
Notes	—

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**Low**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**

Odukoya 1996 *(Continued)*

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Ohata 2008

Study characteristics	
Patient sampling	*Primary objective*: to determine if serum concentration of serum IL-8 can be found in ovarian endometrioma and if this is a useful tool for diagnosing this disease *Participants*: women who underwent laparoscopy or laparotomy for endometrioma or other benign ovarian cysts *Selection criteria*: inclusion criteria: preoperative imaging suggestive of ovarian cyst; exclusion criteria: suspected infectious diseases, chronic or acute inflammatory diseases, malignancy, autoimmune diseases, artificial grafts or ruptured endometrioma *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: mean 35.5 ± 8.0, range 20-48 years (endometriosis); 36.0 ± 10.6, range 20-50 years (controls) *Number of participants enrolled*: 91 women *Number of participants available for analysis*: 91 women (44 follicular and 37 luteal phase) *Setting*: Tottori University Hospital *Place of study*: Yonago, Japan *Period of study*: 2001-2006 *Language*: English
Index tests	*Index test*: IL-8 and CA-125 *Details of the index test procedure as stated*: serum concentrations were measured with immunoassays: IL-8 (Quantikine; R&D Systems Inc, Minneapolis, MN), range 3.5-2,000 pg/ml; CA-125 (ChemiLumi ACS-CA-125 II; Bayer Medical Co. Ltd, Tokyo, Japan), range 2 to 600 U/ml; sample processing and laboratory techniques described *Threshold for positive result*: IL-8 ≥ 25 pg/ml; CA-125 ≥ 30 U/ml, not pre-specified *Examiners*: no information provided; unclear if blinded to the results of reference standard *Interobserver variability*: Inter- and intra-assay CV < 10% for both tests

Ohata 2008 *(Continued)*

Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 70/91 (77%), all stage III-IV; controls n = 21 *Reference standard*: laparoscopy/laparotomy + histology N = 91 (100%) *Description of positive case definition by reference standard test as reported*: Surgical diagnosis confirmed by pathologic examination; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained before surgery *Withdrawals*: for CA-125 the data was missing for 5 cases and 3 controls, withdrawals not explained
Comparative	
Key conclusions by the authors	Serum levels of IL-8 could improve diagnostic reliability; further studies are needed for IL-8 to be used as a reliable serum marker in the clinical management of endometriosis
Conflict of interest	Not reported
Notes	—

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			

Ohata 2008 *(Continued)*

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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Oku 2004

Study characteristics	
Patient sampling	*Primary objective*: to elucidate the role of IL-18 in the pathogenesis of endometriosis *Participants*: women undergoing surgery for suspected endometriosis, ovarian mass or infertility *Selection criteria*: not specified *Study design*: observational, single-gate design, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: endometriosis: not specified, controls: benign ovarian cysts - 13, fibroid uterus - 2, infertility - 4; all the women had normal ovulatory cycles and did not take hormonal medication for at least 3/12 months before surgery *Age*: mean age 33.8 ± 6.8 years, range 24-48 years (endometriosis group), 31.7 ± 6.7 years, range 20-46 years (controls) *Number of participants enrolled*: 58 women *Number of participants available for analysis*: 58 women (all in follicular cycle phase) *Setting*: Department of O&G, Institute for Advanced Medical Sciences and Hyogo College of Medicine *Place of study*: Hyogo, Japan *Period of study*: not stated *Language*: English
Index tests	*Index test*: IL-18, IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, TNF-α, GM-CSF and IFN-γ

Oku 2004 *(Continued)*

	Details of the index test procedure as stated: serum IL-18 levels were determined by using ELISA commercial kit (MBL Co. Ltd, Nagoya, Japan); IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, TNF-α, GM-CSF and IFN-γ were determined by Bio-Plex Protein Array System (Bio-Rad Laboratories, Hercules, CA) using Human Cytokine Assay reagents (Bio-Rad) *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 39/58 (67%): stage I-II 6, stage III-IV 33; controls n = 19 *Reference standard*: Surgery (type of surgery not stated), N = 58 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rASRM classification . *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: the samples were collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The elevation of IL-18 in the peritoneal fluid of endometriosis patients and the induction of COX-II in peritoneal monocytes by IL-18 suggest that IL-18 plays a pathogenic role in endometriosis
Conflict of interest	Not reported
Notes	For IL-18 and IL-1β there was no statistically significant difference between the groups - no data available for meta-analysis For IL-2, IL-4, IL-6, IL-8, IL-10, TNF-α, GM-CSF and IFN-γ there was no statistically significant difference between the groups, but there was insufficient information to confirm negative findings - not included in the review The reported diagnostic estimates for peritoneal fluid biomarkers are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**

Oku 2004 *(Continued)*

DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Olkowska-Truchanowicz 2013

Study characteristics	
Patient sampling	*Primary objective*: to investigate the levels of CD4+ CD25+FOXP3+ Treg cells in the peripheral blood and peritoneal fluid of patients with endometriosis *Participants*: women undergoing laparoscopy for suspected endometriosis or ovarian cyst *Selection criteria*: not specified *Study design*: cross-sectional, single-gate, prospective sample collection

Olkowska-Truchanowicz 2013 *(Continued)*

Patient characteristics and setting	*Clinical presentation*: endometriosis group - not specified; controls - indications for surgery: benign ovarian cysts or diagnostic laparoscopy; none of the participants suffered from any other chronic inflammatory or autoimmune disorder and was not subjected to pharmacological treatment which would affect immune response for at least 3/12 months prior to the study *Age*: mean age 31 years, range 19-39 years (endometriosis group), 34 years, range 18-46 years (controls) *Number of participants enrolled*: 32 participants *Number of participants available for analysis*: 32 participants (all in follicular cycle phase, day 5-10) *Setting*: Department of O&G, Militay institute of Medicine and research laboratory, Medical University of Warsaw *Place of study*: Warsaw, Poland *Period of study*: not reported *Language*: English
Index tests	*Index test*: CD4+ CD25+ FOXP3+ regulatory T cells (Treg cells) *Details of the index test procedure as stated*: Treg cells were measured by flow cytometry using chlorophyll protein-conjugated anti-CD4 and allophycocyanin conjugated anti-CD25 monoclonal antibodies (all from BD Biosciences, San Jose, USA); followed by intracellular staining of FOXP3 using the fluorescein isothiocyanate (FITC) Anti-Human Foxp3 Staining Set (eBioscience Inc, San Diego, USA) according to the manufacturer's instructions; sample handling and laboratory technique described *Threshold for positive result*: not reported *Examiners*: not information provided, unclear if were blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 17/32 (53%): all stage III-IV; controls n = 15 *Reference standard*: laparoscopy, N = 32 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection with histological confirmation; staging according to rAFS classification *Examiners*: not reported
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Treg cells may play a part in immunopathogenesis of endometriosis, being responsible for abrogated local cellular immune responses and facilitation and development of autoimmune reactions. Treg cells may be thus a potential target in the treatment of endometriosis
Conflict of interest	The authors declared no conflict of interests; the work was supported by 1M15/N/2011 and NK1W grants from the I Faculty of Medicine, Warsaw Medical University
Notes	For CD25+ FOXP3+ and $CD25^{low}$ FOXP3+ cells, there was no statistically significant difference between the groups - no data available for meta-analysis

Olkowska-Truchanowicz 2013 *(Continued)*

For CD4+ and CD4+ CD25+ Treg cells. there was no statistically significant difference between the groups, but there was insufficient data to confirm negative findings - not included in this review

For $CD25^{high}$ FOXP3+, there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review

The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**

Olkowska-Truchanowicz 2013 *(Continued)*

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Othman 2008

Study characteristics	
Patient sampling	*Primary objective*: to test the ability of a group of serum cytokines, either individually or in combination, to serve as biomarkers for the non-surgical diagnosis of endometriosis *Participants*: women undergoing laparoscopy for the evaluation of infertility or pelvic pain *Selection criteria*: inclusion criteria: regular menstrual cycles, not on hormonal medications at least 3 months prior to enrolment, not been pregnant or hysterosalpingography done at least 3 months prior to enrolment *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility, pelvic pain *Age*: median 34.0, range 29.0–38.5 years (endometriosis group) and 32.0, range 28.5–36.5 years (controls) *Number of participants enrolled*: 131 women *Number of participants available for analysis*: 131 women (60 in follicular, 78 in luteal cycle phase) *Setting*: gynaecologic endoscopy unit, institution not specified Place of study: not stated; authors' affiliations include universities in USA, Germany, Egypt *Period of study*: not stated *Language*: English
Index tests	*Index test*: MCP-1, IL-6, VEGF, TNF-α, GM-CSF, INF-γ *Details of the index test procedure as stated*: serum cytokine concentrations were determined using the Bio-Plex Protein Array System (Bio-Rad, Hercules, CA, USA) with cytokine-specific antibody-coated beads (Bio-Rad) detecting range 0.2–32,000 pg/ml; sample processing and laboratory techniques described *Threshold for positive result*: IL-6 > 1.03 pg/ml, > 1.9 pg/ml, > 2.6 pg/ml; not pre-specified Examiners: no information provided; unclear if blinded to the results of reference standard

Othman 2008 *(Continued)*

	Interobserver variability: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 68/138 (49%): stage I-II 32, stage III-IV 36; controls n= 70 *Reference standard*: laparoscopy + histology n = 138 (100%) *Description of positive case definition by reference standard test as reported*: surgical diagnosis confirmed by pathologic examination, reference to the source on morphological criteria; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Serum IL-6 provided a good means of discrimination between subjects with endometriosis and controls; adding MCP-1 and IFN-γ to IL-6 did not improve the discrimination between subjects with endometriosis and controls over that achieved by using IL-6 alone
Conflict of interest	Not reported
Notes	For VEGF, TNF-α, GM-CSF there was no statistically significant difference between the groups - no data available for meta-analysis For MCP-1 and INF-γ there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review For IL-2, IL-8, IL-15 the concentrations were below the detection limit of the assay in each group

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		

Othman 2008 *(Continued)*

If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Ozhan 2014

Study characteristics	
Patient sampling	*Primary objective*: to investigate the diagnostic potentials of the serum levels of 9 different biomarkers in endometriosis *Participants*: women undergoing laparoscopy or laparotomy for evaluation of chronic pelvic pain, severe dysmenorrhoea, infertility, pelvic endometriosis or pelvic mass *Selection criteria*: exclusion criteria: autoimmune diseases, pelvic inflammatory disease, any malignancy, a history of delivery or abortion within the last 6/12 months, any endocrine disease, menopause, premature ovarian failure, menses, other pelvic masses out of endometrial adhesions or endometrioma, any anti-inflammatory or hormone medication within 3/12 months of operation *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility, pelvic pain, dysmenorrhoea, ovarian mass *Age*: mean age 32.3 ± 7.01 years (endometriosis group) and 34.2 ± 6.88 years (controls)

Ozhan 2014 *(Continued)*

	Number of participants enrolled: 80 women *Number of participants available for analysis*: 80 women (cycle phase not reported) *Setting*: Department of O&G, University of Ondokuz Mayis *Place of study*: Samsun, Turkey *Period of study*: over 1 year, dates not reported *Language*: English
Index tests	*Index test*: enolase, MIF, leptin, IL-8, AEA, PDPK1, CA-125, STX-5, LN-1 *Details of the index test procedure as stated*: serum biomarkers were measures using micro-ELISA method by the ELISA reader (awareness technology well model, USA); detection range for Enolase 1.25–80.00 ng/ml, for MIF 125–8000 pg/ml, for leptin > 0,04 ng/ml, for IL-8 > 1,1 pg/ml, for PDPK1 0.156–10.00 ng/ml, for CA-125 15–300 U/ml, for STX-5 23.4–1500.0 ng/ml, for LN-1 78–5000 pg/ml; sample processing described *Threshold for positive result*: CA-125 > 43 U/ml, STX-5 > 55 ng/ml, LN-1 > 1110.0 pg/ml; not pre-specified *Examiners*: no information provided; unclear if blinded to the results of reference standard *Interobserver variability*: the intra- and interassay CV for enolase, MIF and STX-5 was < 8% and < 10%; for AEA, < 15%; for PDPK1 and LN-1, < 10% and < 12%, for CA-125, < 15% and < 20%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 60/80 (75%): stage I-II - 18, stage III-IV - 42; controls n = 20 *Reference standard*: laparoscopy/ laparotomy N = 80 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained 1-2 hours before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Concurrent measurement of CA-125, syntaxin-5 and laminin-1 might be a useful non-invasive test in strengthening the diagnosis of endometriosis and in predicting its severity
Conflict of interest	Not reported; the study was supported by the scientific research funding of the University of Ondokuz Mayis: PYO.TIP.1904.12.038
Notes	For enolase, MIF, leptin, IL-8, AEA and PDPK1 there was no statistically significant difference between the groups - no data available for meta-analysis When the data are available for the whole group of endometriosis versus controls, the diagnostic estimates for separate stages of endometriosis are not included For CA-125 and LN-1 the diagnostic estimates were reported only for certain stages of endometriosis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Ozhan 2014 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Ozhan 2014 *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Paiva 2014

Study characteristics	
Patient sampling	*Primary objective*: to develop a test to discriminate between women suffering from pelvic pain associated with presence or absence of endometriosis, using symptom visual analogue scale (VAS) scores, demographic and lifestyle factors and known and novel plasma biomarkers *Participants*: women undergoing laparoscopy for evaluation of chronic pelvic pain, dysmenorrhoea, or dyspareunia *Selection criteria*: exclusion criteria: women on current hormonal therapy, failure to complete questionnaire, *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain, dysmenorrhoea, dyspareunia *Age*: mean age 27 years, range 18-44 years (endometriosis group) and 30 years, range 19-43 years (controls) *Number of participants enrolled*: 172 women *Number of participants available for analysis*: 101 women (in menstrual, proliferative or secretory cycle phase) *Setting*: Department of O&G, Royal Women's Hospital, University of Melbourne *Place of study*: Melbourne, Australia *Period of study*: May 2006 - February 2009 *Language*: English
Index tests	*Index test*: CA-125, MIF, GM-CSF, MCP-1, VEGF, IL-17, CNTF, GDNF, SOD3, GSH, NT4, vitamin E, annexin V, glycodelin, nitrotyrosine, NGF, leptin, sICAM *Details of the index test procedure as stated*: serum CA-125 and MIF were measures using 2-plex magnetic human circulating cancer biomarker panel, GM-CSF, MCP-1, VEGF, IL-1 - using 4-plex magnetic human cytokine panel kit (Millipore, USA), CNTF, GDNF, SOD3, GSH, NT4, vitamin E, annexin V, glycodelin, nitrotyrosine, NGF - using ELISA kits (Life Research, Australia) and leptin, sICAM - using ELISA kits (R&D Systems, USA); detection limit for CA-125 - 0.26 pg/ml, MIF - 30 pg/ml, CNTF - 3.2 pg/ml, GDNF - 39 pg/ml, SOD3 - 3.9 pg/ml, GSH - 0.8 ug/ml, NT4 - 0.3 ng/ml, vitamin E - 0.2 µmol/ml, annexin V - 1.6 ng/ml, glycodelin - 0.78 ng/ml, nitrotyrosine - 0.16 ng/ml, NGF - 78 pg/ml, leptin - 31 pg/ml, sICAM - 24 pg/ml; laboratory methods and sample processing described *Threshold for positive result*: not reported *Examiners*: no information provided; unclear if blinded to the results of reference standard *Interobserver variability*: the intra- and interassay CV < 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 69/101 (68%): stage I-II 45, stage III-IV 24; controls n = 32 *Reference standard*: laparoscopy N = 101 (100%) + histopathology

Paiva 2014 *(Continued)*

	Description of positive case definition by reference standard test as reported: visual inspection confirmed by histological demonstration of endometrial glands and stroma; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained preoperatively *Withdrawals*: 71 participants were excluded: 16 due to current hormone treatment, 31 - not completed questionnaire, 24 - no samples available due to laboratory freezer failure
Comparative	
Key conclusions by the authors	Combining symptom scores, historical measures and CA-125 provides a reasonable means to discriminate between women with pelvic pain associated with presence or absence of endometriosis, but greater specificity is needed before such a model could replace laparoscopy
Conflict of interest	The authors declared no conflict of interests; the study was supported by several research grants
Notes	For MIF, GM-CSF, MCP-1, VEGF, IL-17, CNTF, GDNF, SOD3, GSH, NT4, vitamin E, annexin V, glycodelin, nitrotyrosine, NGF, leptin, sICAM there was no statistically significant difference between the groups - no data available for meta-analysis For CA-125 there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables solely for this marker - not included in this review The diagnostic estimates for a diagnostic model based on combination of demographic data, symptoms and CA-125 are not included in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it prespecified?	No		

Paiva 2014 *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Patton 1986

Study characteristics	
Patient sampling	*Primary objective*: to determine the efficacy of CA-125 measurements as a screening procedure for endometriosis *Participants*: women who underwent laparoscopy *Selection criteria*: inclusion criteria: no systemic diseases *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*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

Patton 1986 *(Continued)*

	Number of participants enrolled: 113 women *Number of participants available for analysis*: 113 women (menstrual cycle phase not specified) *Setting*: Department of O&G, Mayo Clinic, tertiary care centre *Place of study*: Rochester, Minnesota *Period of study*: January 1985 - June 1985 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: 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) *Threshold for positive result*: CA-125 > 35 U//ml; unclear if pre-specified *Examiners*: no information provided; unclear if blinded to the results of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 37/113 (33%): stage I-II - 22, stage III-IV - 15; controls n = 76: normal pelvis - 45, adhesions - 26, other - 5 *Reference standard*: laparoscopy + histology N = 113 (100%) *Description of positive case definition by reference standard test as reported*: surgical diagnosis confirmed by pathologic examination; endometriosis, pelvic adhesions, or other pelvic pathology were prospectively recorded; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained immediately before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The analysis of proteins with antigenic determinant CA-125 in patients with endometriosis and other disorders may be useful
Conflict of interest	Not reported
Notes	The reported diagnostic estimates for advanced endometriosis (stage III-IV) are nor presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Patton 1986 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Philippoussis 2004

Study characteristics

Philippoussis 2004 *(Continued)*

Patient sampling	*Primary objective*: to evaluate whether the levels of the circulating factors involved in gynaecologic cancers, such as AFP, IGFBP-3, c-erbB-2 and EGF are modulated in the serum of patients with endometriosis *Participants*: women who were scheduled to undergo laparoscopy or celiotomy at one of the 8 clinical institutions of the Montreal area (for various indications) *Selection criteria*: inclusion criteria: pre-menopausal age, not currently menstruating, regular menstrual cycles, no acute salpingitis, not pregnant, not under hormonal treatment for the past 3 months *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: indications for surgery: tubal ligation or reanastomosis - 40%, hysterectomy/ovariectomy - 22%, diagnostic laparoscopy - 38%; symptoms not specified; history of acute infection - 39% controls, 36% cases; leiomyoma - 11% controls, 17% cases *Age*: mean 35.2 ± 6.5 years (endometriosis group) and 36.3 ± 5.4 years (controls) *Number of participants enrolled*: 72 women *Number of participants available for analysis*: 72 women (all in luteal phase of menstrual cycle) *Setting*: biotech firm - MetrioGene BioSciences (a subsidiary of PROCREA BioSciences) *Place of study*: Montreal, Quebec, Canada *Period of study*: not specified *Language*: English
Index tests	*Index test*: AFP, IGFBP-3, c-erbB-2, EGF *Details of the index test procedure as stated*: serum levels of AFP, IGFBP-3, c-erbB-2, EGF were determined in ELISA commercial kits (AFP and IGFBP-3 Diagnostic Systems Laboratories, TX), (c-erB-2 Bender MedSystems, Austria), (EGF Quantikine, R&rDdingaSlystems, MN); the sensitivity of AFP, IGFBP-3, c-erbB-2, EGF assays was 0.7 pg/ml, 0.04 ng/ml, 0.1 ng/ml, 0.7 pg/ml; sample handling and laboratory techniques described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: Intra-and interassay CVs < 10% for all the assays
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 36/72 (50%), stage I-II 26, stage III-IV 10; controls - 36 *Reference standard*: laparoscopy/laparotomy N = 72 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rAFS system *Examiners*: gynaecologists collaborating in this study were trained surgeons experienced with the management of endometriosis and skilled to detect and identify all forms of endometriotic lesions
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	

Philippoussis 2004 *(Continued)*

Key conclusions by the authors	Although AFP, IGFBP-3, c-erbB-2, and EGF are not altered in the circulation of patients with endometriosis, their involvement in the development of endometriotic lesions cannot be excluded
Conflict of interest	Not reported; the authors are affiliated to the biomedical company; supported by a grant #15453Q of IRAP from the NSERC and by internal resources at PROCREA BioSciences, Canada
Notes	For AFP, IGFBP-3, c-erbB-2, and EGF there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	No		
		High	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			

Philippoussis 2004 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Pittaway 1989

Study characteristics	
Patient sampling	*Primary objective*: to determine whether serum CA-125 would be useful in differentiating between pelvic pain caused by endometriosis and that from other causes *Participants*: reproductive-aged women scheduled for laparoscopy or laparotomy for investigation of chronic pelvic pain with or without infertility *Selection criteria*: inclusion criteria: reproductive age, pain lasting at least 3 months *Study design*: cross-sectional single-gate design, prospective recruitment
Patient characteristics and setting	*Clinical presentation*: pelvic pain ± infertility *Age*: mean age 28.9 years, range 16-39 (endometriosis) and 26.7 years, range 14-44 years (controls) *Number of participants enrolled*: 180 women *Number of participants available for analysis*: 163 women (all in in late follicular phase of menstrual cycle, day 7-10) *Setting*: Section on Reproductive Endocrinology, Wake Forest School of Medicine, tertiary referral centre *Place of study*: Winston Salem, North Carolina, USA *Period of study*: over 30 months period, dates not provided *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum CA-125 was measured in duplicate in using an immunoradiometric assay (Centocor, Malvern, PA); sample handling described, reference to a source describing laboratory technique *Threshold for positive result*: CA-125 ≥16 U/ml; pre-specified *Examiners*: no information provided; operators of index test were blinded to surgical data *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis

Pittaway 1989 *(Continued)*

	Prevalence of target condition in the sample: n = 82/163 (50%): stage I-II 54, stage III-IV 28; controls n = 81: normal pelvis - 15, adhesions - 27, chronic PID - 28, other - 11 *Reference standard*: laparoscopy n = 163 (100%) *Description of positive case definition by reference standard test as reported*: surgical diagnosis; staging according to the rASRM classification *Examiners*: no information provided; CA-125 levels were not known at the time of surgery
Flow and timing	*Time interval between index test and reference standard*: preoperative 7-10 days before onset of last menses *Withdrawals*: 17 women were excluded from the study (were still menstruating on a day of sample collection)
Comparative	
Key conclusions by the authors	Determination of CA-125 may assist in the evaluation and treatment of women with chronic pelvic pain
Conflict of interest	Not reported
Notes	The reported data enabled calculation of the diagnostic estimates per severity of endometriosis - not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Low	**Low**
DOMAIN 3: Reference Standard			

Pittaway 1989 *(Continued)*

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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Podgaec 2007

Study characteristics	
Patient sampling	*Primary objective*: to analyse the interaction between Th1 and Th2 immune response patterns and endometriosis by evaluating a panel of cytokines *Participants*: women undergoing laparoscopy for suspected endometriosis *Selection criteria*: inclusion criteria: age 18–40 years, histologically confirmed endometriosis (study group), absence of autoimmune disease, menstrual cycles of 26–32 days, no use of hormone therapy in 3/12 months before surgery *Study design*: cross-sectional single-gate, prospective collection of samples, consecutive patients
Patient characteristics and setting	*Clinical presentation*: clinically suspected endometriosis *Age*: mean age 32.1 ± 5.4 years (endometriosis group), 32.9 ± 5.1 years (controls) *Number of participants enrolled*: 98 women *Number of participants available for analysis*: 98 women (in follicular or luteal cycle phase) *Setting*: endometriosis clinic, Department of O&G, Universidade de São Paulo *Place of study*: São Paulo, Brazil *Period of study*: January 2004 - November 2005 *Language*: English
Index tests	*Index test*: TNF-α, IFN-γ, IL-2, IL-4, IL-10

Podgaec 2007 *(Continued)*

	Details of the index test procedure as stated: serum biomarkers assessed by using the BD Cytometric Bead Array (CBA), (Pharmingen, Becton Dickinson, USA) and carried out using a flow cytometer (BD FACSCalibur, USA); sample handling and laboratory methods described *Threshold for positive result*: presence or absence of the selected mass protein peaks, not pre-specified *Examiners*: no information provided *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 63/98 (66%): stage I-II - 28, stage III-IV - 37; controls n = 33 *Reference standard*: laparoscopy N = 98 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed on histopathology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Endometriosis is an inflammatory disease involving a possible shift towards Th2 immune response component, as demonstrated by the relative increase in cytokines characteristic of this pattern of immune response
Conflict of interest	Not reported; the work was supported by grant 05/01218-3 from the SP State Foundation
Notes	For TNF-α, IFN-γ, IL-2, IL-4 and IL-10, there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge	Unclear		

Podgaec 2007 *(Continued)*

of the results of the reference standard?			
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Ramos 2012

Study characteristics	
Patient sampling	*Primary objective*: to evaluate serum concentrations of CA-125 and soluble CD-23 and to correlate them with clinical symptoms, localisation and stage of pelvic endometriosis and histological classification of the disease *Participants*: patients undergoing laparoscopy for suspected endometriosis based on symptoms, examination or imaging findings *Selection criteria*: inclusion criteria: age 18-45 years, no hormone therapy within 3 months prior to consultation, no autoimmune diseases confirmed by history and laboratory tests, evidence of ovarian function *Study design*: cross-sectional, single-gate design, prospective collection of samples; consecutive series

Ramos 2012 *(Continued)*

Patient characteristics and setting	*Clinical presentation*: chronic pelvic pain - 59/104, deep dyspareunia - 43/104, dysmenorrhoea - 82/104 *Age*: range 18-45 years *Number of participants enrolled*: 104 women *Number of participants available for analysis*: 102 women (all in menstrual and all in late proliferative cycle) *Setting*: endometriosis division, Department of O&G, Universidade de São Paulo *Place of study*: São Paulo, Brazil *Period of study*: June 2007 - October 2010 *Language*: English
Index tests	*Index test*: CA-125, sCD-23 *Details of the index test procedure as stated*: serum concentrations of CA-125 and sCD-23 were measured by using a a commercial sandwich ELISA kit (Elecsys®, Roche, USA and Bender MedSystems, Vienna, Austria) according to manufacturer's instructions; the analyte range of CA-125 and sCD-23 is 25-35 IU/ml and 10-91 U/ml, respectively; sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 44/102 (43%): stage I-II 19, stage III-IV 25; controls n = 58 *Reference standard*: laparoscopy n = 102 (100%) + histology *Description of positive case definition by reference standard as reported*: visual inspection, histology of the excised lesions; classification according to the rASRM score; referenced to the source of histological criteria *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected up to 3 months before surgery *Withdrawals*: 2 participants left the study
Comparative	
Key conclusions by the authors	The concentrations of CA-125 were higher in patients with endometriosis than in patients without the disease. There were no significant differences for soluble CD-23 levels between groups
Conflict of interest	Not reported
Notes	For sCD-23 there was no statistically significant difference between the groups - no data available for meta-analysis For CA-125 there was no statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review

Ramos 2012 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Ramos 2012 *(Continued)*

	Low

Randall 2007

Study characteristics	
Patient sampling	*Primary objective*: to investigate the relationship between laparoscopic diagnosis of endometriosis and results of a serum anti-endometrial antibody (AEA) assay *Participants*: patients presenting to their physicians with dysmenorrhoea, chronic pelvic pain or infertility, who subsequently underwent laparoscopy *Selection criteria*: not specified *Study design*: cross-sectional single-gate, multicentre, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain, dysmenorrhoea or both, n = 145, infertility, n = 382 *Age*: mean age 31.8 ± 6.5 years *Number of participants enrolled*: 2609 women *Number of participants available for analysis*: 527 women (cycle phase not specified) *Setting*: several medical centres - not specified; the authors' institutions include Department of O&G West Virginia University School of Medicine; Fertility and Endocrinology Center, Bristol, TN; the New Hope Center for Reproductive Medicine, Virginia; Canterbury Women Health Care, Fresno, CA; Abingdon Healthcare for Women; Appalachian Ob/Gyn Associates, Kingsport, TN *Place of study*: USA *Period of study*: not specified *Language*: English
Index tests	*Index test*: IgG anti-endometrial Abs *Details of the index test procedure as stated*: anti-endometrial Ab immunoreactivity measured by indirect immunofluorescence assay, which utilised frozen sections of endometrium from hysterectomy specimens (performed for pelvic pain); AEA reactions were ranked as negative, positive or strongly positive based on fluorescence difference between negative controls and tested sera; sample handling and laboratory technique described *Threshold for positive result*: positive results were defined as glandular epithelial immunofluorescence greater than background as seen in negative female controls (male serum was utilised to assist in selection of female negative controls); threshold pre-specified *Examiners*: same laboratory investigator performed all analyses without prior knowledge of patient history *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 278/527 (53%): stage not specified; controls n = 249 *Reference standard*: laparoscopy N = 527 (100%) *Description of positive case definition by reference standard test as reported*: not reported

Randall 2007 *(Continued)*

	Examiners: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were taken within 1 year before surgery *Withdrawals*: 2082 women did not undergo surgery and were excluded
Comparative	
Key conclusions by the authors	The AEA assay is a very good screening test for patients suspected of having endometriosis and should be utilised prior to laparoscopy in diagnostic categories of dysmenorrhoea or chronic pelvic pain and infertility
Conflict of interest	Not provided
Notes	The reported data for women with pain, infertility or both who did not undergo laparoscopy are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		

Randall 2007 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes

		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Riley 2007

Study characteristics	
Patient sampling	*Primary objective*: to test local (PF) and systemic inflammatory markers in order to explore what parts of inflammation are activated in endometriosis, and test whether this was related to stage and symptoms of the disease *Participants*: patients with histologically confirmed endometriosis and controls undergoing surgery for benign gynaecological disorders *Selection criteria*: not specified *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: dyspareunia - 12/32, dysmenorrhoea - 21/32, other pelvic pain - 19/32, infertility, fibroids *Age*: median age (95% CI): 33 (29-36) years (endometriosis group), 37 (31-43) years (controls) *Number of participants enrolled*: 32 women (14 in follicular, 14 in luteal cycle phase; 3 women were menopausal and 1 had undetermined cycle phase due to AUB) *Number of participants available for analysis*: 30 women *Setting*: Department of O&G, St. Olavs University Hospital *Place of study*: Trondheim, Norway *Period of study*: not provided *Language*: English
Index tests	*Index test*: CA-125, CRP *Details of the index test procedure as stated*: serum concentrations of CA-125 and CRP were measured by using the commercial kits (Elecsys, CA-125II Roche/Roche/Hitachi Modular Analytics E170, Germany and Tina-quant

Riley 2007 *(Continued)*

	CRPLX, Roche/Hitachi Modular Analytics E170, Roche) on the day of collection; sample processing not described *Threshold for positive result*: CA-125 > 35 kU/l, not pre-specified *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 18: stage I - 10, stage III-IV - 8; controls n = 14 *Reference standard*: laparoscopy n = 32 (100%) *Description of positive case definition by reference standard as reported*: staging according to the rAFS score *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: 2 women were excluded (1 - ovarian abscess diagnosed at surgery, 1 - on NSAIDs for rheumatoid arthritis)
Comparative	
Key conclusions by the authors	Neutrophil granulocytes in endometriosis patients may have a lowered ability to respond to weak activation signals, while in more extensive endometriosis stronger neutrophil activation may be related to a pro-inflammatory effect of endometriotic tissue
Conflict of interest	Not reported
Notes	For CA-125 and CRP there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge	Unclear		

Riley 2007 *(Continued)*

of the results of the reference standard?			
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Rosa E Silva 2007

Study characteristics	
Patient sampling	*Primary objective*: to define the serum CA-125 values that best indicate the presence and stage of endometriosis *Participants*: pre-menopausal women who had undergone diagnostic laparoscopy for pelvic pain or infertility *Selection criteria*: exclusion criteria were ovarian tumour (except endometriomas), pregnancy, PID, myomas or adenomyosis on echographic examination and hormonal treatment in the preceding 3 months *Study design*: cross-sectional single-gate design, prospective sample collection, consecutive series
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility or both

Rosa E Silva 2007 *(Continued)*

	Age: range 18-40 years *Number of participants enrolled*: 201 women *Number of participants available for analysis*: 201 women (all in follicular phase of menstrual cycle) *Setting*: Division of Human Reproduction and Gynecological Endoscopy, University of São Paulo, a tertiary referral centre *Place of study*: São Paulo, Brazil *Period of study*: not stated *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: no information provided *Threshold for positive result*: CA-125 > 10 IU/ml; > 20 U/ml; not pre-specified *Examiners*: no information provided *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 148/201 (74%): stage I-II 63, stage III-IV 85; controls n = 53 *Reference standard*: laparoscopy N = 201 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected one to two months preceding surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	In conclusion, it is not advisable to use serum levels of CA-125 as a diagnostic tool; sensitivity of CA-125 as a marker can be increased if used with other non-invasive methods such as TVUS or MRI
Conflict of interest	Not reported
Notes	The reported diagnostic estimates per severity of endometriosis are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Rosa E Silva 2007 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it prespecified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Unclear**
DOMAIN 3: Reference Standard			
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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Rosa E Silva 2014

Study characteristics	
Patient sampling	*Primary objective*: to assess the changes secondary to chronic inflammation in women with and without pelvic endometriosis by the determination of serum thiols and carbonyls *Participants*: women undergoing laparoscopy for suspected endometriosis or tubal ligation

Rosa E Silva 2014 *(Continued)*

	Selection criteria: exclusion criteria: smoking, use of anti-inflammatory medications in 2/12 months before surgery, ovarian tumour, PID, adenomyosis, fibroid uterus, pregnancy, hormonal therapy in 3/12 months preceding surgery *Study design*: cross-sectional two-gate, prospective collection of samples, consecutive patients
Patient characteristics and setting	*Clinical presentation*: pelvic pain, dyspareunia, dysmenorrhoea, infertility; controls - symptomatic or asymptomatic women requesting tubal ligation *Age*: mean age 33.22 ± 6.22 years (endometriosis group), 32.49 ± 4.74 years (controls) *Number of participants enrolled*: 138 women *Number of participants available for analysis*: 108 women (cycle phase not specified) *Setting*: University Hospitals: Division of O&G, Faculty of Medicine of Ribeirao Preto, University of São Paulo and hospital Santa Casa de Misericordia of Curitiba *Place of study*: São Paulo, Brazil *Period of study*: not stated *Language*: English
Index tests	*Index test*: antioxidant substances: total thiols and carbonyls *Details of the index test procedure as stated*: serum thiols and carbonyls were determined using DTNB method; sample handling and laboratory methods described *Threshold for positive result*: thiols <396.44 µM; carbonyls <14.9 µM; not pre-specified *Examiners*: no information provided *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 67/108 (62%): stages of endometriosis not specified; controls n = 41 *Reference standard*: laparoscopy N = 108 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: endometriosis diagnosed at laparoscopy with histologic confirmation *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before surgery (personal communication with the authors) *Withdrawals*: 30 women were excluded before analysis due to haemolysis or high lipid concentration in the samples
Comparative	
Key conclusions by the authors	The serum thiol levels revealed an increase in oxidative stress related to the development of pelvic endometriosis
Conflict of interest	Not reported
Notes	—
Methodological quality	

Rosa E Silva 2014 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Salehpour 2009

Study characteristics	
Patient sampling	*Primary objective*: to evaluate non-invasive and practical diagnostic methods by measuring serum and peritoneal fluid CA-125 levels in patients with endometriosis *Participants*: women who underwent laparoscopy because of infertility, chronic pelvic pain, or recurrent abortion *Selection criteria*: exclusion criteria: hormonal therapies within 6 months prior to laparoscopy, ovarian neoplasia and other cancers, PID or large uterine myomas *Study design*: cross-sectional single-gate, multicentre, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation*: primary infertility - 46/60, secondary infertility - 10/60, chronic pelvic pain - 7/60, dysmenorrhoea - 23/60, dyspareunia - 10/60, recurrent abortion - 3/60 patients *Age*: mean age 28.94 ± 4.34 years (endometrioma group), 28.36 ± 4.02 years (controls) *Number of participants enrolled*: 60 women *Number of participants available for analysis*: 60 women (all in early follicular cycle phase) *Setting*: Infertility and Reproductive Health Research Centre, Shahid Beheshti University of Medical Sciences *Place of study*: Tehran, Iran *Period of study*: 2008-2009 *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum CA-125 levels measured in duplicate by using a 2010 Elecsys kit (Roche Diagnostic GmbH, USA) by ECLIA method with sensitivity of assay of 0.60 IU/ml; sample handling described *Threshold for positive result*: > 14.70 IU/ml - not pre-specified *Examiners*: no information provided *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 35/60 (58%): stage I-II 25, stage III-IV 10; controls n = 25 *Reference standard*: laparoscopy N = 60 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by pathologic evaluation of the biopsy specimen; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before general anaesthesia *Withdrawals*: none

Salehpour 2009 *(Continued)*

Comparative	
Key conclusions by the authors	Serum and peritoneal fluid CA-125 levels are simple and non-surgical tools for diagnosing and staging pelvic endometriosis. These markers are of greater diagnostic value in higher stages of the disease
Conflict of interest	Not reported
Notes	The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			

Salehpour 2009 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Seeber 2008

Study characteristics	
Patient sampling	*Primary objective*: to evaluate whether a combination of putative markers of inflammation and CA-125 could serve as a multiple-marker screening test for endometriosis in a heterogeneous population of patients *Participants*: women undergoing laparoscopy for infertility, pelvic pain, tubal sterilisation or tubal reversal, or other benign aetiology *Selection criteria*: inclusion criteria: reproductive age, at least stage II of endometriosis *Study design*: cross-sectional two-gate, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation*: pain 61/141, infertility 27/141, BTL 27/141, other benign conditions 6/141; OCP use 31/141 *Age*: mean age 34 years, range 18–48 years (endometriosis group), 33 years, range 23–48 years (controls) *Number of participants enrolled*: 197 women *Number of participants available for analysis*: 141 women (91 in follicular, 25 in luteal and 25 in unknown cycle phase) *Setting*: Center for Research in Reproduction and Women's Health, Department of O&G, University of Pennsylvania School of Medicine *Place of study*: Philadelphia, Pennsylvania *Period of study*: December 2003 - November 2005 *Language*: English
Index tests	*Index test*: IL-6, TNF-α, MIF, MCP-1, IFN-γ, leptin, and CA-125 *Details of the index test procedure as stated*: serum concentrations of 7 markers evaluated by using commercially available ELISA kits (R&D Systems, Inc, MN and Panomics, Inc, CA for CA-125); the sensitivities of the IL-6, TNF-α, MIF, MCP-1, IFN-γ, leptin, and CA-125 ELISAs were 0.70, 1.60, 0.017, 5.00, 8.00, and 780.00 pg/ml and 5.0 U/ml, respectively; sample handling described; diagnostic performance of the markers then was evaluated jointly by using CART analysis with automatic self-validation procedures *Threshold for positive result*: CA-125 > 20 mIU/ml; MCP-1 > 76.4 pg/ml, > 152.744 pg/ml, > 53.451 pg/ml; leptin >3.14 pg/ml, > 29.1 pg/ml; MIF >14.7 ng/ml - all not pre-specified *Examiners*: no information provided

Seeber 2008 *(Continued)*

	Interobserver variability: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 63/141 (45%): stage II - 22, stage III-IV - 41; controls n = 78 *Reference standard*: laparoscopy N = 141 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: serum was obtained on the day of surgery *Withdrawals*: 56 participants were excluded before analysis (diagnosed with stage I endometriosis)
Comparative	
Key conclusions by the authors	Using the serum concentration of 4 markers in a 2-tiered decision rule, nearly half of the subjects in this population would have been diagnosed (and could have avoided surgery) with 93% accuracy
Conflict of interest	Not reported
Notes	The diagnostic estimates were reported only for the combination of biomarkers The reported diagnostic estimates were calculated by using a marker classification tree For TNF-α and IL-6 levels, there was no difference between the groups, and these markers were not included in the diagnostic model; there were insufficient data available to present these 2 markers in the review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		

Seeber 2008 *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Seeber 2010

Study characteristics	
Patient sampling	*Primary objective*: to identify potential novel biomarkers that differ between subjects with and without endometriosis and that might aid in developing a non-invasive, serum-based diagnostic test *Participants*: women undergoing laparoscopy for the indications of infertility, pelvic pain, tubal sterilisation or tubal reversal, or other benign aetiology *Selection criteria*: inclusion criteria: reproductive age, at least stage II of endometriosis *Study design*: cross-sectional two-gate, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation*: pain 61/141, infertility 27/141, BTL 27/141, other benign conditions 6/141; OCP use 31/141 *Age*: mean age 34 years, range 18–48 years (endometriosis group), 33 years, range 23–48 years (controls) *Number of participants enrolled*: 197 women

Seeber 2010 *(Continued)*

	Number of participants available for analysis: 139 women (91 in follicular, 25 in luteal and 25 in unknown cycle phase) *Setting*: Center for Research in Reproduction and Women's Health, Department of O&G, University of Pennsylvania School of Medicine *Place of study*: Philadelphia, Pennsylvania *Period of study*: December 2003 - November 2005 *Language*: English
Index tests	*Index test*: serum proteome by SELDI-TOF-MS (5 proteins with molecular mass of 1629.00 Da, 3047.00 Da, 3526.00 Da, 3774.00 Da, 5046.00 Da and 5068.00 Da) *Details of the index test procedure as stated*: serum proteome assessed by using 8-spot CM10 chip arrays. Mass spectrometry analysis was performed using a PBS-II ProteinChip reader (Ciphergen Biosystems, CA); spectra were collected using ˜ 165 laser shots (laser intensity of 170, detector sensitivity of 8, molecular mass range of 1000–10,000 Da); autodetection settings for peak determination with signal-to-noise ratio 5:1 on first pass and 2:1 of second pass; sample handling and method described; diagnostic performance of the markers then was evaluated jointly by using CART analysis with automatic self-validation procedures and 10-fold cross validation *Threshold for positive result*:**p**resence or absence of the selected mass protein peaks, not pre-specified *Examiners*: no information provided *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 63/141 (45%): stage II - 22, stage III-IV - 41; controls n = 78 *Reference standard*: laparoscopy n = 141 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: serum was obtained on the day of surgery *Withdrawals*: 56 participants were excluded before analysis (diagnosed with stage I endometriosis) and 2 participants excluded due to poor sample quality
Comparative	
Key conclusions by the authors	This study is the critical first step in the identification of potential novel biomarkers of endometriosis. Future identification of the proteins and further validation in a second population is needed before applying these findings in clinical practice
Conflict of interest	Not reported
Notes	The reported diagnostic estimates were calculated by using a 2-step classification tree

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Seeber 2010 *(Continued)*

Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Somigliana 2002

Study characteristics	
Patient sampling	*Primary objective*: to investigate the hypothesis that sICAM-1 may be used as a new serum marker of endometriosis *Participants*: women who underwent gynaecologic laparoscopy at the authors' institution *Selection criteria*: inclusion criteria: reproductive age, no hormonal treatment for at least 3 months before surgery; no history of endometritis or autoimmune, liver, endocrine, or neoplastic disorders; exclusion criteria: laparoscopic diagnosis of PID or malignancy *Study design*: cross-sectional single-gate, prospective recruitment and collection of samples, consecutive series
Patient characteristics and setting	*Clinical presentation*: endometriosis group - not specified; control group: pelvic pain, infertility or both - 13/49, uterine fibroids - 7/49, benign ovarian cysts - 25 *Age*: reproductive age, not specified *Number of participants enrolled*: 120 women *Number of participants available for analysis*: 120 women (different cycle phases, not specified) *Setting*: an academic department specialising in gynaecologic laparoscopy - University of Milan, Istituto Auxologico Italiano, and Istituti Clinici di Perfezionamento *Place of study*: Milan, Italy *Period of study*: December 1998 - January 2000 *Language*: English
Index tests	*Index test*: sICAM-1, CA-125 *Details of the index test procedure as stated*: serum sICAM-1 levels assessed by using a commercially available ELISA kit (Bender MedSystem, Austria); serum CA-125 level measured by using a commercially available chemiluminescent immunometric assay (Diagnostic Products Corporation, CA); sample handling described *Threshold for positive result*: sICAM-1 > 381 ng/ml; CA-125 > 37 IU/ml - not pre-specified *Examiners*: no information provided *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 71/120 (59%): stage I-II - 24, stage III-IV - 47, DIE - 21; controls n = 49) *Reference standard*: laparoscopy N = 120 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection (DIE, defined as lesions infiltrating to a depth of at least 5 mm beneath the peritoneal surface), histological confirmation of other benign pelvic conditions; staging according to the rASRM classification *Examiners*: surgery was performed by 1 of the 3 physicians active in the evaluation and treatment of endometriosis
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before laparoscopy

Somigliana 2002 *(Continued)*

	Withdrawals: none
Comparative	
Key conclusions by the authors	Although the present study tends to support a role of sICAM-1 in the development of endometriosis, serum concentrations of this molecule do not seem to be an effective indicator for the diagnosis of either the early or advanced stage of endometriosis. However, an integrated clinical and laboratory approach using both CA-125 and sICAM-1 may be helpful in specifically identifying women with deep peritoneal endometriosis
Conflict of interest	Not reported
Notes	For sICAM-1 levels there was no statistically significant difference between the groups - no data available for meta-analysis The reported diagnostic estimates for CA-125 and sICAM were calculated for a subgroup with DIE versus the remaining cohort (women with and without endometriosis) and the estimates for endometriosis versus controls were not available, hence these estimates were not included in the review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly	Yes		

Somigliana 2002 *(Continued)*

classify the target condition?			
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Somigliana 2004

Study characteristics	
Patient sampling	*Primary objective*: to verify the clinical value of serum CA-125, CA-19.9 and IL-6 levels, either by themselves or combined, in the detection of endometriosis *Participants*: women who underwent gynaecologic laparoscopy for benign gynaecological pathologies *Selection criteria*: inclusion criteria: reproductive age, gynaecological indications for laparoscopic surgery; exclusion criteria: suspected or ascertained diagnosis of malignancy, pregnancy, menopausal age, refusal to participate in the study *Study design*: cross-sectional single-gate, prospective recruitment and collection of samples, consecutive series
Patient characteristics and setting	*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 - Department of O&G, Clinica L.Mangiagalli, University of Milano *Place of study*: Milan, Italy

Somigliana 2004 *(Continued)*

	Period of study: October 2002 - January 2003 *Language*: English
Index tests	*Index test*: CA-125, IL-6, CA-19.9 *Details of the index test procedure as stated*: serum levels of CA-125 and CA-19.9 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 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 *Threshold for positive result*: CA-125 >35 IU/ml, CA-19.9 >37 IU/ml, IL-6 >2 pg/ml - all pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*:**I**ntra-and interassay CV for IL-6: 2.5% and 4.5% (ELISA); 4% and 7% (immunometric assay)
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 45/80 (59%): stage I-II - 14, stage III-IV - 31; controls n = 35) *Reference standard*: laparoscopy N = 80 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the rASRM classification *Examiners*: the surgeries were performed at the department specialising in gynaecological laparoscopy
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The concomitant dosage of CA-125, CA-19.9 and IL-6 does not add significant information with respect to the CA-125 test alone in diagnosing either early or advanced stages of endometriosis
Conflict of interest	Not reported
Notes	For IL- and CA- 19.9 levels there was no statistically significant difference between the groups, but the diagnostic estimates were reported by the authors and presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Somigliana 2004 *(Continued)*

Was a 'two-gate' design avoided?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Steff 2004a

Study characteristics	
Patient sampling	*Primary objective*: to compare the measurements of serum levels of IGF-1, sTNFR-1 and angiogenin in serum of patients with endometriosis and controls *Participants*: patients who underwent laparoscopy or laparotomy for different indications *Selection criteria*: inclusion criteria: pre-menopausal age, not currently menstruating, regular menstrual cycles (21-35 days), no acute salpingitis, no pregnancy, hormonal treatment or IUD for the last 3 months *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified; indications for surgery included diagnostic evaluation, tubal ligation or reanastomosis, or hysterectomy *Age*: mean age 37.5 ± 5.9 years (endometriosis group), 35.7 ± 6.3 years (controls) *Number of participants enrolled*: 148 women *Number of participants available for analysis*: 148 women (77 in follicular, 71 in luteal cycle phase) *Setting*: MetrioGene BioSciences (a subsidiary of PROCREA BioSciences; patients recruited from several collaborating medical institutions - not specified *Place of study*: Montreal, Quebec, Canada *Period of study*: not provided *Language*: English
Index tests	*Index test*: IGF-1, sTNFR-1, angiogenin *Details of the index test procedure as stated*: serum concentrations were measured by using the commercial ELISA kits for sTNFR-1, angiogenin (Quantikine; R&D Systems, MN, USA) and for IGF-1 (Diagnostic Systems Laboratories, TX, USA); sample processing and laboratory methods described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 77/148 (52%): stage I-II - 52, stage III-IV - 25; controls n = 71 *Reference standard*: laparoscopy n = 148 (100%) *Description of positive case definition by reference standard as reported*: staging according to the rAFS score *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	

Steff 2004a *(Continued)*

Key conclusions by the authors	sTNFR-1 and angiogenin represent potential blood markers for endometriosis
Conflict of interest	Not reported (the authors' affiliation is MetrioGene BioSciences, a biotech firm)
Notes	For IGF-1 there was no statistically significant difference between the groups - no data available for meta-analysis For sTNFR-1 and Angiogenin in follicular cycle phase there was statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**

Steff 2004a *(Continued)*

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Suen 2014

Study characteristics	
Patient sampling	*Primary objective*: to investigate the modulatory role of IL-10 in the development of endometriosis *Participants*: patients who underwent surgery for various indications *Selection criteria*: exclusion criteria: any autoimmune disease, allergic disease, malignancy, or hepatitis B virus or hepatitis C virus infection, or any medical treatment or surgery within 3 months before the study-related surgery *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: indications for surgery: endometriosis group - for the treatment of advanced endometriosis; controls - for benign gynaecological conditions; all controls had regular menstrual cycles *Age*: mean age 34.0 ± 7.1 years (endometriosis group), 36.6 ± 7.9 years (controls) *Number of participants enrolled*: 67 women *Number of participants available for analysis*: 67 women (cycle phase not specified) *Setting*: Departments of O&G and Medical Research, Kaohsiung Medical University Hospital *Place of study*: Kaohsiung, Taiwan *Period of study*: not provided *Language*: English
Index tests	*Index test*: IL-6, IL-12, IL-10 *Details of the index test procedure as stated*: serum levels of IL-10 and IL-6 were determined using ELISA, with 2.0 pg/ml as the limit of detection; sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 41/67 (61%): all stage III-IV; controls n = 26

Suen 2014 *(Continued)*

	Reference standard: surgery, not specified N = 67 (100%) *Description of positive case definition by reference standard as reported*: staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected within 24 h before surgery (personal communication with the authors) *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	IL-10 may suppress immunity against endometrial implants, contributing to development of endometriosis
Conflict of interest	Not reported; the work supported by grants NSC-99-2628-B-037-009-MY3, NSC100-2314-B-037-043 and NSC 102-2628-B-037-011-MY3 from the National Science Council (Taiwan), and by grants KMUH101-1R27, KMUH100-0R24, KMUH 99-9I04 and KMUH 99-9R30 from Kaohsiung Medical University Hospital
Notes	For IL-6 and IL-12 there was no statistically significant difference between the groups - no data available for meta-analysis For IL-10 there was statistically significant difference between the groups, but there were insufficient data to construct 2 x 2 tables - not included in this review The data for healthy controls (N = 11) who did not undergo abdominal surgery were not included in this review The data for animal model of endometriosis are not included in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		

Suen 2014 *(Continued)*

If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Szczepanska 2001a

Study characteristics	
Patient sampling	*Primary objective*: to measure the levels of anti-gamete antibodies in serum and peritoneal fluid of women with endometriosis, infertility or both *Participants*: women who underwent laparoscopy for infertility, suspected endometriosis, chronic pelvic pain *Selection criteria*: not specified (only patients with minimal endometriosis were included) *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility, chronic pelvic pain or both *Age*: mean age 29 years, range 23–38 years *Number of participants enrolled*: 98 women *Number of participants available for analysis*: 98 women (all in luteal cycle phase)

Szczepanska 2001a *(Continued)*

	Setting: Clinic of Reproduction, Institute of Gynaecology and Obstetrics Poznan *Place of study*: Poznan, Poland *Period of study*: not provided *Language*: English
Index tests	*Index test*: anti-gamete Abs (anti-ZP Abs and antisperm Abs) *Details of the index test procedure as stated*: serum levels of anti-gamete Abs were assessed by using the quantitative ELISA (absorbance at 492 nm was determined by Multiscan Plus spectrophotometer (Labsystems Multiscan, Finland) and standard curve was plotted; protein concentration was was extrapolated from the standard curve and calculated per cell in both performed assays; sample processing and laboratory methods described in details *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 50/98 (51%): all stage I; controls n = 48 *Reference standard*: laparoscopy N = 98 (100%) *Description of positive case definition by reference standard as reported*: staging according to the rAFS score *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Antizona antibodies locally produced in the peritoneal fluid have diagnostic value for infertility status; however, they cannot be treated as a marker or prognostic factor for minimal endometriosis or its treatment
Conflict of interest	Not reported; supported by the Committee of Scientific Research of Poland and the Ministry of Health, Warsaw, Poland
Notes	For serum anti-ZP and anti-sperm antibodies there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Szczepanska 2001a *(Continued)*

Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Szczepanska 2001b

Study characteristics

Szczepanska 2001b *(Continued)*

Patient sampling	*Primary objective*: to measure the levels of anti-gamete antibodies in serum and peritoneal fluid of women with endometriosis, infertility or both *Participants*: women who underwent laparoscopy for suspected endometriosis or endometriosis recurrence *Selection criteria*: inclusion criteria: regular menstrual cycles, no hormonal therapy for 3/12 months preceding surgery *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain *Age*: mean age 29 years, range 23–38 years *Number of participants enrolled*: 64 women *Number of participants available for analysis*: 64 women (all in luteal cycle phase) *Setting*: Clinic of Reproduction, Institute of Gynaecology and Obstetrics Poznan *Place of study*: Poznan, Poland *Period of study*: 1998-1999 *Language*: Polish
Index tests	*Index test*: IL-12 *Details of the index test procedure as stated*: serum levels of IL-12 were assessed by using ELISA; sample processing and laboratory methods described in details *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 53/64 (83%): stage I-II - 21, stage III-IV - 20; recurrent endometriosis 12; controls n = 11 *Reference standard*: laparoscopy n = 64 (100%) + histopathology *Description of positive case definition by reference standard as reported*: visual inspection with histological confirmation; staging according to the rAFS score *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	There were no statistically significant differences in IL-12 levels in peritoneal fluid nor in serum in any of studied groups
Conflict of interest	Not reported

Szczepanska 2001b *(Continued)*

Notes	For serum IL-2 there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it prespecified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Szczepanska 2001b *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Szubert 2012

Study characteristics	
Patient sampling	*Primary objective*: to evaluate CA-125 in serum and peritoneal fluid (PF) as an indicator of endometriosis *Participants*: patients admitted for diagnostic or therapeutic laparoscopy *Selection criteria*: exclusion criteria: any conditions known as influencing CA-125 concentration and with ovarian malignancy established by intraoperative histopathological examination; luteal phase of the cycle *Study design*: cross-sectional single-gate, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: adnexal mass, infertility, pelvic pain, suspected endometriosis *Age*: reproductive age *Number of participants enrolled*: 59 women *Number of participants available for analysis*: 59 women (all in follicular cycle phase) *Setting*: Department of O&G, Clinic of Gynaecological Surgery and Oncology, Medical University of Lodz *Place of study*: Lodz, Poland *Period of study*: not provided *Language*: English
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum CA-125 levels were measured in accordance with the manufacturer's instructions (VIDAS CA-125 II).; sample handling described *Threshold for positive result*: CA-125 > 11 U/ml - not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 44/59 (75%): stage I-II 22, stage III-IV 22; controls n = 15 *Reference standard*: laparoscopy N = 59 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection, in some cases peritoneal biopsy or ovarian cyst excision was conducted; staging according to the ASRM classification *Examiners*: no information provided

Szubert 2012 *(Continued)*

Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	CA-125 cut-off value in serum suggesting endometriosis with 68% sensitivity is 11 U/ml. This value is normal range for CA-125 concentration
Conflict of interest	Not reported; the study was supported by grant no. 2431/B/P01/2009/37 from Polish Ministry of Science and Higher Education
Notes	The diagnostic estimates for the subgroups by severity of endometriosis are not included in the review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowl-	Yes		

Szubert 2012 *(Continued)*

edge of the results of the index tests?			
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Szubert 2014

Study characteristics	
Patient sampling	*Primary objective*: to evaluate these two processes in women with endometriosis who had been treated with danazol to determine the sensitivity of a non-invasive test in diagnosing endometriosis *Participants*: patients admitted for diagnostic or therapeutic laparoscopy for infertility, pelvic pain or both *Selection criteria*: not reported *Study design*: cross-sectional single-gate, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: infertility, pelvic pain; 9 patients (13%) with endometriosis did not report any pain; none had any disorders in the pelvis minor that may have increased the concentrations of the markers under investigation (e.g. ovarian cysts, ovarian tumours or myomas) *Age*: mean age 31.76 ± 5.09 years (median 31 years; range 22–47 years) *Number of participants enrolled*: 103 women *Number of participants available for analysis*: 102-84 women; number of the samples varied for different tests (all in follicular cycle phase) *Setting*: Department of O&G, Clinic of Gynaecological Surgery and Oncology, Medical University of Lodz *Place of study*: Lodz, Poland *Period of study*: February-November 2010 *Language*: English
Index tests	*Index test*: CA-125, VEGF, IL-1β, CRP *Details of the index test procedure as stated*: serum CA-125 levels were measured in accordance with the manufacturer's instructions (VIDAS CA-125 II); plasma CRP concentrations were determined using an immunoturbidimetric assay (PROTILINE kit; bioMérieux, Poland), CA-125 was assessed by enzyme immunofluorescence (VIDAS II automatic quantitative test; bioMérieux, France); VEGF and IL-1β were analysed by ELISA (the QUANTIKINE Human immunoassays; R&D Systems, MN, USA); sample handling and laboratory methods described

Szubert 2014 *(Continued)*

	Threshold for positive result: CA-125 > 11 U/ml - not pre-specified *Examiners*: no information provided; biomarkers were evaluated before laparoscopy *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: peritoneal endometriosis *Prevalence of target condition in the sample*: n = 71/103 (69%): stages I-IV, number per subgroups of severity not reported; controls n = 32 *Reference standard*: laparoscopy N = 103 (100%) *Description of positive case definition by reference standard test as reported*: staging according to the ASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: "Blood samples were collected prior to surgery and evaluated before laparoscopy", time frame not reported, but the context suggests short time before surgery *Withdrawals*: some samples were missing from the analysis (n = 1 for CA-125 and CRP; n = 19 for VEGF, n = 18 for IL-1β) - reason not explained
Comparative	
Key conclusions by the authors	For the diagnosis of endometriosis, none of the combinations of given markers had a sensitivity > 60%. Danazol treatment is highly effective in relieving pain and decreasing CA-125 concentrations in the plasma. Higher plasma concentrations of VEGF after treatment could imply stimulation of angiogenesis
Conflict of interest	Not reported; the study was supported by grant no. 2431/B/P01/2009/37 from the Polish Ministry of Science and Higher Education, a grant from European Funds for Foundation for Polish Science and a doctoral grant from Polfarma Scientific Foundation
Notes	For CRP and IL-1β there was no statistically significant difference between the groups - no data available for meta-analysis For CA-125 and VEGF there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid and endometrium are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		

Szubert 2014 *(Continued)*

		Unclear	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Thubert 2014

Study characteristics

Thubert 2014 *(Continued)*

Patient sampling	*Primary objective*: to evaluate CA-125 in serum and peritoneal fluid (PF) as an indicator of endometriosis *Participants*: non-pregnant patients < 42 years old who underwent pelvic surgery *Selection criteria*: exclusion criteria: visually diagnosed with endometriosis the absence of histologic confirmation *Study design*: cross-sectional two-gate, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: dysmenorrhoea, dyspareunia, non-cyclic pelvic pain, gastro and urinary complains; 224 women from endometriosis group had previous surgery for endometriosis; controls underwent surgeries for various reasons (ovarian cysts, n = 117; tubal defects, n = 81; fibroids, n = 172; and other benign conditions, n = 94); no infectious or inflammatory diseases at the time of serum collection *Age*: mean age 31.9 ± 5.3 years (endometriosis group), 32.2 ± 5.8 years (controls) *Number of participants enrolled*: 1439 women *Number of participants available for analysis*: 834 women (215 in follicular and 207 in luteal cycle phase; in 412 cycle phase was unclear) *Setting*: Department of O&G and Reproductive Medicine, Centre Hospitalier Universitaire Cochin, a tertiary care university hospital *Place of study*: Paris, France *Period of study*: January 2005 - December 2009 *Language*: English
Index tests	*Index test*: hs-CRP *Details of the index test procedure as stated*: CRP levels were assayed in fresh serum using the hs-CRP method, performed on a Cobas Integra 400 Plus analyser using a particle-enhanced immunoturbidimetric technique (Roche Diagnostics, Germany); the lower detection limit of assay was 0.03 mg/L with functional sensitivity of 0.11 mg/l; sample handling described *Threshold for positive result*: > 10 ng/ml, not pre-specified *Examiners*: all measurements were performed in the same laboratory (Laboratoire Port Royal, Paris) preoperatively *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 370/834 (44%): stage I-II - 130, stage III-IV - 240; controls n = 464 *Reference standard*: surgery (not specified) n = 834 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology in all cases, in some cases peritoneal biopsy or ovarian cyst excision was conducted; histological criteria for different types of endometriosis described; staging according to the ASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were acquired and analysed before surgical intervention *Withdrawals*: 605 women were excluded: 133 refused to participate, 365 - missing serum samples, 21 - incomplete surgical excision of endometriosis, 86 - no histologic proof of endometriosis

Thubert 2014 *(Continued)*

Comparative	
Key conclusions by the authors	Although endometriosis is an inflammatory disease, we failed to identify any systemic changes in hs-CRP serum levels; therefore, hs-CRP analysis appears to be irrelevant to the diagnosis and staging of endometriosis
Conflict of interest	Not reported
Notes	The diagnostic estimates for the subgroups by severity of endometriosis are not included in the review The data for different anatomical distributions of endometriosis are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of	Yes		

Thubert 2014 *(Continued)*

the results of the index tests?			
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
		High	

Tokmak 2011

Study characteristics	
Patient sampling	*Primary objective*: to establish the value of a new molecule, urocortin, in the diagnosis of endometrioma and compare with CA-125 to identify superiority of urocortin *Participants*: patients who underwent laparoscopy for adnexal mass in the authors' institution *Selection criteria*: not specified (only moderate-severe endometriosis included) *Study design*: cross-sectional single-gate, multicentre, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation*: adnexal mass, infertility - 28/88, concurrent diseases - 30/88 *Age*: mean age 34.3 ± 7.7 years (endometrioma group), 33.2 ± 11.8 years (controls) *Number of participants enrolled*: 88 women *Number of participants available for analysis*: 88 women (all in follicular cycle phase) *Setting*: Department of Reproductive Endocrinology, Zekai Tahir Burak Women's Health Research and Education Hospital *Place of study*: Ankara, Turkey *Period of study*: January 2009 - June 2009 *Language*: English
Index tests	*Index test*: urocortin, CA-125 *Details of the index test procedure as stated*: plasma urocortin levels measured by using urocortin (Human) EIA kit (range 0–100 ng/dl), Phoenix Pharmaceuticals Inc, Burlingame, CA, USA); serum CA-125 levels were measured with the electro chemiluminescence immunoassay method (Roche Elecsys 1010/2010,Roche Diagnostics, Germany); sample handling described

Tokmak 2011 *(Continued)*

	Threshold for positive result: urocortin > 4.16 ng/dl; CA-125 > 21.38 U/l - not pre-specified *Examiners*: biomarkers were analysed at the Hospital Biochemistry Laboratory; no other information provided *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 42/88 (48%): all stage III-IV; controls n = 46 *Reference standard*: laparoscopy n = 88 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection: the diameter of all the ovarian cysts was measured and peritoneal invasion; staging according to the AFS classification; for pathological examination specimens were obtained by total cystectomy, partial cystectomy or biopsy *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected in the morning before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Urocortin was not found to be efficient in distinguishing endometrioma from other benign ovarian cysts or to be superior to CA-125 in the diagnosis of endometrioma
Conflict of interest	The authors have no conflicts of interest
Notes	—

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**High**
DOMAIN 2: Index Test All tests			
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		

Tokmak 2011 *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Tuten 2014a

Study characteristics	
Patient sampling	*Primary objective*: to determine whether serum copeptin levels were altered in women with endometriosis and played a role in the pathophysiology of the disease *Participants*: women who had undergone laparoscopy or laparotomy due to suspected ovarian endometriosis, infertility and pelvic pain *Selection criteria*: inclusion criteria: reproductive age, regular menstrual cycle; exclusion criteria: postmenopausal FSH levels, pregnancy, suspicion of a malignant ovarian disease, history of any hormone therapy in past 3/12 months, presence of any non-endometriotic ovarian cyst/mass *Study design*: cross-sectional single-gate, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: adnexal mass, infertility, pelvic pain; none had a history of a previous ovarian surgery and any other endocrine or autoimmune disease *Age*: mean age 31.9 ± 8.2 years (endometrioma group), 30.7 ± 7.8 years (controls) *Number of participants enrolled*: 92 women *Number of participants available for analysis*: 86 women (menstrual cycle phase not reported)

Tuten 2014a *(Continued)*

	Setting: Department of O&G, Istanbul University Cerrahpasa School of Medicine *Place of study*: Istanbul, Turkey *Period of study*: May 2012 - July 2013 *Language*: English
Index tests	*Index test*: Copeptin, CRP, WBC, CA-125, CA-19.9, CA-15.3 *Details of the index test procedure as stated*: serum CA-125, CA-19.9, CA-15.3 were measured using an IMMULITE 2000 (DPC, Los Angeles, CA): chemiluminescent immunometric assay for CA-125 and CA-15-3 and immunometric assay for CA-19.9; serum copeptin was measured by using Human Vasopressin-neurophysin 2-copeptin ELISA kit (EIAab Wuhan EIAab Science Co. Ltd, China); with minimum detectable dose of Human Vasopressin-neurophysin 2-copeptin was < 10 pg/ml, detection rate of 15.6–1000.0 pg/ml; CRP was measured using an automated CRPLX Tina-quant C-Reactive Protein (Latex) assay (Roche, Belgium) with lower detection limit of 0.425 mg/L and the functional sensitivity of 0.88 mg/L; sample handling described *Threshold for positive result*: CA-125 > 26.29 IU/ml, CA-19.9 >10.67 IU/ml, CA-15-5 >15.04 IU/ml; copeptin >251.18 pg/ml - not pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 50/86 (58%): stage I-II - 27, stage III-IV - 23; controls n = 36 *Reference standard*: laparoscopy/laparotomy n = 88 (100%) + histology *Description of positive case definition by reference standard test as reported*: thorough examination of the abdominopelvic cavity with histological confirmation; staging according to the ASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before surgery *Withdrawals*: 5 patients were excluded (met exclusion criteria)
Comparative	
Key conclusions by the authors	Serum copeptin levels were significantly higher in patients with endometriosis as compared to healthy controls and severity of the disease was correlated with serum copeptin levels
Conflict of interest	The authors disclosed no conflict of interests
Notes	For CRP and WBC there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		

Tuten 2014a *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Vercellini 1993

Study characteristics	
Patient sampling	*Primary objective*: to determine whether in vivo levels of tumour necrosis factor a in plasma and peritoneal fluid differ in infertile subjects with and without endometriosis *Participants*: women undergoing laparoscopy for infertility *Selection criteria*: inclusion criteria: regular menstrual cycles, no previous pelvic surgery, no hormonal treatment in preceding 3 months *Study design*: cross-sectional, single-gate design, prospective recruitment and collection of samples, consecutive patients
Patient characteristics and setting	*Clinical presentation*: primary infertility - 70/94, secondary infertility - 24/94 *Age*: mean age 30 ± 6 years (endometriosis group), 29 ± 5 years (controls) *Number of participants enrolled*: 94 women *Number of participants available for analysis*: 94 women (cycle phase not specified) *Setting*: Department of O&G, University of Milano *Place of study*: Milan, Italy *Period of study*: not provided *Language*: English
Index tests	*Index test*: TNF-α *Details of the index test procedure as stated*: plasma levels of TNF-α were assessed by using enzyme immunoassay test (Biokine, T Cell Sciences, Mas, USA); sensitivity 10 pg/ml; sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; blinded to the result of reference standard *Interobserver variability*: intra-assay CV < 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 46/94 (49%): stage I-II 38, stage III-IV 8; controls n = 48 *Reference standard*: laparoscopy n = 94 (100%) *Description of positive case definition by reference standard as reported*: staging according to the rAFS score *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before anaesthesia *Withdrawals*: none
Comparative	

Vercellini 1993 *(Continued)*

Key conclusions by the authors	In our series, plasma and peritoneal tumour necrosis factor a levels were not different in infertile women with and without endometriosis.
Conflict of interest	Not reported; supported by Italian National Research Council, grant N 91.00131.PF41.115.05532
Notes	For plasma TNF-α there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			

Vercellini 1993 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Verit 2008

Study characteristics	
Patient sampling	*Primary objective*: to compare the serum PON-1 activity in women with endometriosis versus controls and to assess whether PON-1 activity can be used as a diagnostic test for endometriosis *Participants*: women undergoing laparoscopy or laparotomy for evaluation of infertility, pelvic pain, pelvic mass, tubal ligation or endometriosis *Selection criteria*: inclusion criteria: reproductive age, regular menstrual cycle; exclusion criteria: age > 35 years, pregnancy, hormonal therapy, smoking, alcohol drinking, CAD, unstable angina, myocardial infarction, any operation or cardiovascular intervention within the previous 3 months, hypertension, hyperlipidaemia, rheumatological or endocrine conditions, liver diseases, renal dysfunction, anaemia, obesity, parasitic diseases, systemic or local infection, any history of cancer in the past 5 years and therapeutic interventions known to influence antioxidants such as supplemental vitamins *Study design*: cross-sectional two-gate, prospective recruitment and collection of samples, consecutive series
Patient characteristics and setting	*Clinical presentation*: preoperative indications: infertility 50 (57.5%), pelvic pain 9 (10.3%), pelvic mass 16 (18.4%), tubal ligation 12 (13.8%) *Age*: mean age 24.4 ± 4.0 years (endometriosis group), 24.8 ± 3.8 years (controls) *Number of participants enrolled*: 87 women *Number of participants available for analysis*: 87 women (all in follicular cycle phase) *Setting*: tertiary referral centre - Department of O&G, Harran University Faculty of Medicine *Place of study*: Sanliurfa, Turkey *Period of study*: November 2006 - May 2007 *Language*: English
Index tests	*Index test*: PON-1 *Details of the index test procedure as stated*: PON-1 enzymatic activity determined by using paraoxon as a substrate and measured by increases in the absorbance at 412 nm due to the formation of 4-nitrophenol (referenced to the primary source); sample handling and laboratory technique described *Threshold for positive result*: < 141.5 U/l, not pre-specified (different thresholds for diagnosis of minimal-mild and moderate-severe disease) *Examiners*: no information provided; unclear if were blinded to the result of reference standard

Verit 2008 *(Continued)*

	Interobserver variability: Intra-and interassay CV 3%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 47/87 (54%): stage I-II - 24, stage III-IV - 23; controls n = 40 *Reference standard*: laparoscopy n = 71 (81.6%)/laparotomy 16 (18.4%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rAFS classification *Examiners*: all procedures were performed by the same surgeon in a tertiary referral centre
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected less than 12 months before surgery (personal communication with the author) *Withdrawals*: none
Comparative	
Key conclusions by the authors	Reduced serum PON-1 activity and increased LOOH might contribute to the increased susceptibility for the development of atherosclerosis. PON-1 activity can be used as a diagnostic test to detect endometriosis
Conflict of interest	Not reported; no financial support was accepted for this study
Notes	The reported diagnostic estimates per severity of endometriosis are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	No		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpreta-	Yes		

Verit 2008 *(Continued)*

tion of the result of index test?			
		High	**Low**
DOMAIN 3: Reference Standard			
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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Vigano 2002

Study characteristics	
Patient sampling	*Primary objective*: to evaluate whether leptin may be used as a new serum marker of endometriosis *Participants*: women who underwent laparoscopy for infertility, pelvic pain or adnexal mass *Selection criteria*: inclusion criteria: reproductive age (17–46 years), normal regular menstrual cycle (25–35 d), day 5 LH/FSH <2, no hormone therapy for at least 3 months before surgery, no evidence of endometritis or previous autoimmune, liver, endocrine or malignant disease; exclusion criterion: intraoperative diagnosis of malignancy *Study design*: cross-sectional, single-gate design, prospective collection of samples, consecutive patients
Patient characteristics and setting	*Clinical presentation*: infertility, pelvic pain, adnexal mass *Age*: median age 32.2 years, range 23–46 years (endometriosis), 33 years, range 17-40 years (controls)

Vigano 2002 *(Continued)*

	Number of participants enrolled: 67 women *Number of participants available for analysis*: 67 women (8 in menstrual, 28 in follicular, 31 in luteal cycle phase) *Setting*: II Department of O&G, University of Milan *Place of study*: Milan, Italy *Period of study*: February 2000 - October 2000 *Language*: English
Index tests	*Index test*: leptin *Details of the index test procedure as stated*: serum levels of leptin were assessed by using using a commercially available RIA kit (DRG Instruments GmbH, Germany) with a sensitivity of 0.5 ng/ml; sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: Intra- and interassay CV 3.4%–8.3% and 6.5%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 42/67 (63%): stage I-II 20, stage III-IV 22; controls n = 25 *Reference standard*: laparoscopy N = 67 (100%) + histology *Description of positive case definition by reference standard as reported*: visualisation and histological confirmation in all cases of atypical, deep and adnexal lesions; classification according to the rASRM score *Examiners*: 3 physicians active in the evaluation and treatment of endometriosis
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Serum concentrations of the obese gene product, leptin, cannot reliably be used for the diagnosis of endometriosis
Conflict of interest	Not reported; supported by the EndoBank program of Arevia GmbH
Notes	For leptin there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		

Vigano 2002 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Low	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Vigil 1999

Study characteristics

Vigil 1999 *(Continued)*

Patient sampling	*Primary objective*: to study a prevalence of endometriosis in women of reproductive age presenting with dysmenorrhoea, infertility or both; to evaluate relationship between CA-125 and laparoscopic finding and to identify the most frequent grade endometriosis by age group *Participants*: women who underwent laparoscopy for dysmenorrhoea and pelvic pain not responding to medical management, with or without infertility *Selection criteria*: not specified *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*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 *Language*: Spanish
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: CA-125 levels analysed by the IRMA-COUNT OM-MA method; sample handling and laboratory technique not described *Threshold for positive result*: > 35 IU/ml, pre-specified *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 45/49 (92%): stages I-IV, number of patients per group provided only for stage IV - 20; controls n = 4 *Reference standard*: laparoscopy N = 49 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: not specified, but the context suggests that the samples were taken peri-operatively *Withdrawals*: 1 patient excluded from the analysis (reason not specified)
Comparative	
Key conclusions by the authors	CA-125 is not correlated with the presence and staging of endometriosis. Laparoscopy remains the best alternative
Conflict of interest	Not reported

Vigil 1999 *(Continued)*

Notes	Translated from Spanish The reported diagnostic estimates per age group (< 25 years and 26-41 years) are not reported in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Unclear**
DOMAIN 3: Reference Standard			
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		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		

Vigil 1999 *(Continued)*

Were all patients included in the analysis?	Yes
	Unclear

Vodolazkaia 2011

Study characteristics	
Patient sampling	*Primary objective*: to compare the diagnostic performance of the hs-CRP assay and the classical CRP assay to detect low grade inflammation in plasma of women with endometriosis *Participants*: women who underwent laparoscopy for subfertility with or without pain - identified through electronic database of the biobank samples *Selection criteria*: exclusion criteria: samples collected from women who were on hormonal medication at the time of collection, who had been operated within 6 months prior to the time of collection or who had other pelvic inflammatory disease or general diseases at the time of collection *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility or both *Age*: reproductive age *Number of participants enrolled*: 295 women *Number of participants available for analysis*: 295 women (60 in menstrual, 119 in follicular and 116 in luteal cycle phase) *Setting*: Department of O&G, University Hospital Gasthuisberg *Place of study*: Leuven, Belgium *Period of study*: not specified; samples collected since 1999 *Language*: English
Index tests	*Index test*: CRP and hsCRP *Details of the index test procedure as stated*: plasma CRP level was measured twice by 2 methods: the classical automated CRPLX Tina-quant assay (Roche, Vilvoorde, Belgium) (CRP), and HS Tina-quant high sensitive assay (Roche, Vilvoorde, Belgium) (hsCRP), both performed on a Roche Modular P instrument; the lower detection limit was 0.425 mg/L (CRP) and 0.03 mg/L (hsCRP); sample handling and method described *Threshold for positive result*: CRP > 0.71 mg/l; hs-CRP > 0.62 mg/l all phases, > 0.70 ng/ml for luteal phase, > 0.61 for follicular phase, > 0.73 for menstrual phase; not pre-specified *Examiners*: the assays were performed at the central laboratories of the University Hospitals Leuven *Interobserver variability*: the within-run CV was 1.34%-0.28% for hs-CRP and 2.5%-0.76% for CRP; total imprecision CV was 5.70%-2.51% for hsCRP and 2.53%-1.8% for CRP
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 204/295 (69%): stage I-II 135, stage III-IV 69; controls n = 91 *Reference standard*: laparoscopy N = 295 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection with histological confirmation for most of the samples; staging according to the rASRM classification

Vodolazkaia 2011 *(Continued)*

	Examiners: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples collected before anaesthesia *Withdrawals*: none
Comparative	
Key conclusions by the authors	The hsCRP assay was superior to the classical CRP assay for the detection of low CRP levels and for revealing subclinical inflammation in plasma of women with endometriosis
Conflict of interest	The authors declare that they have no competing interests; supported by a TBM (Toegepast Biomedisch Onderzoek met Primair Maatschappelijke Finaliteit) grant from the Institute for Innovative Science and Technology IWT (Innovatie door Wetenschap en technologie) in Flanders, Belgium
Notes	The reported diagnostic estimates according to severity of endometriosis are not presented in this review The reported diagnostic estimates for CRP assay are demonstrated as inferior to hs-CRP, since both assays test the same marker - less accurate classical CRP is not presented in this review The diagnostic estimates for hs-CRP were reported for the overall group and per menstrual cycle phase The diagnostic estimates for hs-CRP in luteal cycle phase were also reported for the same cohort in Mihalyi 2010 but the cut-off threshold in the later study was not provided, therefore the data from both studies are included but not combined in the meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
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		

Vodolazkaia 2011 *(Continued)*

Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Vodolazkaia 2012

Study characteristics	
Patient sampling	*Primary objective*: to develop and validate a non-invasive diagnostic test with a high sensitivity (80% or more) for symptomatic endometriosis patients, without ultrasound evidence of endometriosis, since this is the group most in need of a non-invasive test *Participants*: women who underwent laparoscopy for subfertility with or without pain - identified through electronic database of the biobank samples *Selection criteria*: inclusion criteria: minimal sample volume (2.5 ml) and essential clinical information available; exclusion criteria: samples collected from women who were on hormonal medication, had other pelvic inflammatory disease or general diseases at the time of collection or who had been operated within 6 months prior to collection *Study design*: cross-sectional single-gate, prospective collection of samples

Vodolazkaia 2012 *(Continued)*

Patient characteristics and setting	*Clinical presentation*: pelvic pain, infertility or both *Age*: mean age 31.2 ± 4.02 years, range 24-44 years (endometriosis), 31.7 ± 5.28 years, range 19-46 years (controls) *Number of participants enrolled*: 353 women - independent training and test set, with equal distribution of controls (34%) and endometriosis (66%) patients *Number of participants available for analysis*: 296 women (67 in menstrual, 111 in follicular and 118 in luteal cycle phase; all had normal preoperative ultrasound) *Setting*: Department of O&G, University Hospital Gasthuisberg *Place of study*: Leuven, Belgium *Period of study*: not specified; samples collected since 1999 *Language*: English
Index tests	*Index test*: Panel of 28 bio markers: IL-1β, IL-4, IL-6, IL-8, IL-10, IL-17, TNF-α, RANTES, NGF, β-FGF, IFN-γ, MIF, MCP-1, VCAM, VEGF, M-CSF, HGF, osteopontin, IGFBP-3, leptin, sICAM-1, follistatin, annexin V, IL-21, glycodelin, CA-125, CA-19.9, hs-CRP *Details of the index test procedure as stated*: plasma levels of the biomarkers were assessed by using Bio-Plex Protein Array System (Bio-Rad Laboratories, USA) for IL-1β, IL-4, IL-6, IL-8, IL-10, IL-17, TNF-α, RANTES, NGF, β-FGF, IFN-γ, MIF; multiplexing sandwich-ELISA (Aushon Biosystems, USA) for osteopontin, IGFBP-3, leptin; single ELISAs for sICAM-1 and follistatin (R&D Systems, USA), annexin V (American Diagnostica, Inc, USA), IL-21 (Bender Med Systems, Austria) and glycodelin (Bioserv Diagnostics, Germany); automated immunoassays (Roche, Vilvoorde, Belgium) for CA-125, CA-19.9 and hs-CRP; the analyses were performed separately for training and for test sets using univariate analysis for individual markers as well as the multivariate logistic regression and LS-SVM models for predictive models of the combined biomarkers *Threshold for positive result*: CA-125 > 12.5 U/ml, glycodelin > 18 ng/ml, VEGF > 1.5 pg/ml, IGFBP-3 > 210 ng/ml, sICAM-1 < 243 ng/ml (all phases) and < 254.6 ng/ml (menstrual), CA-19.9 > 9.5 IU/ml CRP > 0.71 mg/l; hs-CRP > 0.62 mg/l all phases, > 0.70 ng/ml for luteal phase, > 0.61 for follicular phase, > 0.73 for menstrual phase; prespecified (for validation test set) *Examiners*: the assays were performed at the central laboratories of the University Hospitals Leuven *Interobserver variability*: glycodelin: intra- and interassay CV 12.6%-15.3% and 6.8%-18.8%, not reported for other tests
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 175/296 (59%): stage I-II 146, stage III-IV 29; controls n = 121 *Reference standard*: laparoscopy N = 296 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: visual inspection with histological confirmation for most of the samples; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples collected before anaesthesia *Withdrawals*: 57 women were excluded prior to analysis as had endometriosis-related findings on preoperative ultrasound (outside the study objectives)
Comparative	
Key conclusions by the authors	The hs-CRP assay was superior to the classical CRP assay for the detection of low CRP levels and for revealing subclinical inflammation in plasma of women with endometriosis

Vodolazkaia 2012 *(Continued)*

Conflict of interest	The authors declare that they have no competing interests; supported by a TBM grant from the Institute for Innovative Science and Technology in Flanders, Belgium, Research Council KUL: ProMeta, GOA MaNet, CoE EF/05/007 SymBioSys, GOA 08/16 KUL PFV/10/016 SymBioSys, START 1, several PhD/postdoc and fellow grants Flemish Government (the list is not presented in full)
Notes	The reported diagnostic estimates according to severity of endometriosis are not presented in this review The reported diagnostic estimates for each marker are presented for only for ultrasound-negative endometriosis (univariate analysis for single markers and multivariate analysis/LS-SVM model for combination of the biomarkers) The diagnostic estimates are presented only for validation test set except for CA-19.9 (only data for training set available) The diagnostic estimates are presented for the overall group, per specific menstrual cycle phase or both only for the best performing markers, as selected for reporting by the authors The diagnostic estimates for CA-125 for each cycle phase were also reported in the overlapping but not identical cohort in Mihalyi 2010, but the cut-off threshold in that study was not provided, therefore the data from both studies are included but not combined in a meta-analysis IL-4, NGF and M-CSF were not detectable in 90% of the samples and have been excluded from the statistical analysis.

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it prespecified?	Yes		
Was a cycle phase considered in inter-	Yes		

Vodolazkaia 2012 *(Continued)*

pretation of the result of index test?			
		Unclear	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Vouk 2012

Study characteristics	
Patient sampling	*Primary objective*: to analyse the plasma metabolomes of endometriosis patients by comparing them with healthy controls *Participants*: patients with ovarian endometriosis who underwent laparoscopic surgery and a control group of healthy women who underwent sterilisation *Selection criteria*: not specified *Study design*: cross-sectional two-gate, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified; concomitant findings: adenomyosis - 1; fibroids - 5 in endometriosis group; fibroids - 3 in controls; not on hormonal treatment - 75% endometriosis group, 62% controls

Vouk 2012 *(Continued)*

	Age: mean age 33.3 ± 6.06 years, range 22-44 years (endometriosis group), 40.6 ± 3.1 years, range 32-45 years (controls) *Number of participants enrolled*: 111 women *Number of participants available for analysis*: 92 women (29 in follicular, 19 in late follicular/early luteal, 41 in luteal cycle phase; for 3 participants cycle phase was not determined) *Setting*: Department of O&G, University Clinical Centre, University of Ljubljana *Place of study*: Ljubljana, Slovenia *Period of study*: March 2008 - October 2009 *Language*: English
Index tests	*Index test*: metabolome (model of SMOH C16:1+ PCaa C36:2/ PCae C34:2, corrected for age and BMI) *Details of the index test procedure as stated*: plasma metabolome evaluated by electrospray ionisation mass spectrometry (ESI-MS/MS) measurements with the AbsoluteIDQTM p150 kit (BIOCRATES Life Sciences AG, Austria); referenced to the sources with description of the assay and quality measures; experiments sample handling described; diagnostic model defined by using backward stepwise-regression selection procedure *Threshold for positive result*: not provided *Examiners*: no information provided; blinded to the result of reference standard; "randomly assigned samples" *Interobserver variability*: CV < 0.25 (otherwise excluded)
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 40/92 (44%): all stage III-IV; controls n = 52 *Reference standard*: laparoscopy N = 92 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected on the day of surgery *Withdrawals*: 19 patients were excluded prior to analysis for the following reasons: the absence of ovarian endometriosis (11 patients), pregnancy (1 control), menopause (1 patient), surgery did not take place (2 controls) and errors in the sampling procedure (2 patients and 2 controls)
Comparative	
Key conclusions by the authors	Endometriosis is associated with elevated levels of sphingomyelins and phosphatidylcholines, which might contribute to the suppression of apoptosis and affect lipid-associated signalling pathways
Conflict of interest	The authors have nothing to disclose; supported by a J3-4135 grant from the Slovenian Research Agency and by the Deutsche Forschungsgemeinschaft grant AD127/10-1
Notes	The evaluated diagnostic model was selected by using multiple regression procedure

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Vouk 2012 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Wang 2013a

Study characteristics	
Patient sampling	*Primary objective*: to detect the serum microRNAs that are differentially expressed between endometriosis patients and negative controls to evaluate the potential of these microRNAs as diagnostic markers for endometriosis *Participants*: patients attending the hospital with complaints of severe dysmenorrhoea and pelvic mass as well as infertility and subsequently underwent laparoscopy *Selection criteria*: inclusion criteria: aged 20–60 years, no hormone therapy for at least 3/12 months, non-smoker, no history of other inflammatory disease; exclusion criteria: malignancy, benign ovarian cyst except endometrioma, severe PID, known chronic, systemic, metabolic, and endocrine disease including PCOS *Study design*: cross-sectional single-gate, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility - 48/85, dysmenorrhoea - 44/85 *Age*: mean age 33.3 ± 6.06 years, range 22-44 years (endometriosis), 40.6 ± 3.1 years, range 32-45 years (controls) *Number of participants enrolled*: 85 women *Number of participants available for analysis*: 85 women (64 in follicular, 21 in luteal cycle phase) *Setting*: Department of O&G, Sun Yat-sen Memorial Hospital, Sun Yat-Sen University *Place of study*: Guangzhou, China *Period of study*: 2011 *Language*: English
Index tests	*Index test*: miRNAome with subsequent validation (miR-199a, miR-122, miR-145*, miR-141*, miR-542-3p, miR-9*) *Details of the index test procedure as stated*: plasma microRNA expression evaluated by RT-PCR (screening with Taq- Man microRNA array in pooled samples followed by validation with single assays (SYBR Premix Ex Taq II-based (Takara, Japan) quantified with Roche Light Cycler 480II (Roche, Switzerland)); experiments run in triplicates, normalised to U6; sample handling described; discriminant analysis was used to built the diagnostic model *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 60/85 (71%): stage I-II stage 22, III-IV 38, peritoneal endometriosis - 19, ovarian endometriosis - 41, DIE - 18; controls n = 25 *Reference standard*: laparoscopy N = 85 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to rAFS classification *Examiners*: no information provided

Wang 2013a *(Continued)*

Flow and timing	*Time interval between index test and reference standard*: blood samples were collected 1–3 d before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The circulating microRNAs miR-199a, miR-122, miR-145*, and miR-542-3p could potentially serve as non-invasive biomarkers for endometriosis. miR-199a may also play an important role in the progression of the disease
Conflict of interest	The authors have nothing to declare; supported by the funds from National Science and Technology Department (973, 2011CB811301) and National Science Foundation of China (81270629 and 30500578)
Notes	The predictive models based on combination of microRNAs were defined by discriminant analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted	Yes		

Wang 2013a *(Continued)*

without knowledge of the results of the index tests?			
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Webster 2013

Study characteristics	
Patient sampling	*Primary objective*: to investigate the relationship between circulating angiogenic cells (CACs) and the presence of endometriosis in women, so as to determine whether CACs could be used as a disease biomarker *Participants*: women scheduled for laparoscopy for symptoms or signs suggestive of endometriosis *Selection criteria*: not specified *Study design*: cross-sectional single-gate, prospective recruitment and collection of samples
Patient characteristics and setting	*Clinical presentation*: chronic pelvic pain - 44, subfertility - 36, ovarian cysts - 15; all women were free of exogenous hormones in the preceding 3/12 months *Age*: mean age 35.6 ± 5.0 years (endometriosis group), 32.9 ± 7.3 years (controls) *Number of participants enrolled*: 64 women *Number of participants available for analysis*: 64 women (9 in menstrual, 21 in follicular, 8 in peri-ovulatory, 22 in luteal cycle phase; for 4 participants cycle phase was not determined) *Setting*: Department of O&G, University of Oxford, Women's Centre, John Radcliffe Hospital, a national referral centre for the management of endometriosis *Place of study*: Oxford, UK *Period of study*: July 2010 - May 2012 *Language*: English
Index tests	*Index test*: CAC *Details of the index test procedure as stated*: Peripheral blood CAC was evaluated by flow cytometry according to an established protocol for identifying viable $CD34^{bright}CD133+CD31+ CD45^{dim}$ cells; in a subgroup of women, CAC levels were also assessed using a CFU assay; laboratory methods referenced to a primary source and described

Webster 2013 *(Continued)*

	Threshold for positive result: not provided *Examiners*: no information provided, unclear if blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 42/64 (66%): stage I-II - 21, stage III-IV - 21; controls n = 22 *Reference standard*: laparoscopy n = 64 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rASRM classification *Examiners*: surgeon was blinded to laboratory results
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected on the day of surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	CACs are not a useful biomarker of endometriosis and may be unaffected by the presence of this disease
Conflict of interest	The authors declared no conflict of interests; supported grants from the MRC (New Investigator Award, G0601458), the Oxford Partnership Comprehensive Biomedical Research Centre with funding from the Department of Health's NIHR Biomedical Research Centres Scheme and the Oxfordshire Health Services Research Committee
Notes	—

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		

Webster 2013 *(Continued)*

If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Wei 2005

Study characteristics	
Patient sampling	*Primary objective*: to determine serum and peritoneal fluid leptin levels in women with infertility due to endometriosis *Participants*: women with with infertility or benign ovarian cysts who underwent laparoscopy *Selection criteria*: exclusion criteria: steroid treatment or immunosuppressant treatment 3 months prior to surgery, endometritis, autoimmune disease, endocrine disorders, liver disease, cancer, and abnormalities in reproductive system; other causes of infertility *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility or ovarian cyst *Age*: age range 24-35 years (endometriosis group), 20-35 years (controls) *Number of participants enrolled*: 63 participants

Wei 2005 *(Continued)*

	Number of participants available for analysis: 63 participants (cycle phase not reported) *Setting*: Department of O&G, Gyn Xiangya Hospital of Central South University *Place of study*: Changshang, Hunan province, China *Period of study*: April 2004 - August 2004 *Language*: Chinese
Index tests	*Index test*: leptin *Details of the index test procedure as stated*: serum leptin was measured with RIA (Beijing East - Asian Immune Reagent Institute), minimal detection limit was 0.1 pg/ml; sample handling described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n =33/63 (52%): stage I-II - 14, stage III-IV - 19; controls n = 30 (control group 1 - 15, control group 2 - 15) *Reference standard*: laparoscopy N = 63 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected on the day of surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Peritoneal full leptin level was significantly increased in endometriosis infertility patients, suggesting that leptin may affect fertility via a localised mechanism
Conflict of interest	Not reported
Notes	The data for markers measured in peritoneal fluid are not reported in this review For leptin there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		

Wei 2005 *(Continued)*

Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Wild 1991a

Study characteristics	
Patient sampling	*Primary objective*: to determine whether antibody detection by utilising endometrial carcinoma cell line is more sensitive, specific or both than measurement of circulating CA-125 levels *Participants*: patients undergoing laparoscopy or laparotomy for infertility investigation *Selection criteria*: not specified

Wild 1991a *(Continued)*

	Study design: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility *Age*: mean age 30.7 years, range 18-40 years *Number of participants enrolled*: 93 women (36 gynaecology patients and 73 gynaecological oncology patients were presented as separate groups and not included in this review) *Number of participants available for analysis*: 93 women (cycle phase not specified) *Setting*: Hershey Medical Centre, Pennsylvania State University *Place of study*: Hershey, Pennsylvania *Period of study*: not provided *Language*: English
Index tests	*Index test*: IgG anti-endometrial Abs, CA-125 *Details of the index test procedure as stated*: serum anti-endometrial antibodies evaluated by IIF utilising monolayered cultures of carcinoma cell line; fluorescence evaluated by using Nicon optics (Nicon Inc, NY) and ranked by immunofluorescence intensity (0 to 3+); laboratory method described and referenced to a primary source; serum CA-125 levels determined by IRMA (Centocor, PA) *Threshold for positive result*: IgG Abs: positive fluorescence of 1+ to 3+ (ranked by intensity of immunofluorescence); CA-125: > 16 U/ml; pre-specified *Examiners*: single technician, blinded to the surgical findings *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 72/93 (77%): stage I-II stage 51, III-IV 21; controls n = 21 *Reference standard*: laparoscopy/laparotomy N = 93 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	These initial results suggest that detection of antibodies might be useful for the diagnosis of endometriosis
Conflict of interest	Not reported; supported in part by a contract from Winthrop Pharmaceuticals Division of Sterling Drug, NY
Notes	The reported data for gynaecological patients (wide age range) and gynaecological oncology patients are not included in this review

Wild 1991a *(Continued)*

	The presented data enabled calculation of the diagnostic estimated according to severity of endometriosis - not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Wild 1991a *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Wolfler 2009

Study characteristics	
Patient sampling	*Primary objective*: to evaluate evaluate whether distinct patterns of serum proteins in symptomatic women are of value to predict endometriosis before laparoscopy *Participants*: women presenting for diagnosis or treatment of dysmenorrhoea, dyspareunia, chronic pelvic pain or unexplained infertility *Selection criteria*: exclusion criteria: oestrogen-dependent diseases, previous diagnosis of endometriosis or endocrine therapy such as GnRH analogues or danazol *Study design*: cross-sectional single-gate, prospective recruitment
Patient characteristics and setting	*Clinical presentation*: dysmenorrhea - 74/91, dyspareunia - 14/91, chronic pelvic pain - 28/91, infertility - 31/91 *Age*: mean age 32.3, range 22 - 47 years *Number of participants enrolled*: 91 women *Number of participants available for analysis*: 90 (51 proliferative and 39 secretory phase) *Setting*: tertiary care centre, institution not specified *Place of study*: not stated; authors' affiliations include universities in Aachen and Luebeck, Germany and in Peking, China *Period of study*: not stated *Language*: English
Index tests	*Index test*: proteome by SELDI-TOF MS (5 peaks with molecular weights of 4159.00 Da, 5264.00 Da, 5603.00 Da, 9861.00 Da and 10,533.00 Da) *Details of the index test procedure as stated*: serum proteome was profiled by SELDI-TOF MS, by using Q10 (anionic exchange surface) ProteinChips (Ciphergen, Freemont, CA) and the calibrated protein biologic system IIc SELDI-ProteinChipReader, ProteinChip 3.1 software (Ciphergen), and optimised measuring protocol; sample processing, experimental techniques and analyses described in details; classifying model was created with subsequent cross-validation and application of decision tree algorithm to optimise the classification *Threshold for positive result*: presence or absence of the selected mass protein peaks, not pre-specified *Examiners*: no information provided; unclear if were blinded to surgical data *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 51/90 (57%): stage I-II 19, stage III-IV 32; controls n = 39 women *Reference standard*: laparoscopy N = 90 (100%) + histology

Wolfler 2009 *(Continued)*

	Description of positive case definition by reference standard test as reported: laparoscopic visualisation, followed by histopathologic assessment of putative lesions; staging according to the rAFS classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained before laparoscopy *Withdrawals*: 1 sample was excluded as not eligible (not specified)
Comparative	
Key conclusions by the authors	Screening for serum protein patterns using SELDI-TOF MS before laparoscopy might be of discriminative value in the prediction of disease and partly confirms recently published data. However, both low sensitivity and low specificity disqualify this method as a 'quick fix' diagnostic test
Conflict of interest	The authors reported no conflict of interest; supported by a research grant from Takeda Pharma
Notes	The reported diagnostic estimates according to severity of endometriosis are not presented in this review The diagnostic estimates established by a rule-based selection process using a decision tree algorithm (DTA) are reported in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Unclear**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**

Wolfler 2009 *(Continued)*

DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Wu 1998

Study characteristics	
Patient sampling	*Primary objective*: to investigate the association between concentrations of soluble intercellular adhesion molecule- 1 (ICAM-1) and interferon-gamma (IFN-γ) with regard to the severity of endometriosis *Participants*: women with infertility who underwent laparoscopy for suspected endometriosis *Selection criteria*: not specified *Study design*: cross-sectional, single-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility *Age*: mean age 28.93 ± 2.66 years, range, 24–35 years *Number of participants enrolled*: 71 women *Number of participants available for analysis*: 71 women (cycle phase not reported) *Setting*: Department of O&G, Medical College, National Cheng-Kung University *Place of study*: Tainan, Taiwan *Period of study*: not provided

Wu 1998 *(Continued)*

	Language: English
Index tests	*Index test*: INF-γ, ICAM-1 *Details of the index test procedure as stated*: plasma levels of ICAM-1 and serum levels of INF-γ were assessed by using commercial ELISA kits (Cellfree, ICAM-1 test kit; T Cell Diagnostics, MA); quantification at absorbance at 490 nm; kit sensitivity of 0.3 ng/ml; sample processing and laboratory methods described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 36/71 (51%): stage I-II - 22, stage III-IV - 14; controls n = 35 *Reference standard*: laparoscopy n = 71 (100%) *Description of positive case definition by reference standard as reported*: classification according to rASRM score *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The increased serum levels of ICAM-1 found in patients with endometriosis may indicate the presence of an active disease process. Further, the increased levels of soluble ICAM-1 in peripheral blood were inversely correlated with the concentrations of INF-γ in PF and may be associated with an immunologic feedback response that blocks further infiltration of immune cells. These findings may be of value in the diagnosis and evaluation of endometriosis
Conflict of interest	Not reported; supported by by grant NSC86-2314-B006-080 from the National Science Council, Taipei, Taiwan
Notes	For serum INF-γ there was no statistically significant difference between the groups - no data available for meta-analysis For ICAM-1 there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		

Wu 1998 *(Continued)*

Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Yagmur 2013

Study characteristics	
Patient sampling	*Primary objective*: to evaluate whether the analysis of different pro-inflammatory and angiogenesis-regulating cytokines in a well-defined patient population can be accurate for the diagnosis of endometriosis at different stages

Yagmur 2013 *(Continued)*

	Participants: patients undergoing laparoscopy for infertility investigation *Selection criteria*: exclusion criteria: women on hormonal medication, underwent an operation within 6 months, other pelvic inflammatory disease. *Study design*: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility *Age*: mean age 31.24 ± 7.24 years (endometriosis group), 26.86 ± 9.13 years (controls) *Number of participants enrolled*: 55 women *Number of participants available for analysis*: 55 women (cycle phase not specified) *Setting*: Department of O&G, Istanbul University School of Medicine *Place of study*: Istanbul, Turkey *Period of study*: not provided *Language*: English
Index tests	*Index test*: CA-125, IL-6, Epo, TNF-α *Details of the index test procedure as stated*: plasma concentrations of Epo, IL-6 and TNF-α were determined by using commercially available ELISA kits (R&D Systems Inc, Minneapolis, USA) according to the manufacturer's instructions; plasma levels of the CA-125 were measured using Microparticle Enzyme Immunoassay (MEIA) Abbott AxSYM instrument (Abbott Diagnostics, USA) *Threshold for positive result*: not provided *Examiners*: no information provided *Interobserver variability*: Inter- and intra-assay CV were for Epo < 10% and 5.9%; for IL-6, 6.4% and 4.2%; for TNF-α, 3.5% and 1.8%; for CA-125 < 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 33/55 (60%): stage I-II stage 16, III-IV 17; controls n = 22 *Reference standard*: laparoscopy n = 55 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to the rAFS classification *Examiners*: experienced gynaecologic surgeon
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected on the day of surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	Progression of endometriosis is associated with the elevated level of serum IL-6. Undoubtedly, larger well-designed prospective studies are urgently needed to determine the diagnostic potential of cytokines like IL-6 in endometriosis
Conflict of interest	The authors declared no conflict of interests
Notes	For CA-125 and IL-6 levels there was statistically significant difference between the groups but there were insufficient information to construct 2 x 2 tables - not included in this review

Yagmur 2013 *(Continued)*

	For Epo and TNF-α levels, there was no statistically significant difference between the groups - no data available for meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Yagmur 2013 *(Continued)*

Were all patients included in the analysis?	Yes
	Low

Yang 1994

Study characteristics	
Patient sampling	*Primary objective*: to measure levels of CA-125 and endometrial antibodies (EMAb) in serum and peritoneal fluid *Participants*: women who underwent laparoscopy at the authors' institution for infertility or suspected endometriosis *Selection criteria*: not reported *Study design*: cross-sectional, single-gate, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: infertility - 40, suspected endometriosis - 2 *Age*: mean age 31.36 years, range 24-39 years *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 *Language*: Chinese
Index tests	*Index test*: CA-125, anti-endometrial antibodies *Details of the index test procedure as stated*: 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; endometrial antibodies were assessed with indirect ELISA by using the endometrial antigens (EMAg) and horseradish peroxidase-labelled staphylococcal protein A (HRP-SPA); sample handling and laboratory technique described *Threshold for positive result*: CA-125 > 35,000 U/l, for anti-endometrial antibodies > 0.3 A (492 nm wavelength absorbance value), not pre-specified *Examiners*: not information provided, unclear if were blinded to the result of reference standard *Interobserver variability*: Intra- and inter-observer CV for CA-125 were 4.3%-5.4% and 5.3%-6.6%; for anti-endometrial antibodies 7.9%-9.2% and 10.1%-11.7%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 28/42 (67%): stage I-II - 19, stage III-IV - 9; controls n = 14 *Reference standard*: laparoscopy n = 42 (100%) *Description of positive case definition by reference standard test as reported*: staging according to rAFS classification

Yang 1994 *(Continued)*

	Examiners: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	The sensitivity of CA-125 and EMAb measurements in the diagnosis of endometriosis were 71.43% and 82.14%, and the specificity were 57.21% and 57.14% respectively
Conflict of interest	Not reported
Notes	The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		

Yang 1994 *(Continued)*

		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Yavuzcan 2013

Study characteristics	
Patient sampling	*Primary objective*: to compare the preoperative values of mean platelet volume (MPV) and peripheral systemic inflammatory response (SIR) markers (neutrophil/lymphocyte ratio and platelet/lymphocyte ratio) between patients with advanced-stage (stage 3/4) endometriosis having endometrioma and patients with a non-neoplastic adnexal mass other than endometrioma *Participants*: patients who underwent laparotomy or laparoscopy with the pre-diagnosis of infertility or adnexal mass and who underwent laparoscopic tubal ligation *Selection criteria*: exclusion criteria: patients beyond reproductive age, previous medical therapy for endometriosis, history of past pelvic surgery or PID, myoma uteri, adenomyosis, endometrial polyp, endometrial hyperplasia or borderline ovarian tumour, infectious disease, chronic or acute inflammatory disease, smokers, autoimmune or systemic disorder, any malignancy, endometrioma < 10 mm or other benign adnexal mass < 30 mm *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: infertility - 10, dyspareunia - 14, dysmenorrhoea - 17, ovarian mass - 61 *Age*: mean age 36.21 ± 8.37 years *Number of participants enrolled*: 94 women *Number of participants available for analysis*: 94 women (cycle phase not reported) *Setting*: Department of O&G, Düzce University Faculty of Medicine *Place of study*: Düzce, Turkey *Period of study*: November 2009 - February 2013 *Language*: English
Index tests	*Index test*: haemoglobin, WBC, platelet count, MPV, neutrophil count, lymphocyte count, NLR, PLR, CA-125 *Details of the index test procedure as stated*: haematological parameters were analysed using a haematology analyser (Abbott CELL DYN 3700, Boston, USA); serum CA-125 levels were determined by using electro chemo-illuminescence method (Roche Hitachi Cobas 6000 E 60, Rotkreuz, Switzerland); sample handling described

Yavuzcan 2013 *(Continued)*

	Threshold for positive result: CA-125 >35 IU/ml, for other biomarkers not reported *Examiners*: no information provided; unclear if blinded to the results of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: ovarian endometriosis *Prevalence of target condition in the sample*: n = 33/94 (35%): all stage III-IV; controls n= 61: healthy controls - 33, other ovarian cyst - 28 *Reference standard*: laparoscopy/laparotomy n = 94 (100%) + histopathology *Description of positive case definition by reference standard test as reported*: staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were obtained before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	MPV, NLR and PLR values are not useful for this purpose in patients with advanced stage endometriosis that are proven to develop severe inflammation at either the cellular or molecular level
Conflict of interest	The authors declared no conflict of interests; the study did not receive any financial support
Notes	For haemoglobin, WBC, platelet count, MPV, neutrophil count, lymphocyte count, NLR and PLR, there was no statistically significant difference between the groups - no data available for meta-analysis When the data are available for the whole group of endometriosis versus controls, the diagnostic estimates for separate stages of endometriosis are not included For CA-125 there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		High	**High**
DOMAIN 2: Index Test All tests			
Were the index test results interpreted without knowledge of the	Unclear		

Yavuzcan 2013 *(Continued)*

results of the reference standard?			
If a threshold was used, was it pre-specified?	No		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Zeng 2005

Study characteristics	
Patient sampling	*Primary objective*: to evaluate the diagnostic value of examining endometrial biopsy specimens for aromatase cytochrome P450 and CA-125 for endometriosis *Participants*: patients undergoing laparoscopy or laparotomy for pelvic pain, infertility or both *Selection criteria*: inclusion criteria: reproductive age regular menstrual cycle; exclusion criteria: hormonal treatment for 3/12 months prior reproductive age, preoperative diagnosis of uterine fibroids, adenomyosis.

Zeng 2005 *(Continued)*

	Study design: cross-sectional single-gate, prospective collection of samples
Patient characteristics and setting	*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 *Period of study*: March 2003 - February 2004 *Language*: Chinese
Index tests	*Index test*: CA-125 *Details of the index test procedure as stated*: serum CA-125 was determined by chemiluminescence assay; sample handling and laboratory technique not described *Threshold for positive result*: > 35 U/ml, not pre-specified *Examiners*: no information provided *Interobserver variability*: not stated
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 36/58 (62%): stage I-II 20, stage III-IV 16; controls n = 22 *Reference standard*: laparoscopy/laparotomy N = 58 (100%) *Description of positive case definition by reference standard test as reported*: visual inspection; staging according to rAFS classification *Examiners*: not stated
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected on the day of surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The combination assay of aromatase cytochrome P450 in eutopic endometrium and CA-125 can be used as a diagnostic test for endometriosis, especially for the early stage of endometriosis, which is superior to the assay of CA-125
Conflict of interest	Not reported
Notes	Translated from Chinese The reported diagnostic estimates for combined test of endometrium and blood markers are not presented in this review
Methodological quality	

Zeng 2005 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Yes		
		Unclear	**Low**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Unclear**
DOMAIN 3: Reference Standard			
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?	Yes		
		Unclear	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Zhang 2005a

Study characteristics	
Patient sampling	*Primary objective*: to compare peritoneal fluid and serum IL-16 levels between women with and without endometriosis *Participants*: consecutive patients undergoing laparoscopy *Selection criteria*: not specified *Study design*: cross-sectional two-gate, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: controls: asymptomatic fertile women undergoing tubal sterilisation; endometriosis: women undergoing surgery for pelvic pain (n = 7), infertility (n = 6) or pelvic mass (n = 9) *Age*: mean age 37.1 ± 10.2 years (endometriosis group) and 38.6 ± 10.9 years (control group) *Number of participants enrolled*: 44 women *Number of participants available for analysis*: 44 women (25 in follicular and 19 in luteal phase of menstrual cycle) *Setting*: Department of Gynaecology, Women's hospital, Zhejiang University School of Medicine *Place of study*: Hangzhou, China *Period of study*: December 2001 - December 2002 *Language*: English
Index tests	*Index test*: IL-16 *Details of the index test procedure as stated*: serum IL-16 analysis was by Human IL-16 ELISA kit (Human IL-16 BMS 248, Bender Medsystems, Austria); laboratory technique not described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if were blinded to the result of reference standard *Interobserver variability*: not provided
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 22/44 (50%): stage I-II 8, stage III-IV 14; controls - 22 *Reference standard*: laparoscopy n = 44 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histological examinations; staging according to the rAFS scoring system *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: venous blood was obtained preoperatively *Withdrawals*: none
Comparative	
Key conclusions by the authors	Our results suggest that IL-16 is not involved in the pathogenesis of pelvic endometriosis

Zhang 2005a *(Continued)*

Conflict of interest	Not reported
Notes	For IL-16 there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Yes		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Zhang 2005a *(Continued)*

Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
	Low

Zhang 2005b

Study characteristics	
Patient sampling	*Primary objective*: to assess the levels of IL-18 in peritoneal fluid and blood of patients with endometriosis in correlation with the rAFS classification and to understand the role of IL-18 in pathogenesis of endometriosis *Participants*: women who underwent laparoscopy or laparotomy at the authors' institution and were diagnosed with endometriosis, benign ovarian mass or normal pelvis *Selection criteria*: inclusion criteria: regular menstrual cycle, no hormonal therapy 3 months before surgery, no autoimmune diseases and no malignancy *Study design*: cross-sectional, unclear if single- or two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: not specified *Age*: mean age 33.41 ± 6.53 years, range 23–45 years (endometriosis), 32.49 ± 5.02 years, range 24-44 years (controls) *Number of participants enrolled*: 60 women *Number of participants available for analysis*: 60 women (cycle phase not specified) *Setting*: Xiangya Hospital, Central South University *Place of study*: Changsha, China *Period of study*: April 2004 - September 2004 *Language*: Chinese
Index tests	*Index test*: IL-18 *Details of the index test procedure as stated*: serum levels of IL-18 were assessed by using a commercial ELISA kits (Hysen male biological reagents public division) with assay sensitivity of 6 pg/ml; sample processing described *Threshold for positive result*: not provided *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: Intra-assay CV < 10%
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 39/60 (65%): stage I-II - 12, stage III-IV - 27; controls n = 21 *Reference standard*: laparoscopy/laparotomy n = 60 (100%) + histopathology

Zhang 2005b *(Continued)*

	Description of positive case definition by reference standard as reported: visual inspection confirmed by histopathology; classification according to the rAFS score *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected immediately before surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	IL-18 levels in serum and peritoneal fluid did not correlate with a presence or severity of endometriosis
Conflict of interest	Not reported
Notes	For IL-18 there was no statistically significant difference between the groups - no data available for meta-analysis The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	Unclear		
		High	**Unclear**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Zhang 2005b *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Zhang 2006a

Study characteristics	
Patient sampling	*Primary objective*: to investigate inhibitory and activation motif expression of killer immunoglobulin-like receptor (KIR) by natural killer (NK) cells, which may be pathogenetically involved in endometriosis *Participants*: women undergoing laparoscopy for various indications *Selection criteria*: exclusion criteria: history of pregnancy or history of treatment with GnRH analogues within previous year, complications from apparent pelvic inflammatory disease *Study design*: cross-sectional, two-gate, prospective sample collection
Patient characteristics and setting	*Clinical presentation*: endometriosis group - not specified; controls: benign ovarian cysts - 21, uterine myoma - 35, infertility - 6, paraovarian cysts - 2, chronic abdominal pain - 4 *Age*: mean age 35.1 ± 7.6 years (endometriosis group), 33.9 ± 6.5 years (controls) *Number of participants enrolled*: 68 participants *Number of participants available for analysis*: 68 participants (menstrual cycle phase not reported) *Setting*: Department of O&G, Kochi Medical School *Place of study*: Kochi, Japan *Period of study*: April 2003 - May 2005 *Language*: English
Index tests	*Index test*: T cells (CD3, CD3, CD8), B cells (CD 19), NK cells (CD 56), KIR2DL1^{+}NK (CD158a+NK), KIR2DL2^{+}NK (CD158b+NK), CD94+NK, monocyte/macrophage (CD 14) and their antigen presentation *Details of the index test procedure as stated*: PBMC were measured by flow cytometry using specific mononuclear antibodies (FITC-labeled anti-CD3 and anti-CD4 mAb and PE-labelled anti-CD8 mAb as T-cell markers; PE-labelled anti-CD19 mAb as B cells marker, FITC-labelled anti-CD56 mAb for NK cells, and FITC-labeled anti-CD14 mAb for monocytes/macrophages; PE-labelled anti-CD158a and anti-CD158b as mark-

Zhang 2006a *(Continued)*

	ers for KIR subfamilies KIR2DL1 and KIR2DL2 expressed on NK cells; PE-labeled anti-CD94 mAb for lectin-like receptor; PE-labeled anti-HLA-ABC and -DR mAbs to assess antigen presentation. PE-labeled anti-CD54 mAb, CD40 mAb, CD58 mAb, CD80 mAb, and CD86 mAb - to identify co-stimulatory molecules for antigen presentation (all from Beckman-Coulter Fullerton, CA); laboratory technique described *Threshold for positive result*: not reported *Examiners*: no information provided, unclear if were blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 56/124 (45%): stage I-II - 20, stage III-IV - 36; controls n = 68 *Reference standard*: laparoscopy N = 124 (100%) *Description of positive case definition by reference standard test as reported*: staging according to rAFS classification assigned by the same operator intraoperatively and later finalised by postoperative review of video materials *Examiners*: surgical team included an expert operator who had performed laparoscopy for more than 20 years
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected at surgery *Withdrawals*: none reported
Comparative	
Key conclusions by the authors	Increased CD158a(+) NK cells in PB and PF indicated decreased NK cell cytotoxicity in endometriosis, while decreased HLA expression on PF macrophages suggested impaired antigen presentation. Thus, aberrant immune responses by NK cells and macrophages may represent risk factors for endometriosis
Conflict of interest	Not reported
Notes	For CD3, CD3, CD8, CD 19, CD 56, CD158b+NK, CD94+NK and CD14 there was no statistically significant difference between the groups - no data available for meta-analysis For CD158a+NK there was statistically significant difference between the groups, but there was insufficient data to construct 2 x 2 tables - not included in this review The data for markers measured in peritoneal fluid are not presented in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		

Zhang 2006a *(Continued)*

		High	**High**
DOMAIN 2: Index Test All tests			
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		
Was a cycle phase considered in interpretation of the result of index test?	Unclear		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
		Low	

Zhang 2006b

Study characteristics

Zhang 2006b *(Continued)*

Patient sampling	*Primary objective*: to investigate the levels of soluble intracellular adhesion molecule-1 (sICAM-1) in serum and peritoneal fluid of patients with or without endometriosis, and to discuss the clinical significance of serum sICAM-1 in pelvic endometriosis *Participants*: women who underwent surgical treatment for endometriosis or for benign epithelial ovarian tumours *Selection criteria*: exclusion criteria: steroid treatment 3 months prior to surgery, pelvic inflammatory disorder, autoimmune disease, other known internal medicine or surgical disease *Study design*: cross-sectional, two-gate design, prospective collection of samples
Patient characteristics and setting	*Clinical presentation*: endometriosis group - not specified; controls - benign ovarian mass *Age*: mean age 38.7 ± 9.5 years (endometriosis group), 36.0 ± 8.6 years (controls) *Number of participants enrolled*: 60 participants *Number of participants available for analysis*: 60 participants (cycle phase not reported) *Setting*: Department of O&G, Renmin Hospital of Wuhan University *Place of study*: Wuhan, China *Period of study*: September 2004 - March 2005 *Language*: Chinese
Index tests	*Index test*: sICAM-1 *Details of the index test procedure as stated*: serum sICAM-1 was measured with human sICAM-1 enzyme-linked immunosorbent assay (ELISA) (R&D Systems Germany), working assay range or minimal detection limit were not reported; sample handling described *Threshold for positive result*: cut-off threshold > 241.46 µg/ml, not pre-specified *Examiners*: no information provided; unclear if blinded to the result of reference standard *Interobserver variability*: not reported
Target condition and reference standard(s)	*Target condition*: endometriosis *Prevalence of target condition in the sample*: n = 30/60 (50%): stage I-II 14, stage III-IV 16; controls n = 30 *Reference standard*: laparoscopy/laparotomy N = 60 (100%) + histology *Description of positive case definition by reference standard test as reported*: visual inspection confirmed by histopathology; staging according to the rASRM classification *Examiners*: no information provided
Flow and timing	*Time interval between index test and reference standard*: blood samples were collected on the day of surgery *Withdrawals*: none
Comparative	
Key conclusions by the authors	The sICAM-1 may participate in the inflammatory process in endometriosis. Serum concentrations of sICAM-1 seem to be the effective indicator for the diagnosis of endometriosis.
Conflict of interest	Not reported

Zhang 2006b *(Continued)*

Notes	The data for markers measured in peritoneal fluid are not reported in this review

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Yes		
Was a 'two-gate' design avoided?	No		
		High	**High**
DOMAIN 2: Index Test All tests			
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?	Yes		
Was a cycle phase considered in interpretation of the result of index test?	No		
		High	**Low**
DOMAIN 3: Reference Standard			
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?	Yes		
		Low	**Low**
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Zhang 2006b *(Continued)*

Were all patients included in the analysis?	Yes
	Low

AA: arachidonic acid; **AEA**: anti-endometrial antibodies; **AUB**: abnormal uterine bleeding; **BMI**: body mass index; **BTL**: bilateral tubal ligation; **CAD**: coronary artery disease; **CFU**: colony-forming unit; **CRP**: C-reactive protein; **CV**: coefficient of variation; **Da**: dalton; **DIE**: deep infiltrating endometriosis; **DTA**: decision tree algorithm; **DTNB**: Ellman's reagent (5.5'-dithiobis-(2-nitrobenzoic acid); **ECLIA**: electrochemiluminescence immunoassay; **EIA**: enzyme immunoassay; **ELISA**: enzyme-linked immunosorbent assay; **EPA**: eicosapentaenoic acid; **ESHRE**: European Society of Human Reproduction and Embryology; **ESI-MS/MS**: electrospray ionization mass spectrometry **FACS**: Fluorescence-activated cell sorting; **FSH**: follicle-stimulating hormone; **HGF**: hepatocyte growth factor; **hs-CRP**: high sensitivity C-reactive protein; **IIF**: indirect immunofluorescence; **IRMA**: immunoradiometric assay; **IUD**: intrauterine device; **kd**: kilodalton; **KIR**: killer inhibitory receptor; **LH**: luteinising hormone; **LOOH**: lipid hydroperoxides; **LPS**: lipopolysaccharide; **LS-SVM**: least squares support vector machine; **MEIA**: microparticle enzyme immunoassay; **MF**: menstrual fluid; **miR**: microRNA; **MPV**: mean platelet volume; **MRI**: magnetic resonance imaging; **mRNA**: messenger RNA; **MW**: molecular weight; **n**: number of events/number in study arm; **N**: total sample size; **NK**: natural killer cell; **NSAID**: nonsteroidal anti-inflammatory drugs; **OCP**: oral contraception pill; **OD**: optical density; **O&G**: obstetrics and gynaecology; **PB**: peripheral blood; **PBL**: peripheral blood lymphocytes; **PCaa**: phosphatidylcholine; **PCae**: etherphospholipid; **PCOS**: polycystic ovary syndrome; **PID**: pelvic inflammatory disease; **PF**: peritoneal fluid; **PL**: plasma; **SELDI-TOF-MS**: surface enhanced laser desorption/ionisation time of flight mass spectrometry; **(r)AFS**: (revised) American Fertility Society; **(r)ASRM**: (revised) American Society for Reproductive Medicine; **RCOG**: Royal College of Obstetricians and Gynaecologists; **RIA**: radioimmunoassay; **RDF**: research development fund; **RT-PCR**: real time polymerase chain reaction; **SD:** standard deviation; **SMOH**: hydroxysphingomyelin; **TVUS**: transvaginal ultrasound; **VAS**: visual analogue scale; **WBC**: white blood cell.
For a comprehensive list of all biomarkers with their biological annotation, please see Appendix 1.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Abdallah 2006	Study groups outside inclusion criteria (comparison within endometriosis group pre- and post-surgery; no control group included)
Abrao 1997	Study design outside inclusion criteria (retrospective collection of samples); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Abrao 1999	Study design outside inclusion criteria (retrospective collection of samples)
Acien 2007	Insufficient information of study methods and population (unclear if prospective or retrospective sample collection); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Adamyan 1993	Insufficient description of study methods and population (unclear number of participants tested and if all the controls had abdominal surgery); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Agic 2007	Population outside inclusion criteria (women with pregnancy and malignancy were included)
Alcazar 2011	Index test outside inclusion criteria (lesion level analysis; unable to construct 2 x 2 tables)
Amaral 2006	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Ammendola 2008	Predictive study: test for susceptibility to endometriosis; not for diagnosis of the disease
Anastasi 2013	Target condition outside inclusion criteria (assessment of benign versus malignant ovarian tumours; not specific for endometriosis); population outside inclusion criteria (postmenopausal women included)

Study	Reason for exclusion
Andrade 2010	Study design outside inclusio criteria (retrospective sample collection)
Andrisani 2014	Reference standard outside inclusion criteria (no abdominal surgery in controls)
Antsiferova 2005	Reference standard outside inclusion criteria (no abdominal surgery in the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Arjona Berral 1996	Reference standard outside inclusion criteria (no abdominal surgery in approximately 45% of the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Avcioglu 2014	Study groups outside inclusion criteria (comparison within endometriosis group; no control group included)
Ayers 1987	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Badawy 1984	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population (unclear if prospective sample collection)
Badawy 1987	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Badawy 1990	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Balasch 1985	Study design outside inclusion criteria (retrospective sample collection)
Barbieri 1986	Population outside inclusion criteria (postmenopausal women included)
Barbieri 1987	Review article
Barrier 2002	Study design outside inclusion criteria (retrospective collection of samples); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Basta 2009	Population outside inclusion criteria (postmenopausal women included); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Bedaiwy 2006	Index test outside inclusion criteria (focus on genotype of the biomarker, not its levels)
Berkes 2013	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Bianchi 2003	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Bohler 2007	Reference standard outside inclusion criteria (no abdominal surgery in the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Bordin 2010	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Bourlev 2006a	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Bourlev 2006b	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables or to confirm negative findings)
Bragatto 2013	Study groups outside inclusion criteria (comparison within endometriosis group; no control group included)
Brinton 1996	Study design outside inclusion criteria (retrospective collection of samples)

Study	Reason for exclusion
Brosens 1978	Reference standard outside inclusion criteria (no abdominal surgery in the control group); study design outside inclusion criteria (retrospective collection of samples)
Cai 2005	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Carmona 2012	Insufficient diagnostic accuracy information (unable to construct 2 x 2 tables; presented diagnostic estimates are for ovarian endometriosis versus mixed group of controls and other type of endometriosis; no separate data for endometriosis versus controls)
Cheng 2002	Population outside inclusion criteria (only participants with positive reference standard included)
Chihal 1986	Study design outside inclusion criteria (retrospective collection of samples)
Cho 2008	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Cho 2009	Reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the control group)
Cho 2012	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Chrobak 2004	Reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Chun 2012	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Colacurci 1996b	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); population likely overlapped with Colacurci 1996a (unable to clarify with the study authors)
Confino 1990	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Cunha-Filho 2001	Insufficient description of study methods and population (unclear if prospective sample collection; unable to clarify with the study authors)
D'Amico 2013	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
D'Cruz 1996	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Darai 2003	Population outside inclusion criteria (postmenopausal women included)
Dawood 1988	Reference standard outside inclusion criteria (no abdominal surgery in the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
De Sanctis 2011	Insufficient description of study methods and population (unable to contact the study authors)
Di Stefano 1994	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population (unclear if all the controls had abdominal surgery and if prospective sample collection)
Dias 2006	Population outside inclusion criteria (only participants with positive reference standard included)
Dias 2012	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Dogan 2006	Review article
Dutta 2012	Study design outside inclusion criteria (retrospective collection of samples)

Study	Reason for exclusion
Dutta 2015	Reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the control group)
Ejzenberg 2013	Population outside inclusion criteria (likely only participants with positive reference standard included; unable to clarify with the study authors); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Fallat 1997	Study design outside inclusion criteria (retrospective collection of samples); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Fedele 1988	Population overlapped with Fedele 1989
Fernandez-Shaw 1993	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Fernandez-Shaw 1996	Insufficient description of study methods and population (unable to clarify with the study authors)
Ferrero 2005b	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Fisk 1988	Insufficient description of study methods and population (unable to clarify with the study authors)
Flores 2006	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Fu 2002	Reference standard outside inclusion criteria (no abdominal surgery in the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Fujii 2008	Reference standard outside inclusion criteria (no abdominal surgery in the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Gagne 2003c	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Gajbhiye 2008	Insufficient description of study methods and population (unable to contact the study authors)
Gajbhiye 2012	Insufficient description of study methods and population (unable to contact the study authors)
Galleri 2009	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Galo 2005	Population outside inclusion criteria (postmenopausal women included); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Garcia-Manero 2007	Study groups outside inclusion criteria (comparison within endometriosis group; no control group included)
Garcia-Velasco 2002	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Garza 1991	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Garzetti 1994	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Gebel 1993	Insufficient description of study methods and population (unclear age group, no separate data for women with untreated endometriosis; unable to contact the study authors)
Gebel 1995	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables or to confirm negative findings)
Giudice 1986	Population outside inclusion criteria (postmenopausal women and women with malignancy included)

Study	Reason for exclusion
Gmyrek 2005	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Gorski 2007	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Guerriero 1997	Index test outside inclusion criteria (data for combined blood test and imaging, no separate data for blood biomarker); population overlapped with Guerriero Guerriero 1996a and Guerriero 1996b
Gunev 1981	Insufficient description of study methods and population (unclear if all the participants had abdominal surgery and if prospective sample collection; unable to contact the study authors)
Hammadeh 2003	Study design outside inclusion criteria (retrospective sample collection); reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the control group)
Han 2009	Predictive study (test for susceptibility for endometriosis, not for diagnosis of the disease)
Hatayama 1996	Insufficient description of study methods and population (unable to contact the study authors)
He 1993	Study design outside inclusion criteria (retrospective collection of samples)
Hompes 1996	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Hornstein 1992	Population likely overlapped with Hornstein 1995; unable to contact the study authors
Hrycek 1996	Reference standard outside inclusion criteria (no abdominal surgery in the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Hsu 1997	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables or to confirm negative findings)
Hsu 2014	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Huang 2004	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Hwang 2014	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population (unclear if prospective sample collection)
Ihlenfeld 2007	Full text not available (unable to contact the study authors)
Illera 2001	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population
Izumiya 2003	Index test outside inclusion criteria (data for peritoneal fluid to peripheral blood macrophage ratio; no separate data for blood biomarker)
Jackson 2005	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population (unclear if prospective sample collection)
Jana 2013	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population (unclear if prospective sample collection)
Jedryka 2001	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Jerzak 2002	Insufficient description of study methods and population (unclear if all the controls had abdominal surgery and if prospective sample collection; unable to contact the study authors)

Study	Reason for exclusion
Jing 2009	Reference standard outside inclusion criteria (no surgery in approximately 50% of the control group)
Kabut 2007	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Kadija 2012	Target condition outside inclusion criteria (assessment of benign versus malignant ovarian tumours; not specific for endometriosis)
Kafali 2004	Study design outside inclusion criteria (retrospective collection of samples)
Kang 1988	Study design outside inclusion criteria (retrospective collection of samples); insufficient description of study methods and population (unclear if all the controls had surgery)
Kataoka 2012	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Kharfi 2002	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables or to confirm negative findings)
KhoshdelRad 2014	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Kichuchi 1993	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Kiechle 1994	Reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the control group)
Kilpatrick 1991	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Kim 1995	Population outside inclusion criteria (umbilical cord blood served as control samples)
Kim 2007	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Kim 2014	Study groups outside inclusion criteria (comparison within endometriosis group; no control group included)
Kinugasa 2011	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population (unclear if prospective sample collection)
Kobayashi 1987	Reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the study and control group); insufficient description of study methods and population (unclear if all the participants were of reproductive age and time interval between sample collection and surgery)
Kondera-Anasz 2004	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Kondera-Anasz 2005	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Koninckx 1992	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Kopuz 2014	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Koumantakis 1994	Insufficient description of study methods and population (unclear if all the controls had abdominal surgery)
Kralickova 2007	Target condition outside inclusion criteria (assessment of leukaemia-inhibitory factor mutation-positive versus leukaemia-inhibitory factor mutation-negative women; no separate analysis for endometriosis)

Study	Reason for exclusion
Krasnicki 2001	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Kurt 2014	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Lambert 2014	Unable to locate the full text
Lang 2001	Population outside inclusion criteria (male donors served as controls)
Lee 2014	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Leggieri 2014	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Leng 2002	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Lenhard 2011	Target condition outside inclusion criteria (assessment of benign verus ovarian tumours of low malignant potential; not specific for endometriosis); population outside inclusion criteria (post-menopausal women included)
Lermann 2010	Insufficient description of study methods and population (unclear age group and if prospective collection of samples; unable to clarify with the study authors)
Li 2000	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Li 2010	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Linghu 2004	Reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the control group)
Liu 2007	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Liu 2013	Study design outside inclusion criteria (retrospective sample collection)
Long 2013	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Luo 2005	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Maeda 2004	Population overlapped with Maeda 2002a and Maeda 2002b
Mahdian 2015	Reference standard outside inclusion criteria (no abdomial surgery in the control group); insufficient diagnostic accuracy information (unable to construct 2 x 2 tables)
Malvezzi 2013	Study design outside inclusion criteria (retrospective collection of samples)
Manero 2009	Study groups outside inclusion criteria (comparison within endometriosis group; no control group included)
Manero 2010	Study groups outside inclusion criteria (comparison within endometriosis group; no control group included)
Markham 1997b	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Masahashi 1988	Reference standard outside inclusion criteria (no abdominal surgery in the control group); study design outside inclusion criteria (retrospective collection of samples)

Study	Reason for exclusion
Matalliotakis 1994	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population (unclear if controls had abdominal surgery)
Matalliotakis 1997	Study design outside inclusion criteria (retrospective sample collection); insufficient information on study population (unclear if controls had abdominal surgery)
Matalliotakis 2000	Study design outside inclusion criteria (retrospective sample collection)
Matalliotakis 2001a	Study design outside inclusion criteria (retrospective sample collection); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Matalliotakis 2001b	Study design outside inclusion criteria (retrospective sample collection); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Matalliotakis 2003b	Study design outside inclusion criteria (retrospective sample collection); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Matarese 2000	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Mathur 1982	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Mathur 1990	Descriptive study; no focus on diagnostic performance of the test
Mathur 1998	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Mathur 1999	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Mathur 2000	Review article
Matsuoka 2005	Population overlapped with Zhang 2006a
Medl 1997	Population outside inclusion criteria (postmenopausal women included)
Michaud 2014	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population (unclear if prospective sample collection)
Moloney 1989	Study design outside inclusion criteria (retrospective collection of samples)
Moretuzzo 1988	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Nabeta 2009	Reference standard outside inclusion criteria (no abdominal surgery in approximately 50% of the control group); population outside inclusion criteria (women with known malignancy included)
Nabeta 2011	Reference standard outside inclusion criteria (no abdominal surgery in approximately 50% of the control group); population outside inclusion criteria (women with known malignancy included)
Nagamani 1992	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Nalbanski 2008	Study design outside inclusion criteria (retrospective collection of samples); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Nomiyama 1997	Insufficient description of study methods and population (unclear if prospective sample collection; unable to contact the study authors)
O'Shaughnessy 1993	Insufficient description of study methods and population (unable to contact the study authors); insufficient diagnostic accuracy information (unable to construct 2 x 2 tables)

Study	Reason for exclusion
Odukoya 1995a	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Odukoya 1995b	Insufficient description of study methods and population (unable to contact the study authors)
Ota 1990	Reference standard outside inclusion criteria (no abdominal surgery in the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Ota 1991	Study groups outside inclusion criteria (comparison of endometriosis group with adenomyosis; no control group included)
Ozaksit 1995	Study design outside inclusion criteria (retrospective sample collection)
Ozasa 1987	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Perwira 2009	Full text not available (unable to contact the study authors)
Pittaway 1986	Target condition outside inclusion criteria (evaluation of blood biomarker in various pathological/physiological conditions; unable to obtain separate data for endometriosis)
Pittaway 1987a	Population likely overlapped with Pittaway 1989
Pittaway 1987b	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Pizzo 2002	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Podgaec 2010	Population overlapped with Podgaec 2007
Pupo-Nogueira 2007	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables or to confirm negative findings)
Quaranta 2006	Study question outside inclusion criteria: focus on the impact of environmental contaminants on the dysregulation of immune function in endometriosis
Rajkumar 1992	Insufficient description of study methods and population (unclear age group; unable to contact the study authors)
Ramos 2011	Population overlapped with Ramos 2012
Reis 2012	Reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the control group)
Santulli 2015	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Sengul 2014	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Seo 2010	Reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the control group)
Sha 2009	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Shanti 1999	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Sharma 2010	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)

Study	Reason for exclusion
Sharpe-Timms 1998	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Signorile 2014	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Slabe 2013	Insufficient description of study methods and population (unclear if prospective sample collection; unable to contact the study authors)
Socolov 2011	Population outside inclusion criteria (women with ectopic pregnancy, pelvic inflammatory disease and other known pathologies included)
Steff 2004b	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Suryawanshi 2013	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Szyllo 2001	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Takahashi 1987	Study groups outside inclusion criteria (comparison within endometriosis groups; no controls included)
Takahashi 1988	Reference standard outside inclusion criteria (no abdominal surgery in the control group); study design outside inclusion criteria (retrospective collection of samples)
Takahashi 1989	Reference standard outside inclusion criteria (no abdominal surgery in the control group); study design outside inclusion criteria (retrospective collection of samples)
Takemura 2005	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Tanaka 2000	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables); insufficient description of study methods and population (unclear if retrospective sample collection and if all the controls had abdominal surgery)
Telimaa 1989	Reference standard outside inclusion criteria (no abdominal surgery in the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Tsao 2007	Focus on screening, not on diagnostic performance of the test
Tuten 2014b	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Venturella 2011	Insufficient description of study methods and population (unclear if prospective sample collection; unable to clarify with the study authors); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Vercellini 1992	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Wang 2007	Reference standard outside inclusion criteria (no abdominal surgery in the control group); population outside inclusion criteria (postmenopausal women included)
Wang 2008	Reference standard outside inclusion criteria (no abdominal surgery in the control group); population outside inclusion criteria (postmenopausal women included)
Wang 2009	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Wang 2013b	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Watanabe 1990	Reference standard outside inclusion criteria (no abdominal surgery in subset of subjects within the study and control group); insufficient description of study methods and population (unclear

Study	Reason for exclusion
	if all the participants were of reproductive age and time interval between sample collection and surgery)
Wild 1985	Population overlapped with Wild Wild 1991a; insufficient description of methods and population
Wild 1991b	Population overlapped with Wild 1991a
Wild 1991c	Evaluation of the laboratory techniques; no focus on diagnostic accuracy of the test
Wild 1992	Evaluation of the laboratory techniques; no focus on diagnostic accuracy of the test
Wilson 1994	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Wojcik-Krowiranda 2010	Population outside inclusion criteria (postmenopausal women included)
Xavier 2006	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Yang 2013a	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Yang 2013b	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Yi 2010	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Yin 2000	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Zachariah 2009	Reference standard outside inclusion criteria (no abdominal surgery in the control group); insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Zhang 2006c	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Zhang 2009	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Zhao 2015	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Zheng 2011	Reference standard outside inclusion criteria (no abdominal surgery in the control group)
Zhu 2007	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)
Zomer 2013	Study groups outside inclusion criteria (comparison within endometriosis group; no control group included)
Zong 2003	Insufficient diagnostic test accuracy information (unable to construct 2 x 2 tables)

Characteristics of ongoing studies *[ordered by study ID]*

JPRN-UMIN000009223

Trial name or title	Analysis of miRNA in blood for development of diagnostic biomarkers for endometriosis ClinicalTrials.gov Identifier: JPRN-UMIN000009223 Primary sponsor: Juntendo University Hospital, Department of Obstetrics and Gynecology
Target condition and reference standard(s)	**Objective:** To identify endometriosis-specific microRNAs in blood and to develop a diagnostic test for endometriosis

JPRN-UMIN000009223 *(Continued)*

	Primary outcome measures: Concentration of microRNAs in blood **Study design:** Observational **Target condition:** Endometriosis **Reference standard:** Laparoscopy
Index and comparator tests	Blood
Starting date	February 2013
Contact information	Name: Ikuo Mori DVM, Ph.D Address: 26-1, Muraoka-Higashi, Fujisawa, Kanagawa 251-8555, JAPAN Japan Email: ikuo.mori@takeda.com Affiliation: Takeda Pharmaceutical Company Limited Integrated Technology Research Laboratories, Pharmaceutical Research Division Name: Mari Kitade MD Address: Hongo 3-1-3, Bunkyo-ku, Tokyo 113-8431, Japan Telephone: 03-3813-3111 Email: kitade@juntendo.ac.jp Affiliation: Juntendo University Hospital Department of Obstetrics and Gynecology
Notes	Current status - ongoing, recruiting participants

NCT01301885

Trial name or title	ENDOMET - Novel diagnostic tools and treatments for endometriosis ClinicalTrials.gov Identifier: NCT01301885 Other study name: CA125_VAS_changes
Target condition and reference standard(s)	**Objective:** To identify expression of endometriosis specific RNAs/proteins **Primary outcome measures:** Concentration of protein and DNA in biological fluids and tissues in association with endometriosis **Study design:** Observational case-control, prospective **Target condition:** Endometriosis **Reference standard:** Laparoscopy
Index and comparator tests	Serum, peritoneal fluid, endometrium tissue, healthy peritoneum, tissue of endometriosis (peritoneal, ovarian, deep infiltrating) Extracted DNA, RNA, cDNA and protein from the above samples
Starting date	February 2011
Contact information	Responsible party: Antti Perheentupa, Turku University Hospital
Notes	Current status - ongoing, but not recruiting participants

NCT02091557

Trial name or title	CA-125 and VAS pain score changes to diagnose endometriosis ClinicalTrials.gov Identifier: NCT02091557 Other study name: CA125_VAS_changes
Target condition and reference standard(s)	**Objective:** To assess the diagnostic accuracy for the noninvasive detection of pelvic endometriosis of the combination of two simple parameters: modifications of serum CA-125 and VAS pain score following one dose of GnRH-a **Primary outcome measures:** Serum CA-125 level taken in follicular cycle phase (2nd-3rd day of the menstrual cycle) and VAS score for menstrual pain. During the time passed on surgery waiting list, patients will receive LAD at a dose of 3.75 mg IM on the 21st day of the menstrual cycle. One month later, LAD administration, serum CA-125 levels and VAS score will be assessed again, and then the surgical procedure will be performed in all these patients **Study design:** Observational cohort, prospective **Target condition:** Endometriosis **Reference standard:** Laparoscopy
Index and comparator tests	Blood
Starting date	January 2011
Contact information	Responsible party: Fulvio Zullo, University Magna Graecia
Notes	Current status - completed, results not available

NCT02337816

Trial name or title	Role of metabolomics in the diagnosis of endometriosis ClinicalTrials.gov Identifier: NCT02337816 Other study name: ENDOMETAB01
Target condition and reference standard(s)	**Objective:** To identify an alteration in the expression of the metabolites in women with endometriosis **Primary outcome measures:** Plasma and urine concentration of metabolites (time frame: at least one month after discontinuation of hormonal therapies and before laparoscopic surgery) **Study design:** Non-randomised, parallel assignment, open label **Target condition:** Endometriosis **Reference standard:** Laparoscopy + histopathology
Index and comparator tests	Urine and blood
Starting date	December 2014
Contact information	Responsible party: Stefano Angioni, University of Cagliari
Notes	Current status - ongoing, but not recruiting participants

CA-125: cancer antigen-125; **cDNA**: complementary DNA;**GnRH-a**: gonadotropin-releasing hormone analogue; **IM**: intramuscular; **LAD**: leuprolide acetate depot; **VAS**: visual analogue scale

DATA

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 Glycodelin-A (> 2.07 ng/ml)	1	99
2 Glycodelin (> 9.0 ng/ml)	1	45
3 Glycodelin (> 18 ng/ml)	1	99
4 IGFBP-3 (> 200 ng/ml)	1	45
5 IGFBP-3 (> 210 ng/ml)	1	99
6 VEGF (> 1.5 pg/ml)	1	99
7 VEGF (> 236 pg/ml)	1	95
8 VEGF-A (> 680 pg/ml)	1	60
9 Urocortin (> 29 pg/ml), endometrioma	1	80
10 Urocortin (> 33 pg/ml), endometrioma	1	80
11 Urocortin (> 41.6 pg/ml), endometrioma	1	88
12 Survivin (cut-off not reported)	1	60
13 sICAM-1 (< 243 ng/ml)	1	100
14 sICAM-1 (< 254.6 ng/ml)	1	28
15 sICAM-1 (> 241.46 µg/ml)	1	60
16 LN-1 (> 1110.0 pg/ml)	1	73
17 Metabolome by ESI-MS/MS (SMOH C16:1 + PCaa C36:2/ PCae C34:2) age-/ BMI-adjusted	1	92
18 Proteome by SELDI-TOF-MS (3 peaks with the molecular weight of 3,956.00, 11,710.00 and 6,986.00 Da)	1	31
19 Proteome by SELDI-TOF MS (5 peaks with molecular weights of 4159.00, 5264.00, 5603.00, 9861.00 and 10,533.00 Da)	1	90
20 Proteome by SELDI-TOF MS (5 peaks with molecular weight of 9,926.31, 10,072.2, 6,753.04, 4,302.67, 9,328.49 Da)	1	67

Test	No. of studies	No. of participants
21 Proteome by SELDI-TOF MS (5 peaks with molecular weight of 2,831.02, 7,554.66, 4,241.29, 2,953.25, 9,927.73 Da)	1	98
22 Proteome by SELDI-TOF MS (5 peaks with molecular weight of 11,366.3, 5,712.69, 10,070.7, 3,017.68, 3,824.44 Da)	1	88
23 Proteome by SELDI-TOF-MS (6 peaks with molecular weights of 1629.00 3047.00, 3526.00, 3774.00, 5046.00 and 5068.00 Da)	1	139
24 Prolactin (> 14.8 ng/ml)	1	97
25 Prolactin (> 20 ng/ml)	1	97
26 Anti-endometrial Abs, IgG	4	759
27 Anti-endometrial Abs (MW 26/34/42 kd)	1	36
28 Anti-laminin auto Abs, IgG (> 1 U/ml)	1	68
29 sCD23 (cut-off not reported)	1	97
30 MCP-1 (> 100 pg/ml)	1	101
31 Copeptin (> 251.18 pg/ml)	1	87
32 hs-CRP (> 0.61 mg/l)	1	119
33 hs-CRP (> 0.62 mg/l)	1	295
34 hs-CRP (> 0.70 mg/l)	1	116
35 hs-CRP (> 0.73 mg/l)	1	60
36 hs-CRP (> 438 µg/ml)	1	95
37 hs-CRP (cut-off not reported)	1	116
38 IFN-γ (< 76 pg/ml)	1	45
39 MIF (> 0.57 ng/ml)	1	93
40 TNF-α (> 12.45 pg/ml)	1	95
41 TNF-α (< 45.6 pg/ml)	1	45
42 TNF-α (cut-off not reported)	1	116
43 Neutrophils (> 4058/ml)	1	100
44 NLR (> 2.19)	1	100
45 WBC (> 6400/ml)	1	100
46 IL-1β (< 0.9 pg/ml)	1	45

Test	No. of studies	No. of participants
47 IL-4 (≥ 3 pg/ml)	1	50
48 IL-6 (> 1.03 pg/ml)	1	138
49 IL-6 (> 1.9 pg/ml)	1	138
50 IL-6 (> 2 pg/ml)	2	171
51 IL-6 (> 2.6 pg/ml)	1	138
52 IL-6 (> 4 pg/ml)	1	91
53 IL-6 (> 7.5 pg/ml)	1	91
54 IL-6 (< 10 pg/ml)	1	45
55 IL-6 (> 12.2 pg/ml)	1	95
56 IL-6 (> 15.4 pg/ml)	1	78
57 IL-6 (> 25.75 pg/ml)	1	83
58 IL-6 (cut-off not reported)	1	116
59 IL-8 (> 24 pg/ml)	1	101
60 IL-8 (≥ 25 pg/ml), endometrioma	1	91
61 IL-8 (cut-off not reported)	1	116
62 Follistatin (> 1433 pg/ml), endometrioma	1	104
63 STX-5 (> 55 ng/ml)	1	80
64 Carbonyls (< 14.9 µM)	1	108
65 PON-1 (< 141.5 U/l)	1	87
66 Thiols (< 396.44 µM)	1	108
67 miR-9* (cut-off not reported)	1	85
68 miR-17-5 (< 0.9057)	1	40
69 miR-20a (< 0.6879)	1	40
70 miR-22 (< 0.5647)	1	40
71 miR-122 (cut-off not reported)	1	85
72 miR-141* (cut-off not reported)	1	85
73 miR-145* (cut-off not reported)	1	85
74 miR-199a (cut-off not reported)	1	85

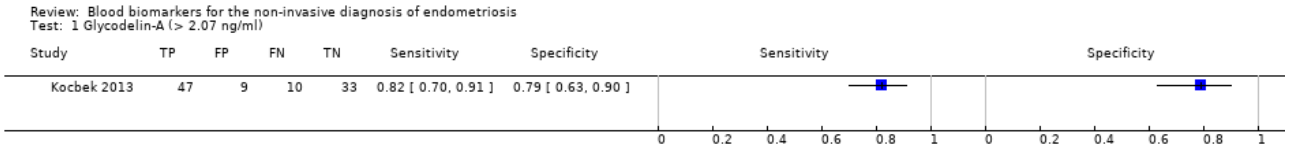
Test	No. of studies	No. of participants
75 miR-532-3p (cut-off not reported)	1	85
76 Ca-15.3 (> 15 IU/ml)	1	88
77 Ca-15.3 (> 30 IU/ml)	1	119
78 CA-19.9 (> 7.5 IU/ml)	1	76
79 CA-19.9 (> 9.5 IU/ml)	1	198
80 CA-19.9 (> 10.67 IU/ml)	1	88
81 CA-19.9 (≥ 12 U/ml), endometrioma	1	118
82 CA-19.9 (> 37 IU/ml)	3	330
83 CA-19.9 (cut-off not reported)	2	176
84 CA-72 (TAG-72) (> 4 U/ml)	1	35
85 CA-72 (TAG-72) (> 6 U/ml)	1	119
86 CA-125 (> 10 IU/ml)	1	201
87 CA-125 (> 11 U/ml)	1	59
88 CA-125 (> 11.5 U/ml)	1	45
89 CA-125 (> 12.5 U/ml)	1	99
90 CA-125 (> 12.8 U/ml)	1	368
91 CA-125 (> 13.5 U/ml)	1	35
92 CA-125 (> 14.7 IU/ml)	1	60
93 CA-125 (> 16 U/ml)	4	335
94 CA-125 (> 17.6 IU/ml)	1	95
95 CA-125 (> 20 IU/ml)	4	1115
96 CA-125 (> 20 U/ml), endometrioma	2	189
97 CA-125 (> 25 U/ml), endometrioma	1	101
98 CA-125 (> 26 IU/ml)	2	862
99 CA-125 (> 30 U/ml)	3	943
100 CA-125 (> 30 U/ml), endometrioma	2	163
101 CA-125 (> 33 U/ml)	1	100
102 CA-125 (> 35 U/ml)	25	3266

Test	No. of studies	No. of participants
103 CA-125 (> 35 U/ml), endometrioma	1	101
104 CA-125 (> 36 U/l) endometrioma	1	80
105 CA-125 (> 42 U/l), endometrioma	1	104
106 CA-125 (> 43 U/ml)	1	62
107 CA-125 (cut-off not reported)	1	59
108 CA-125 (cut-off not reported)	1	119
109 CA-125 (cut-off not reported)	1	60
110 CA-125 (cut-off not reported)	1	116
111 Combined test (CA-125 ≥ 25 U/ml +/or CA-19.9 ≥ 12 U/ml), endometrioma	1	118
112 Combined test (CA-125 ≥ 25 U/ml + Ca-19.9 ≥ 12 U/ml), endometrioma	1	118
113 Combined test (CA-125 > 19.8 U/l + Prolactin > 14.8 ng/ml)	1	97
114 Combined test (CA-125 > 35 U/l + Prolactin > 20 ng/ml)	1	97
115 Combined test (CA-125 > 17.6 IU/ml + VEGF > 236 pg/ml)	1	95
116 Combined test (CA-125 > 20 U/l + Anti-endometrial Abs > 0.3 A-value)	1	42
117 Combined test (CA-125 x NLR; (> 43.1)	1	100
118 Combined test (CA-125 > 30 U/ml +/or IL-8 ≥ 25 pg/ml), endometrioma	1	83
119 Combined test (CA-125 + IL-8) (cut-off not reported)	1	294
120 Combined test (IL-6 > 12.2 pg/ml + TNF-α > 12.45 pg/ml)	1	96
121 Combined test (IL-6 > 12.2 pg/ml + CRP > 438 µg/ml)	1	95
122 Combined test (TNF-α > 12.45 pg/ml + CRP > 438 µg/ml)	1	95
123 Combined test (miR-199a + miR-122) (cut-off not reported)	1	85
124 Combined test (miR-199a + miR-542-3p) (cut-off not reported)	1	85
125 Combined test (Ca-125 + Ca 19-9 + Survivin) (cut-off not reported)	1	60
126 Combined test (CA-125 + STX-5 + LN-1) (cut-off not reported)	1	80
127 Combined test (CA-125 > 35 IU/ml +/or CA-19.9 > 37 IU/ml +/or IL-6 > 2 pg/ml)	1	80
128 Combined test (CA-125 > 50 IU/mL +/ or CCR1 > 1.16 +/or MCP-1 > 140 pg/ml)	1	151

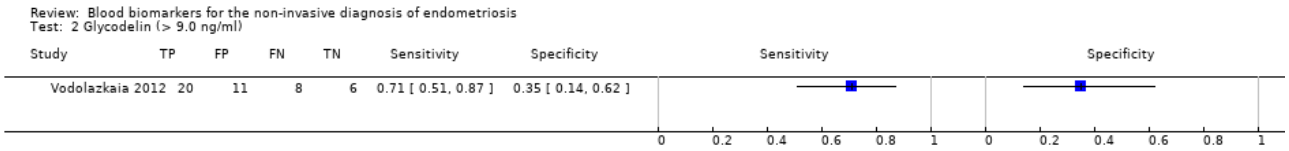
Test	No. of studies	No. of participants
129 Combined test (Ca-125 > 20 mIU/ml + MCP-1 > 152.74 pg/ml + Leptin > 3.14 ng/ml)	1	141
130 Combined test CA-125 + IL-8 + TNF-α) (cut-off not reported)	1	116
131 Combined test (IL-6 > 12.2 pg/ml + TNF-α > 12.45 pg/ml + CRP > 438 µg/ml)	1	95
132 Combined test (CA-125 + VEGF + annexin V + glycodelin] - MLR (cut-off not reported)	1	19
133 Combined test (CA-125 + VEGF + annexin V + glycodelin] - LS-SVM (cut-off not reported)	1	19
134 Combined test (CA-125 + VEGF + annexin V + sICAM-1) - MLR or LS-SVM (cut-off not reported)	1	19
135 Combined test (CA-125 > 20 mIU/ml + MCP-1 > 53.5 pg/ml + Leptin > 29.1 ng/ml + MIF > 14.7 ng/ml)	1	141
136 Combined test (miR-199a + miR-122 + miR-145* + miR-542-3p) (cut-off not reported)	1	85
137 Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)	1	294
138 Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)	1	59
139 Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)	1	119
140 Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported)	1	116
141 CA-125 (> 20 U/ml), Bilibio 2014	1	97
142 CA-125 (> 35 U/ml), Bilibio 2014	1	97
143 CA-125 (> 16 U/ml), Ferreira 1994	1	41
144 CA-125 (> 35 U/ml), Ferreira 1994	1	41
145 CA-125 (> 30 U/ml), Florio 2007	1	80
146 CA-125 (> 36 U/ml), Florio 2007	1	80
147 CA-125 (> 12.8 U/ml), Gagne 2003a	1	368
148 CA-125 (> 35 U/ml), Gagne 2003a	1	368
149 CA-125 (> 20 U/ml), Guerriero 1996b	1	101
150 CA-125 (≥ 25 U/ml), Guerriero 1996b	1	101
151 CA-125 (> 35 U/ml), Guerriero 1996b	1	101

Test	No. of studies	No. of participants
152 CA-125 (> 20 U/ml), Kitawaki 2005	1	775
153 CA-125 (> 26 U/ml), Kitawaki 2005	1	775
154 CA-125 (> 30 U/ml), Kitawaki 2005	1	775
155 CA-125 (> 35 U/ml), Kitawaki 2005	1	775
156 CA-125 (> 10 U/ml), Rosa E Silva 2007	1	201
157 CA-125 (> 20 U/ml), Rosa E Silva 2007	1	201
158 CA-125 (> 20 U/ml), Yang 1994	1	42
159 CA-125 (> 35 U/ml), Yang 1994	1	42
160 IL-6 (> 1.03 pg/ml), Othman 2008	1	138
161 IL-6 (> 1.9 pg/ml), Othman 2008	1	138
162 IL-6 (> 2.6 pg/ml), Othman 2008	1	138
163 IL-6 (> 2 pg/ml), Bedaiwy 2002	1	91
164 IL-6 (> 4 pg/ml), Bedaiwy 2002	1	91
165 IL-6 (> 7.5 pg/ml), Bedaiwy 2002	1	91

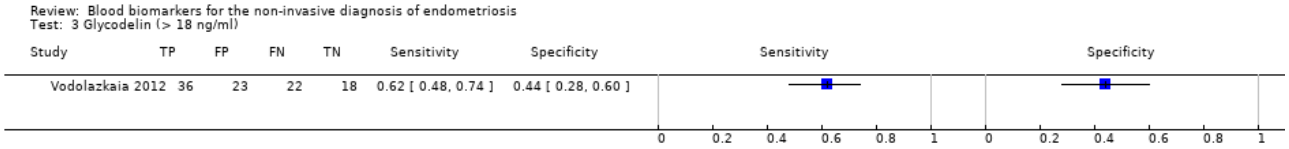
Test 1. Glycodelin-A (> 2.07 ng/ml).



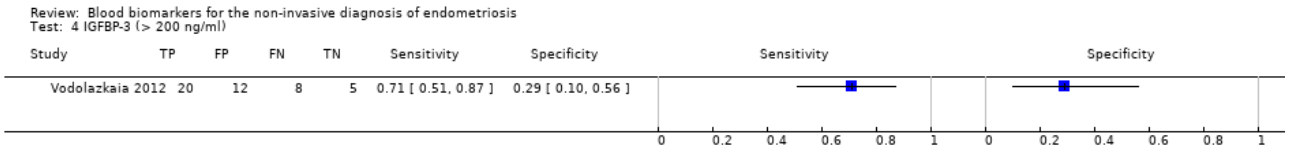
Test 2. Glycodelin (> 9.0 ng/ml).



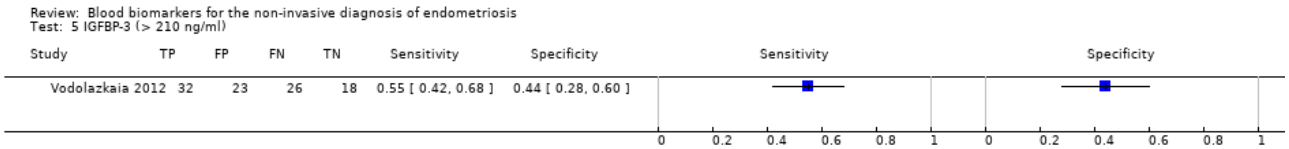
Test 3. Glycodelin (> 18 ng/ml).



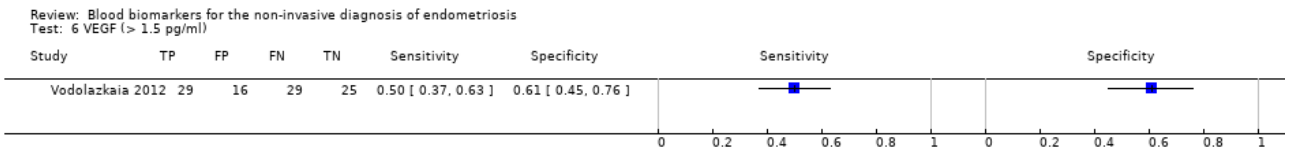
Test 4. IGFBP-3 (> 200 ng/ml).



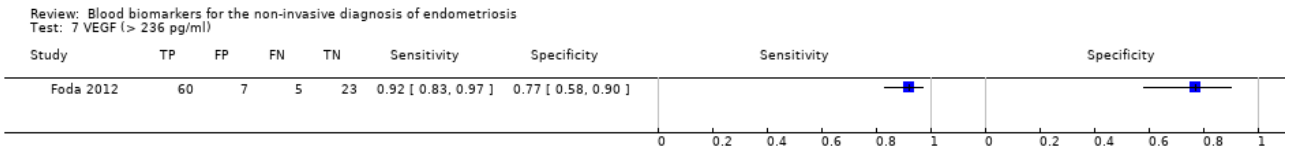
Test 5. IGFBP-3 (> 210 ng/ml).



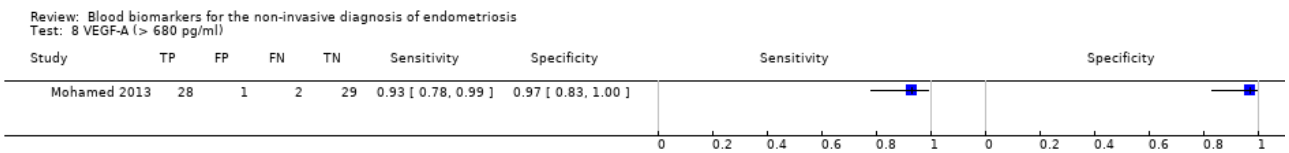
Test 6. VEGF (> 1.5 pg/ml).



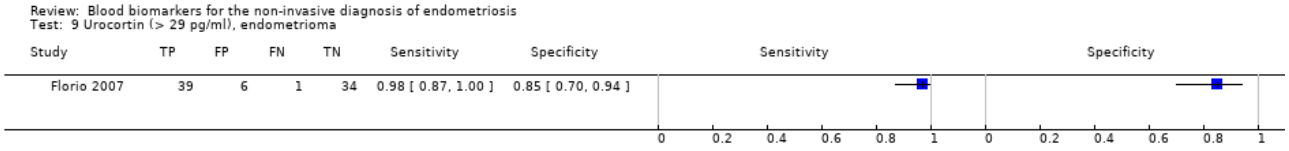
Test 7. VEGF (> 236 pg/ml).



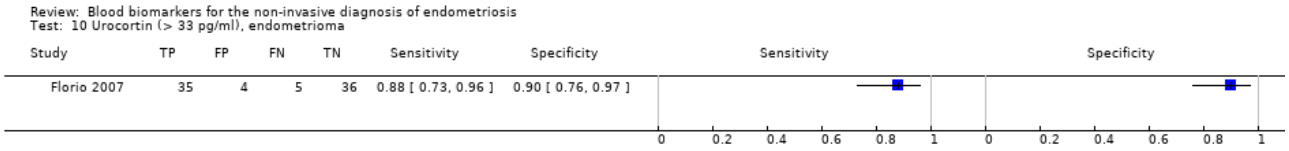
Test 8. VEGF-A (> 680 pg/ml).



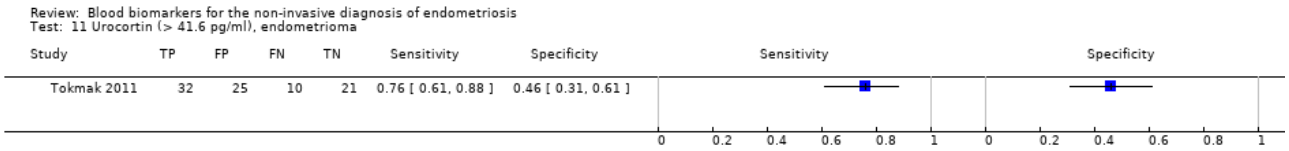
Test 9. Urocortin (> 29 pg/ml), endometrioma.



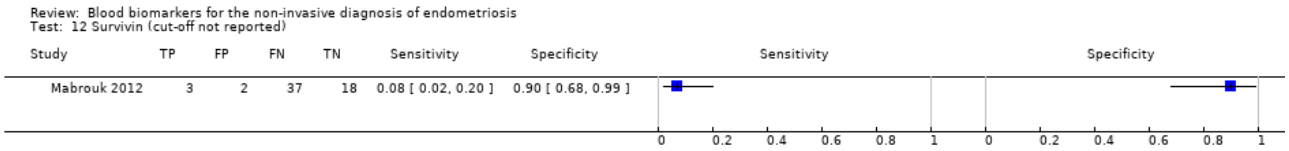
Test 10. Urocortin (> 33 pg/ml), endometrioma.



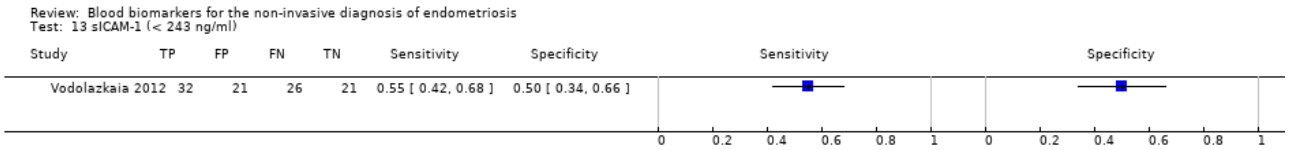
Test 11. Urocortin (> 41.6 pg/ml), endometrioma.



Test 12. Survivin (cut-off not reported).

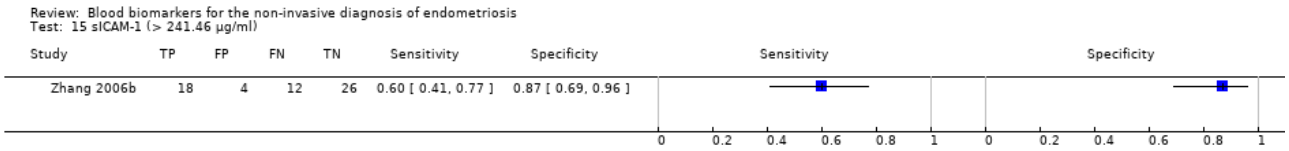


Test 13. sICAM-1 (< 243 ng/ml).

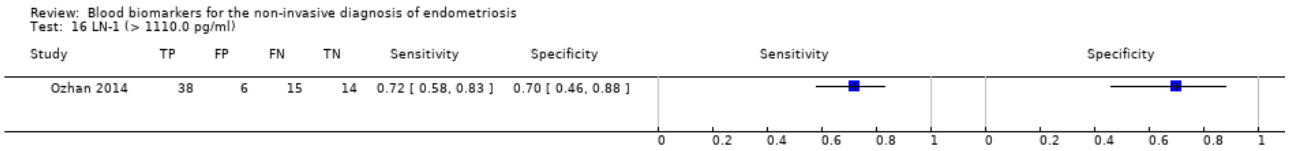


Test 14. sICAM-1 (< 254.6 ng/ml).

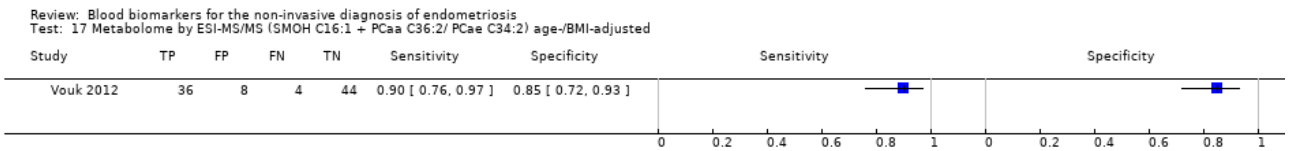
Test 15. sICAM-1 (> 241.46 µg/ml).



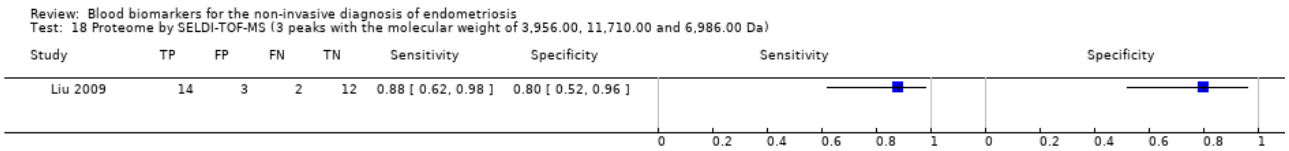
Test 16. LN-1 (> 1110.0 pg/ml).



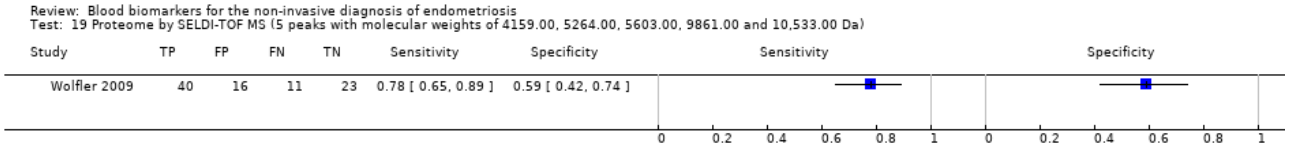
Test 17. Metabolome by ESI-MS/MS (SMOH C16:1 + PCaa C36:2/ PCae C34:2) age-/BMI-adjusted.



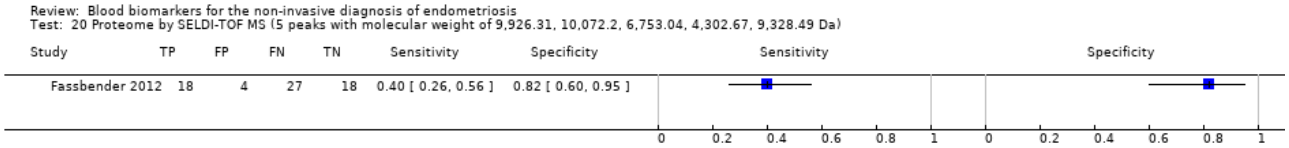
Test 18. Proteome by SELDI-TOF-MS (3 peaks with the molecular weight of 3,956.00, 11,710.00 and 6,986.00 Da).



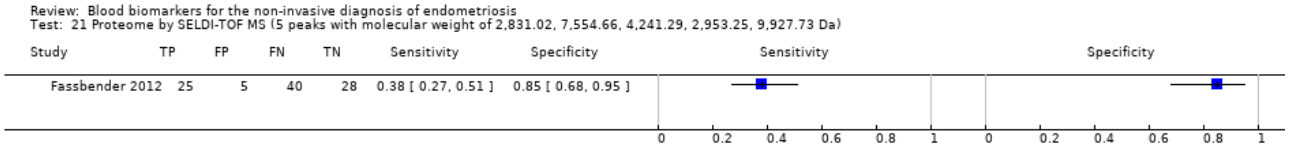
Test 19. Proteome by SELDI-TOF MS (5 peaks with molecular weights of 4159.00, 5264.00, 5603.00, 9861.00 and 10,533.00 Da).



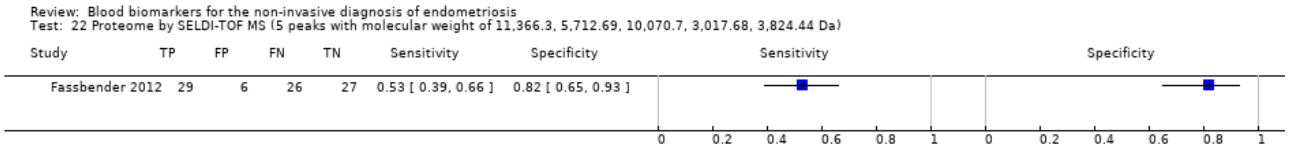
Test 20. Proteome by SELDI-TOF MS (5 peaks with molecular weight of 9,926.31, 10,072.2, 6,753.04, 4,302.67, 9,328.49 Da).



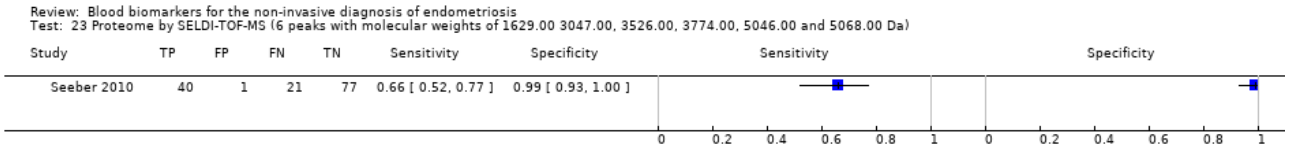
Test 21. Proteome by SELDI-TOF MS (5 peaks with molecular weight of 2,831.02, 7,554.66, 4,241.29, 2,953.25, 9,927.73 Da).



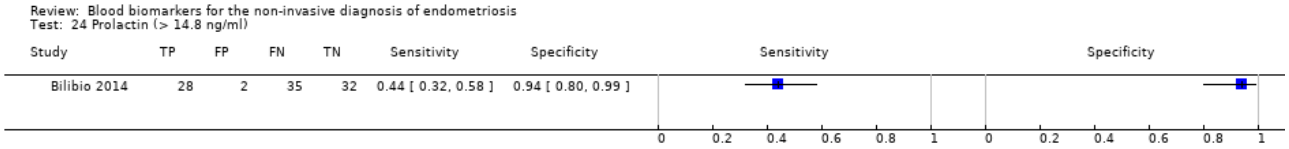
Test 22. Proteome by SELDI-TOF MS (5 peaks with molecular weight of 11,366.3, 5,712.69, 10,070.7, 3,017.68, 3,824.44 Da).



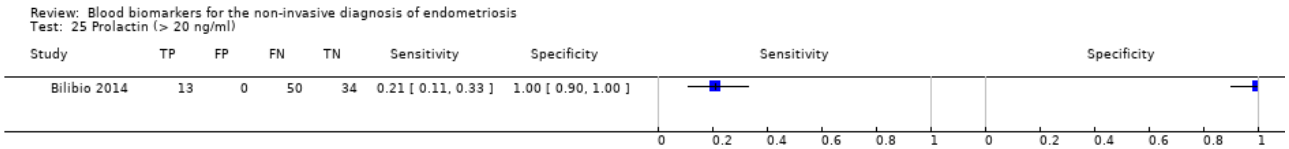
Test 23. Proteome by SELDI-TOF-MS (6 peaks with molecular weights of 1629.00 3047.00, 3526.00, 3774.00, 5046.00 and 5068.00 Da).



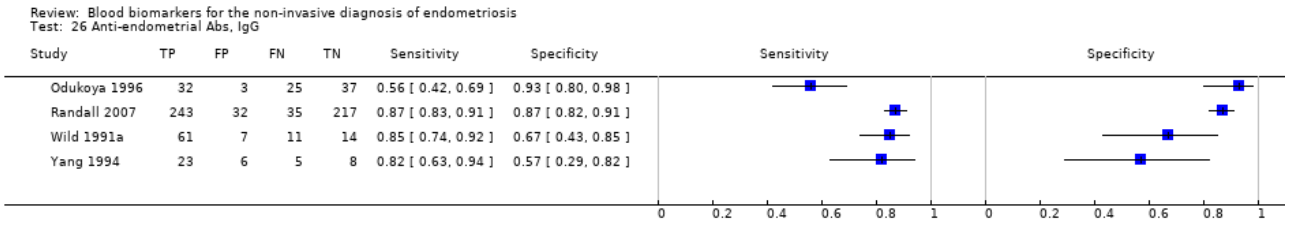
Test 24. Prolactin (> 14.8 ng/ml).



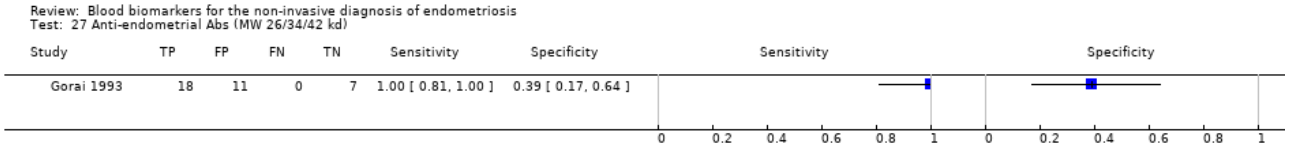
Test 25. Prolactin (> 20 ng/ml).



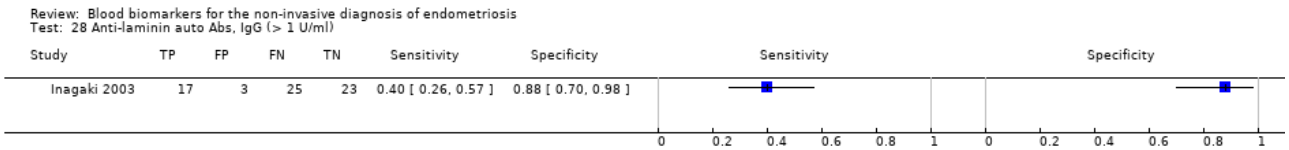
Test 26. Anti-endometrial Abs, IgG.



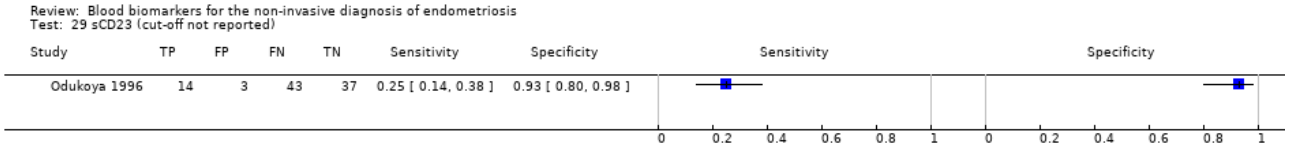
Test 27. Anti-endometrial Abs (MW 26/34/42 kd).



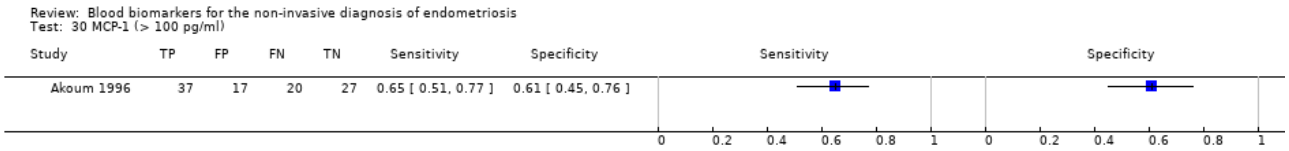
Test 28. Anti-laminin auto Abs, IgG (> 1 U/ml).



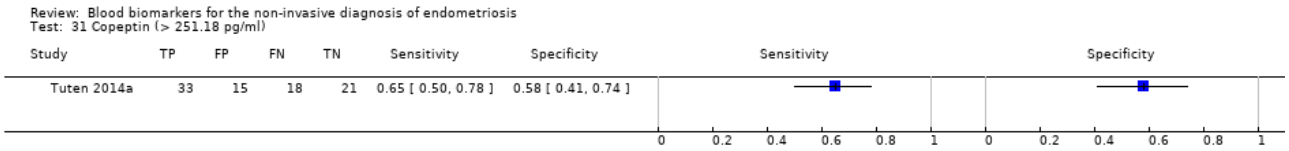
Test 29. sCD23 (cut-off not reported).



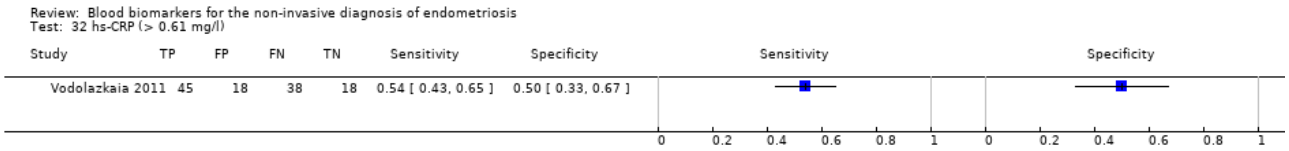
Test 30. MCP-1 (> 100 pg/ml).



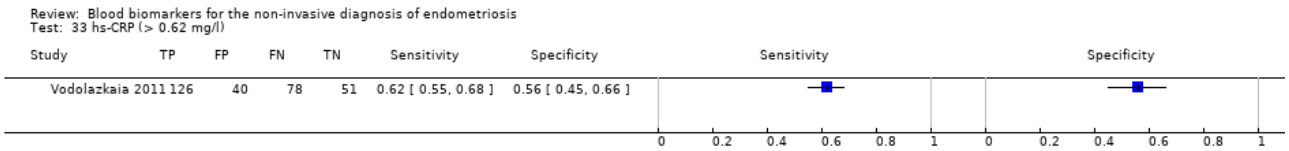
Test 31. Copeptin (> 251.18 pg/ml).



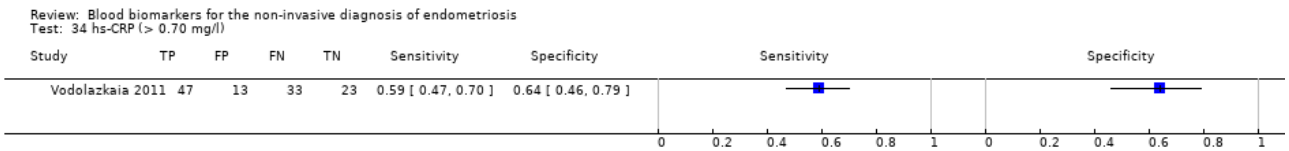
Test 32. hs-CRP (> 0.61 mg/l).



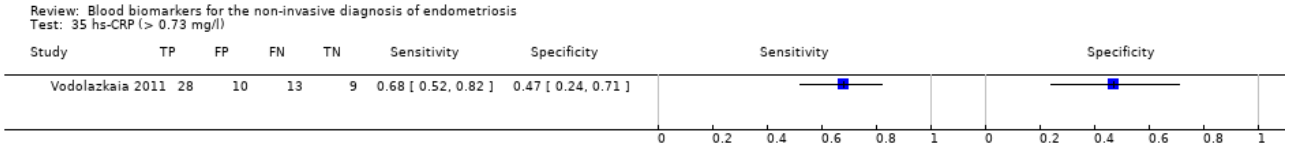
Test 33. hs-CRP (> 0.62 mg/l).



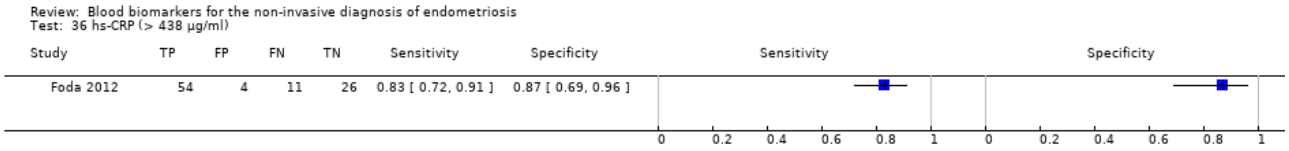
Test 34. hs-CRP (> 0.70 mg/l).



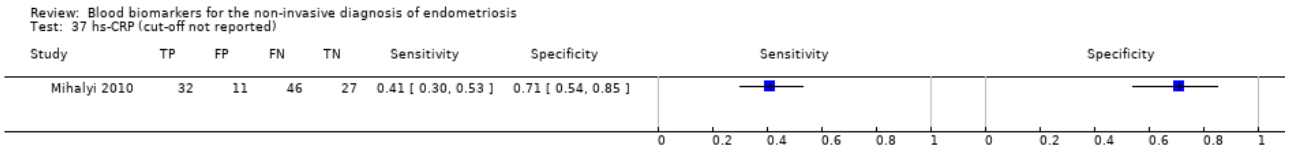
Test 35. hs-CRP (> 0.73 mg/l).



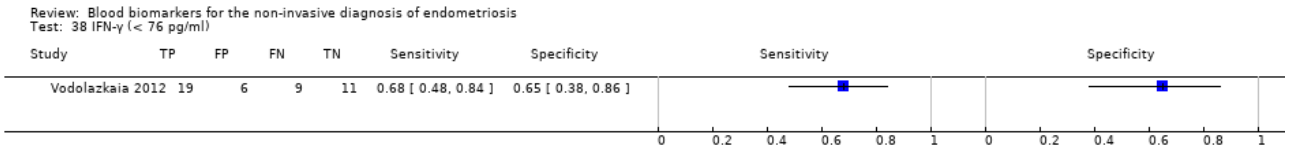
Test 36. hs-CRP (> 438 µg/ml).



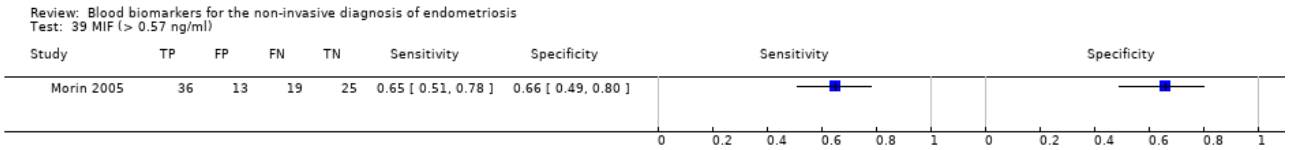
Test 37. hs-CRP (cut-off not reported).



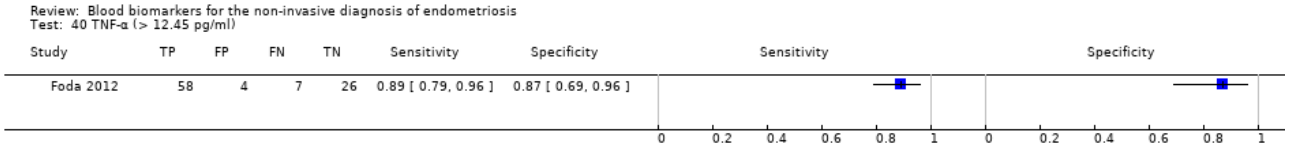
Test 38. IFN-γ (< 76 pg/ml).



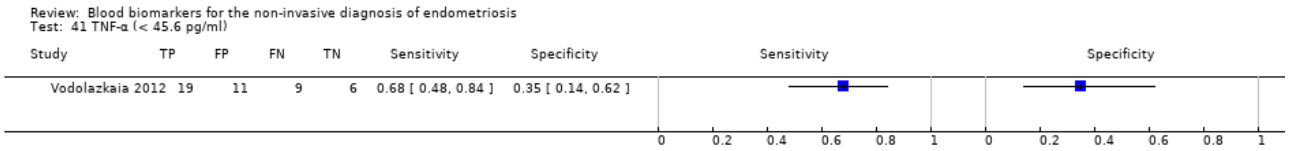
Test 39. MIF (> 0.57 ng/ml).



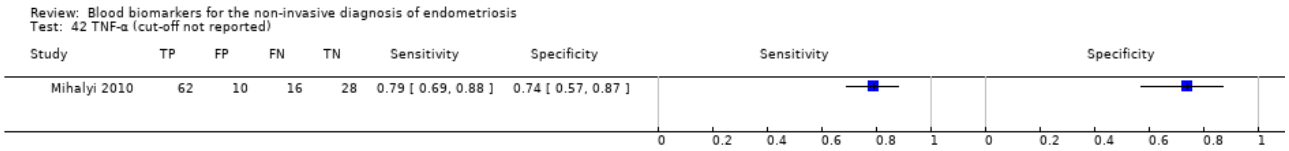
Test 40. TNF-α (> 12.45 pg/ml).



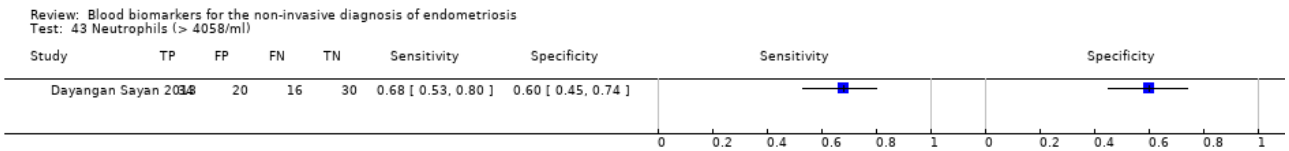
Test 41. TNF-α (< 45.6 pg/ml).



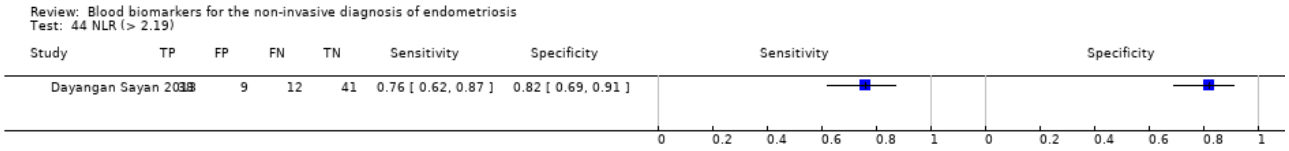
Test 42. TNF-α (cut-off not reported).



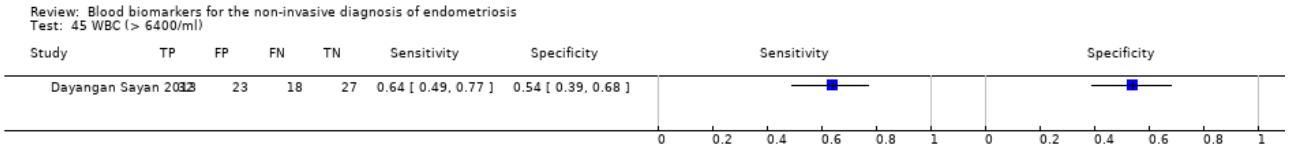
Test 43. Neutrophils (> 4058/ml).



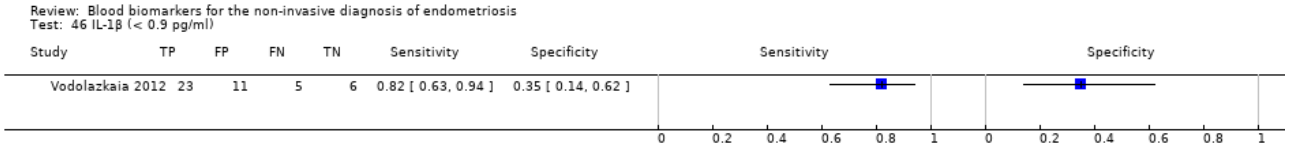
Test 44. NLR (> 2.19).



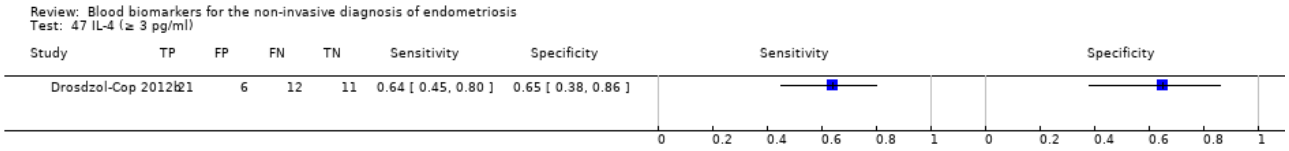
Test 45. WBC (> 6400/ml).



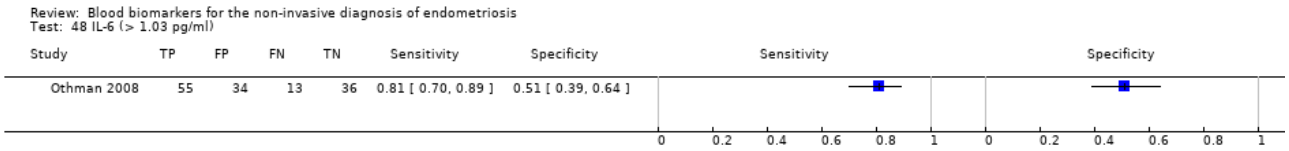
Test 46. IL-1β (< 0.9 pg/ml).



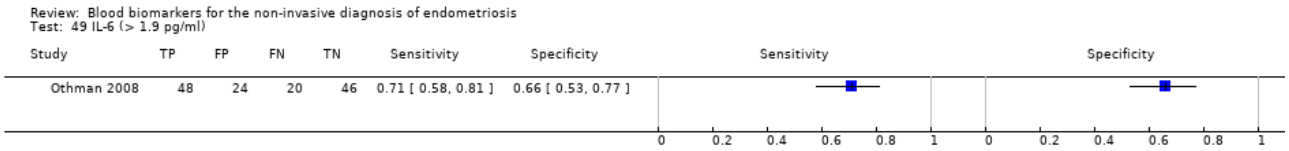
Test 47. IL-4 (≥ 3 pg/ml).



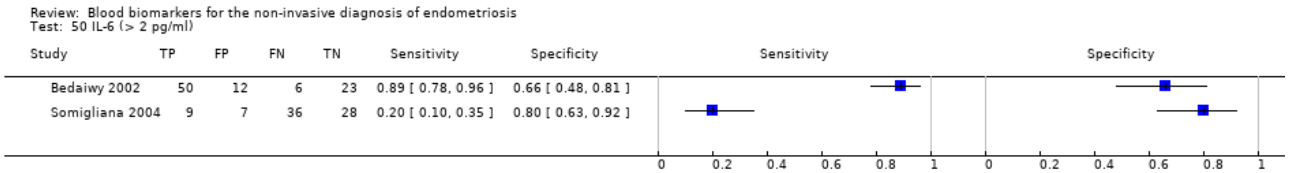
Test 48. IL-6 (> 1.03 pg/ml).



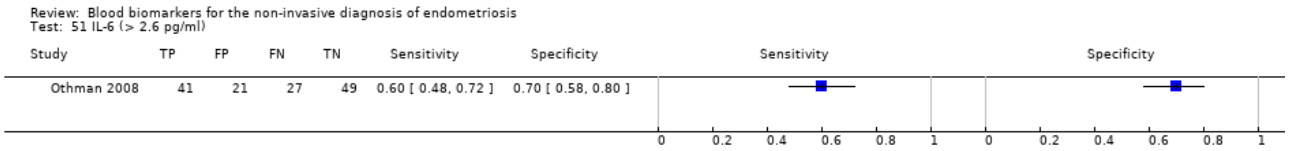
Test 49. IL-6 (> 1.9 pg/ml).



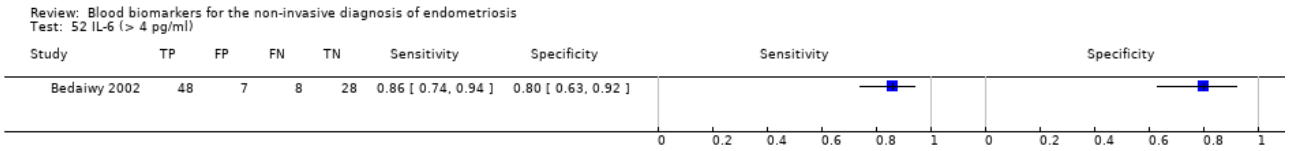
Test 50. IL-6 (> 2 pg/ml).



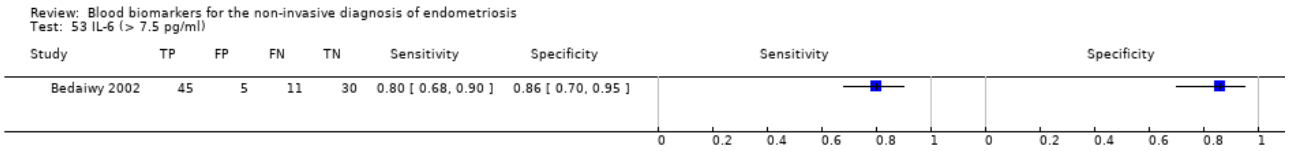
Test 51. IL-6 (> 2.6 pg/ml).



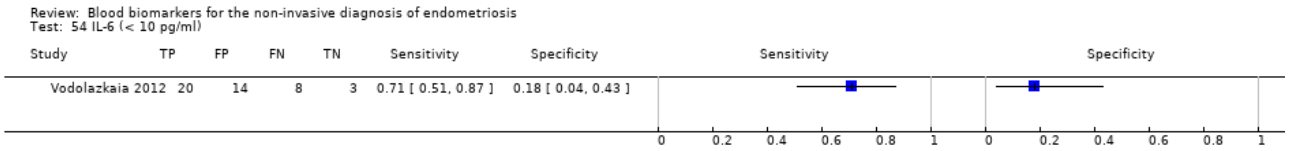
Test 52. IL-6 (> 4 pg/ml).



Test 53. IL-6 (> 7.5 pg/ml).

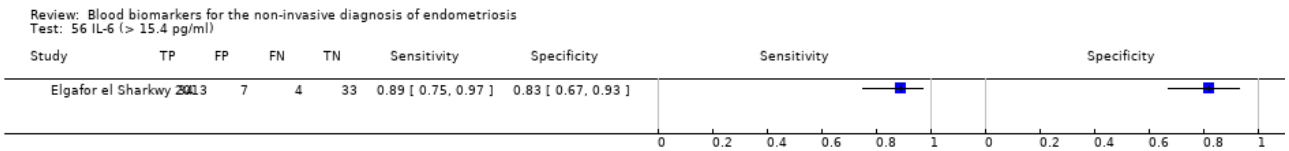


Test 54. IL-6 (< 10 pg/ml).

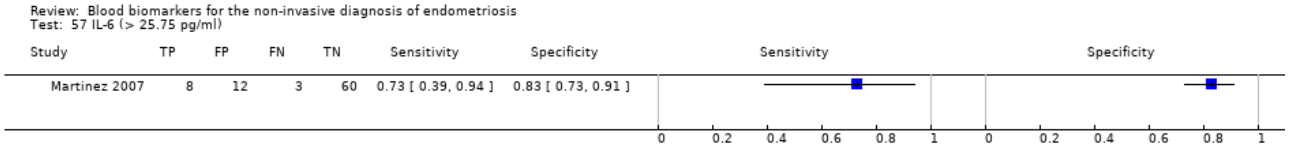


Test 55. IL-6 (> 12.2 pg/ml).

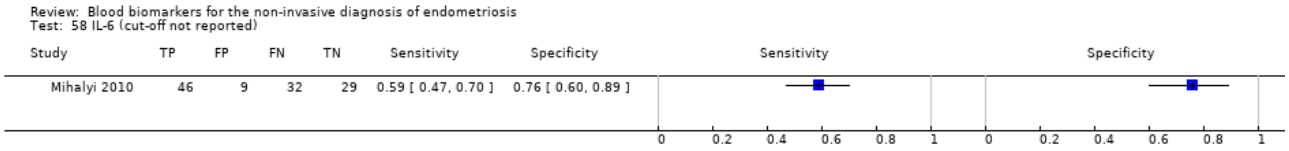
Test 56. IL-6 (> 15.4 pg/ml).



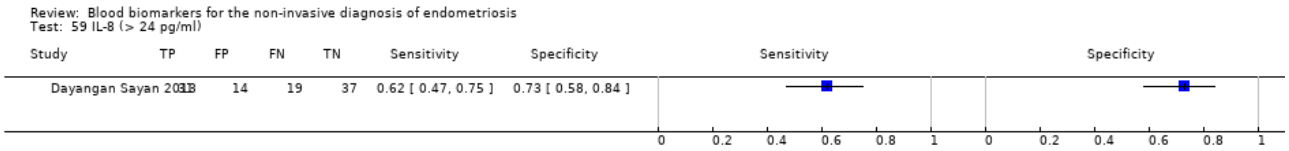
Test 57. IL-6 (> 25.75 pg/ml).



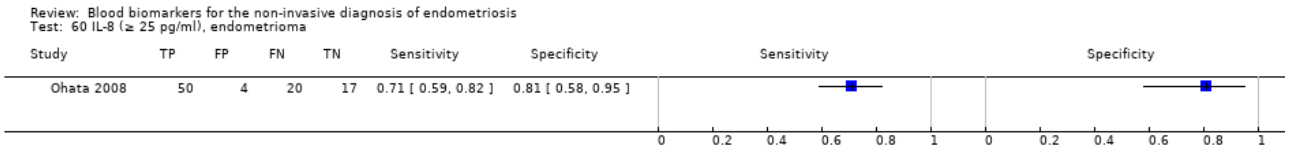
Test 58. IL-6 (cut-off not reported).



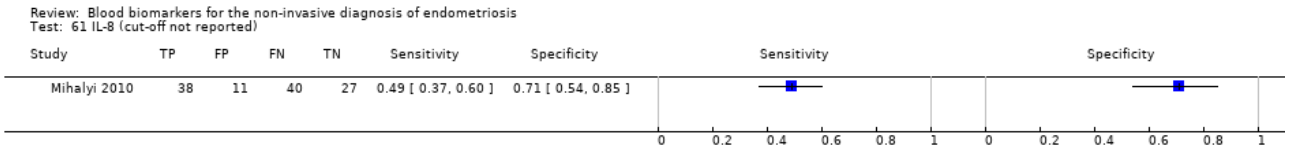
Test 59. IL-8 (> 24 pg/ml).



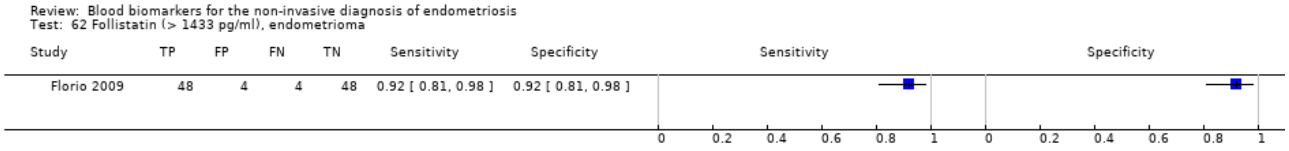
Test 60. IL-8 (≥ 25 pg/ml), endometrioma.



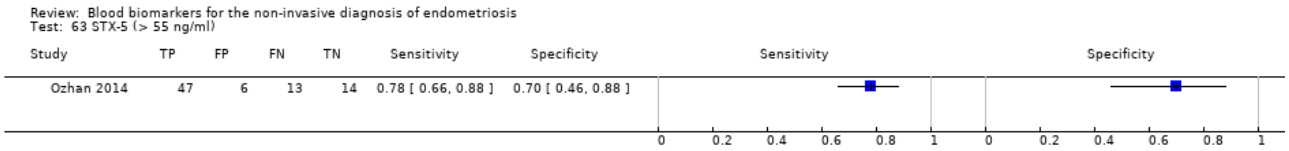
Test 61. IL-8 (cut-off not reported).



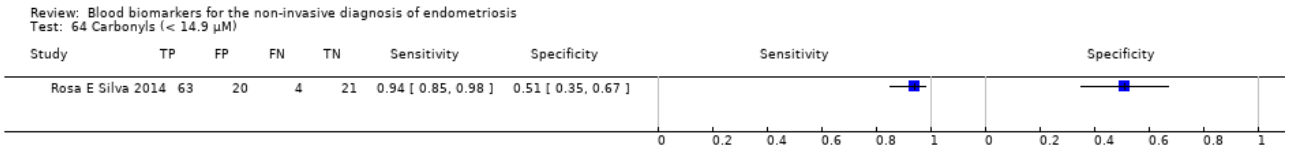
Test 62. Follistatin (> 1433 pg/ml), endometrioma.



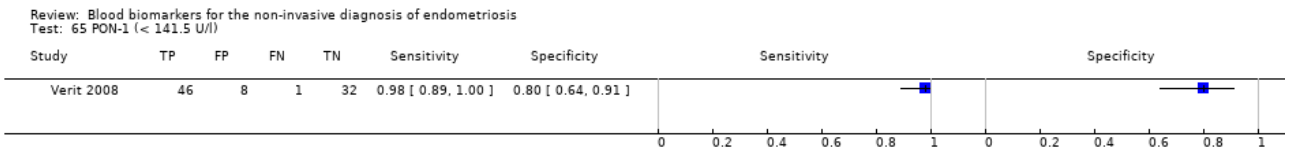
Test 63. STX-5 (> 55 ng/ml).



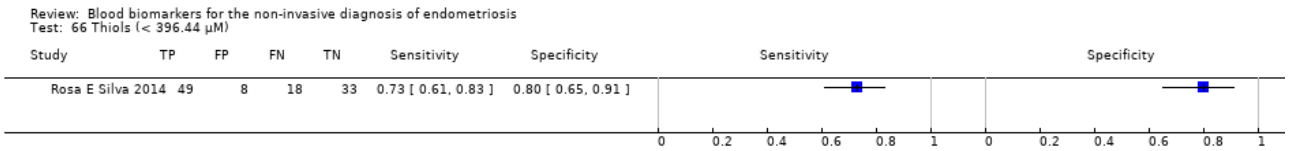
Test 64. Carbonyls (< 14.9 µM).



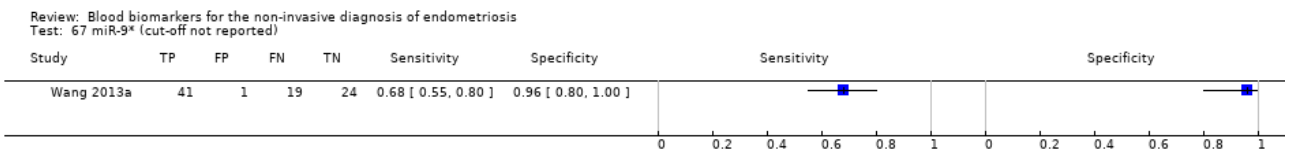
Test 65. PON-1 (< 141.5 U/l).



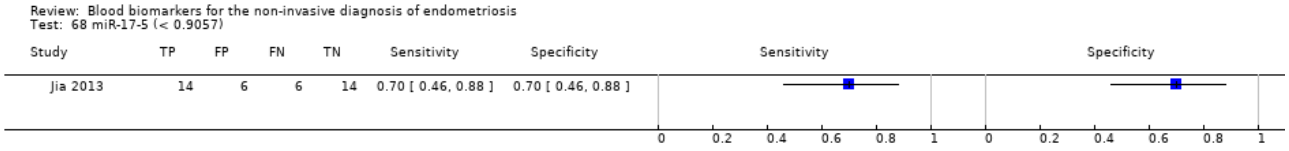
Test 66. Thiols (< 396.44 µM).



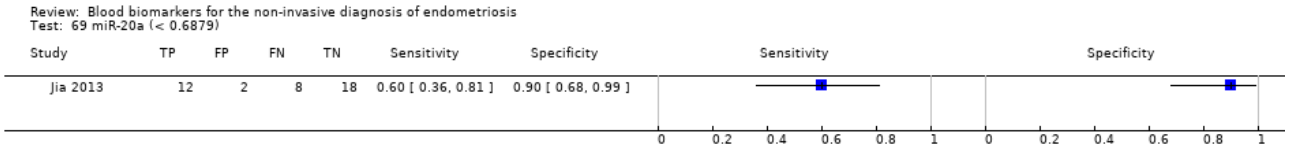
Test 67. miR-9* (cut-off not reported).



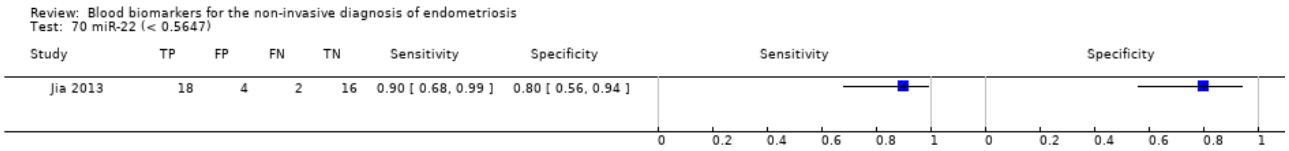
Test 68. miR-17-5 (< 0.9057).



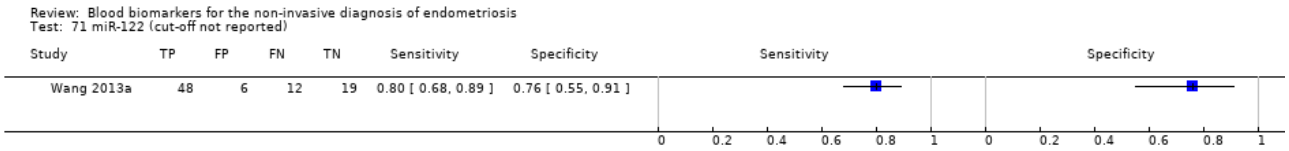
Test 69. miR-20a (< 0.6879).



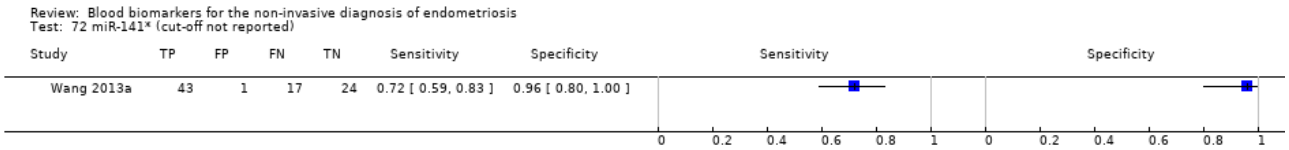
Test 70. miR-22 (< 0.5647).



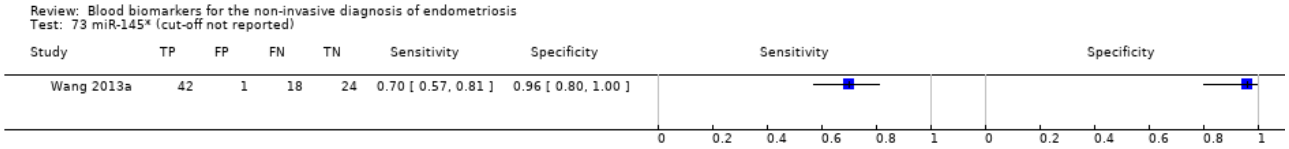
Test 71. miR-122 (cut-off not reported).



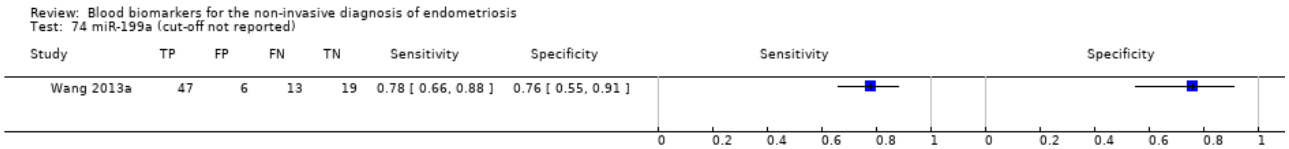
Test 72. miR-141* (cut-off not reported).



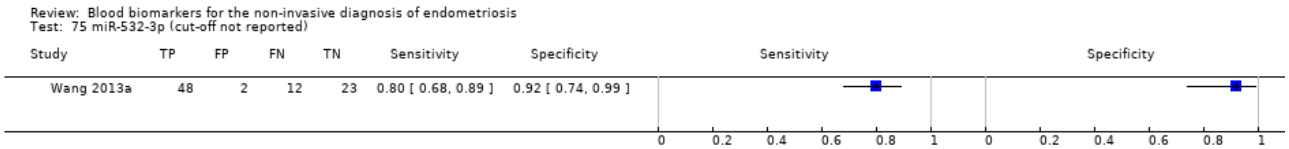
Test 73. miR-145* (cut-off not reported).



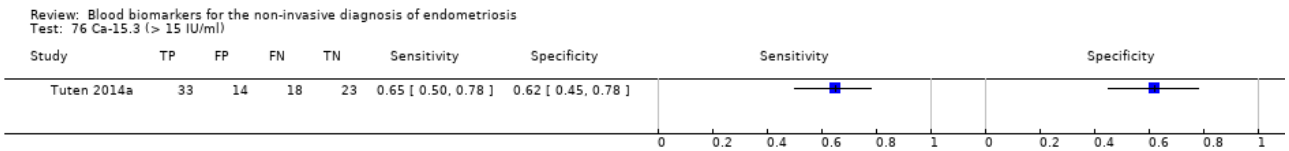
Test 74. miR-199a (cut-off not reported).



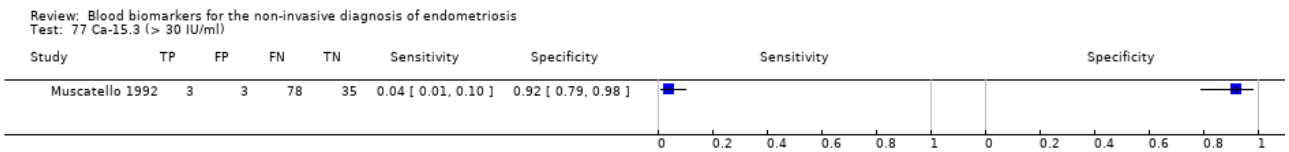
Test 75. miR-532-3p (cut-off not reported).



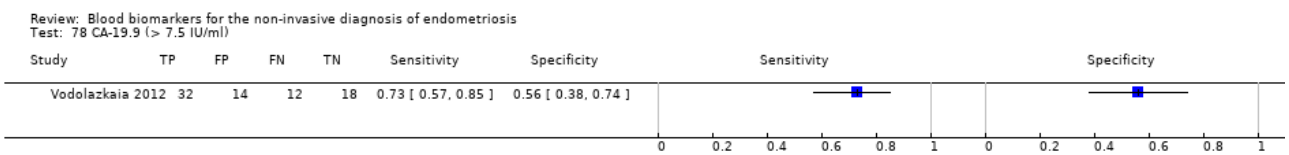
Test 76. Ca-15.3 (> 15 IU/ml).



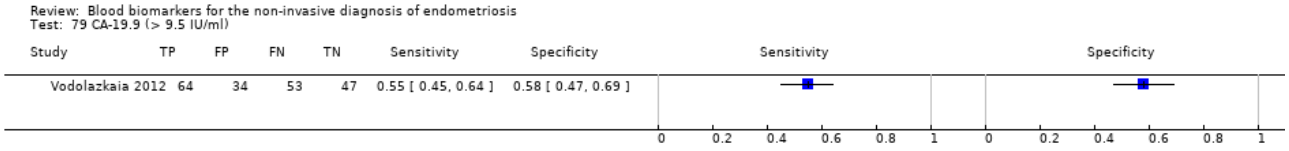
Test 77. Ca-15.3 (> 30 IU/ml).



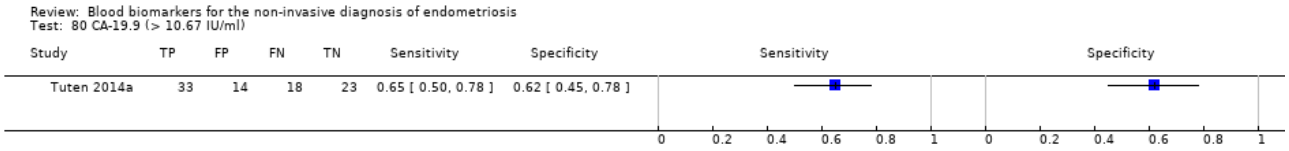
Test 78. CA-19.9 (> 7.5 IU/ml).



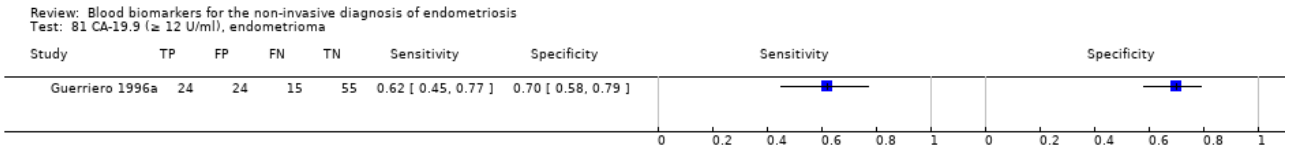
Test 79. CA-19.9 (> 9.5 IU/ml).



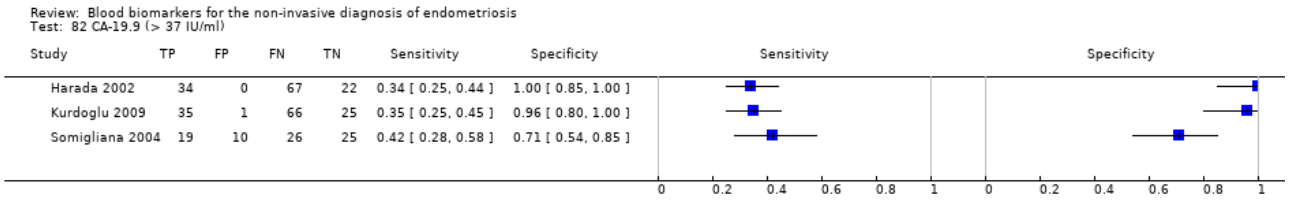
Test 80. CA-19.9 (> 10.67 IU/ml).



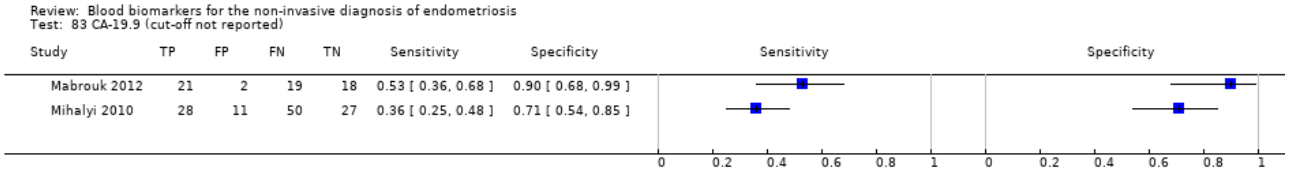
Test 81. CA-19.9 (≥ 12 U/ml), endometrioma.



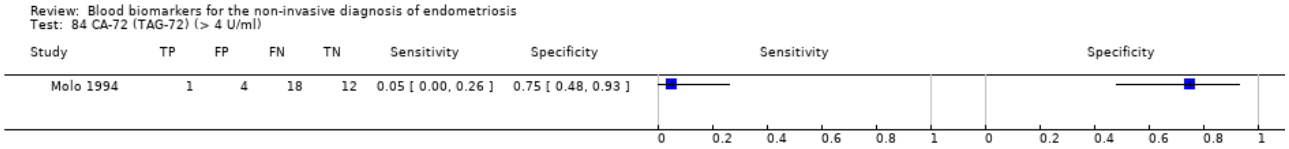
Test 82. CA-19.9 (> 37 IU/ml).



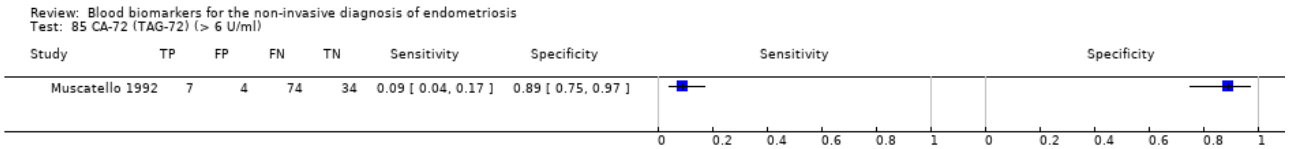
Test 83. CA-19.9 (cut-off not reported).



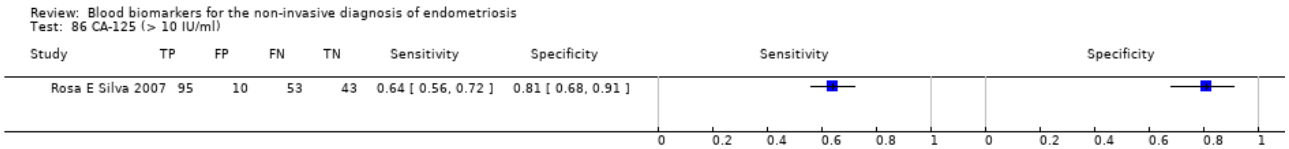
Test 84. CA-72 (TAG-72) (> 4 U/ml).



Test 85. CA-72 (TAG-72) (> 6 U/ml).

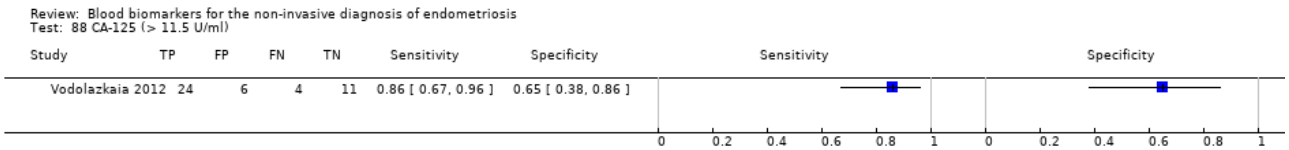


Test 86. CA-125 (> 10 IU/ml).

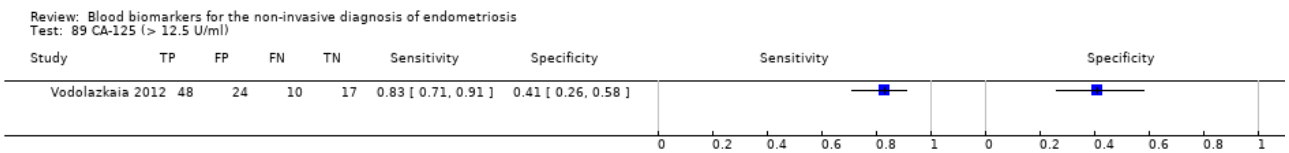


Test 87. CA-125 (> 11 U/ml).

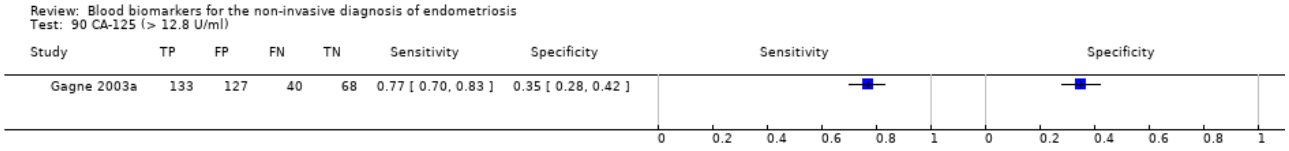
Test 88. CA-125 (> 11.5 U/ml).



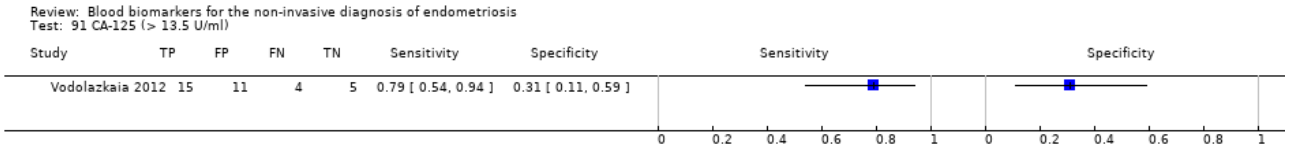
Test 89. CA-125 (> 12.5 U/ml).



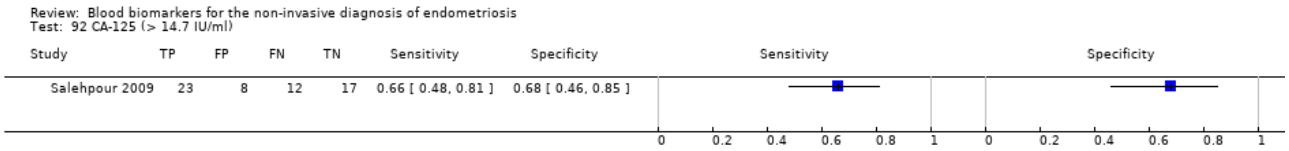
Test 90. CA-125 (> 12.8 U/ml).



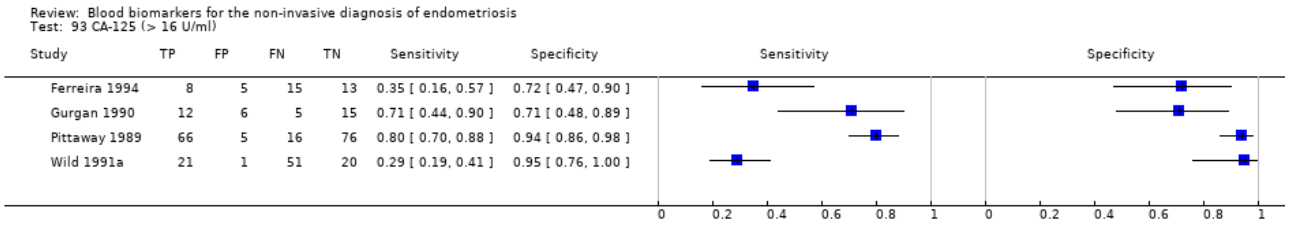
Test 91. CA-125 (> 13.5 U/ml).



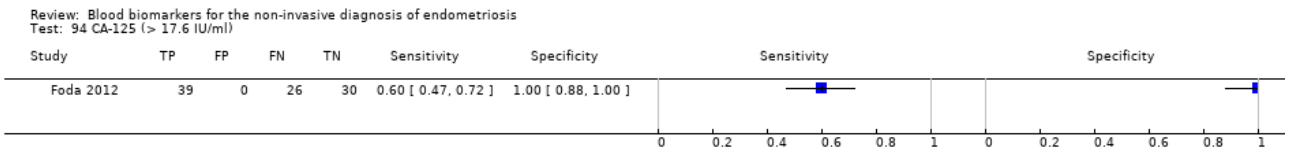
Test 92. CA-125 (> 14.7 IU/ml).



Test 93. CA-125 (> 16 U/ml).



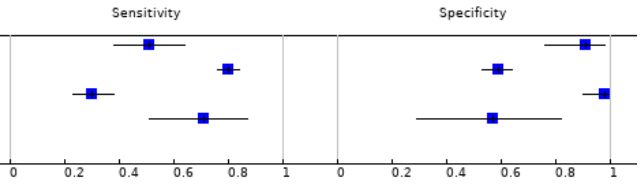
Test 94. CA-125 (> 17.6 IU/ml).



Test 95. CA-125 (> 20 IU/ml).

Review: Blood biomarkers for the non-invasive diagnosis of endometriosis
Test: 95 CA-125 (> 20 IU/ml)

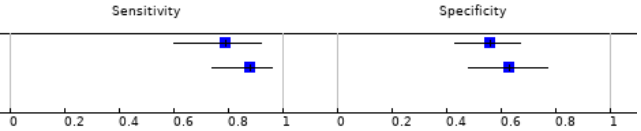
Study	TP	FP	FN	TN	Sensitivity	Specificity
Bilibio 2014	32	3	31	31	0.51 [0.38, 0.64]	0.91 [0.76, 0.98]
Kitawaki 2005	347	141	86	201	0.80 [0.76, 0.84]	0.59 [0.53, 0.64]
Rosa E Silva 2007	45	1	103	52	0.30 [0.23, 0.38]	0.98 [0.90, 1.00]
Yang 1994	20	6	8	8	0.71 [0.51, 0.87]	0.57 [0.29, 0.82]



Test 96. CA-125 (> 20 U/ml), endometrioma.

Review: Blood biomarkers for the non-invasive diagnosis of endometriosis
Test: 96 CA-125 (> 20 U/ml), endometrioma

Study	TP	FP	FN	TN	Sensitivity	Specificity
Guerriero 1996b	23	32	6	40	0.79 [0.60, 0.92]	0.56 [0.43, 0.67]
Tokmak 2011	37	17	5	29	0.88 [0.74, 0.96]	0.63 [0.48, 0.77]



Test 97. CA-125 (> 25 U/ml), endometrioma.

Review: Blood biomarkers for the non-invasive diagnosis of endometriosis
Test: 97 CA-125 (> 25 U/ml), endometrioma

Study	TP	FP	FN	TN	Sensitivity	Specificity
Guerriero 1996b	22	24	7	48	0.76 [0.56, 0.90]	0.67 [0.55, 0.77]

Test 98. CA-125 (> 26 IU/ml).

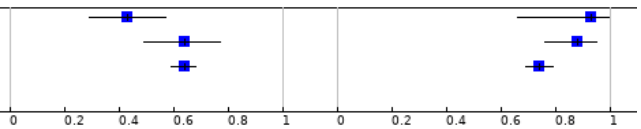
Review: Blood biomarkers for the non-invasive diagnosis of endometriosis
Test: 98 CA-125 (> 26 IU/ml)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Kitawaki 2005	307	108	126	234	0.71 [0.66, 0.75]	0.68 [0.63, 0.73]
Tuten 2014a	44	5	6	32	0.88 [0.76, 0.95]	0.86 [0.71, 0.95]

Test 99. CA-125 (> 30 U/ml).

Review: Blood biomarkers for the non-invasive diagnosis of endometriosis
Test: 99 CA-125 (> 30 U/ml)

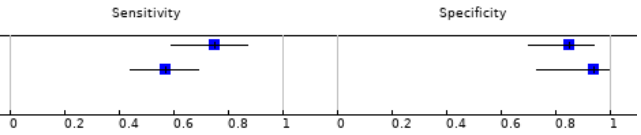
Study	TP	FP	FN	TN	Sensitivity	Specificity
Acien 1989	23	1	31	13	0.43 [0.29, 0.57]	0.93 [0.66, 1.00]
Dayangan Sayan 2013	113	6	18	44	0.64 [0.49, 0.77]	0.88 [0.76, 0.95]
Kitawaki 2005	275	89	158	253	0.64 [0.59, 0.68]	0.74 [0.69, 0.79]



Test 100. CA-125 (> 30 U/ml), endometrioma.

Review: Blood biomarkers for the non-invasive diagnosis of endometriosis
Test: 100 CA-125 (> 30 U/ml), endometrioma

Study	TP	FP	FN	TN	Sensitivity	Specificity
Florio 2007	30	6	10	34	0.75 [0.59, 0.87]	0.85 [0.70, 0.94]
Ohata 2008	37	1	28	17	0.57 [0.44, 0.69]	0.94 [0.73, 1.00]



Test 101. CA-125 (> 33 U/ml).

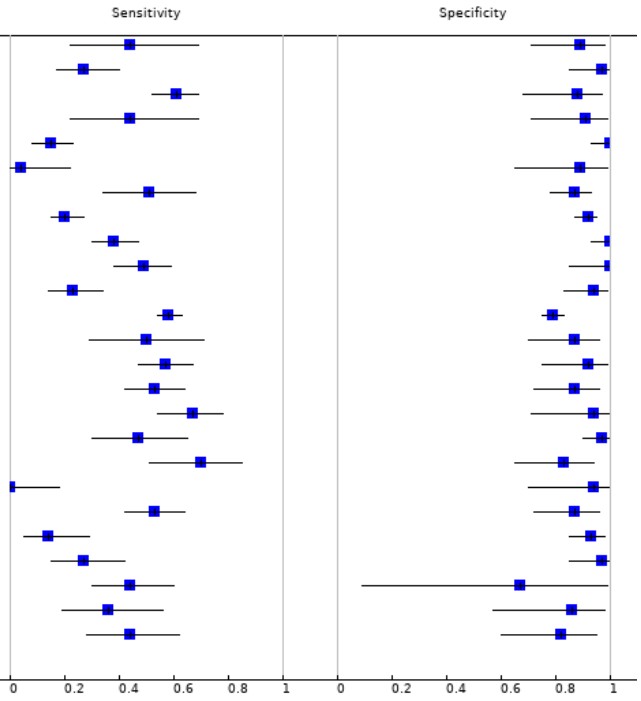
Review: Blood biomarkers for the non-invasive diagnosis of endometriosis
Test: 101 CA-125 (> 33 U/ml)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Matalliotakis 2004	20	0	30	50	0.40 [0.26, 0.55]	1.00 [0.93, 1.00]

Test 102. CA-125 (> 35 U/ml).

Review: Blood biomarkers for the non-invasive diagnosis of endometriosis
Test: 102 CA-125 (> 35 U/ml)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Barbati 1994	8	3	10	24	0.44 [0.22, 0.69]	0.89 [0.71, 0.98]
Bilibio 2014	17	1	46	33	0.27 [0.17, 0.40]	0.97 [0.85, 1.00]
Chen 1998	80	3	51	21	0.61 [0.52, 0.69]	0.88 [0.68, 0.97]
Colacurci 1996a	8	2	10	20	0.44 [0.22, 0.69]	0.91 [0.71, 0.99]
Fedele 1989	15	0	87	52	0.15 [0.08, 0.23]	1.00 [0.93, 1.00]
Ferreira 1994	1	2	22	16	0.04 [0.00, 0.22]	0.89 [0.65, 0.99]
Franchi 1993	19	11	18	72	0.51 [0.34, 0.68]	0.87 [0.78, 0.93]
Gagne 2003a	35	16	138	179	0.20 [0.15, 0.27]	0.92 [0.87, 0.95]
Hallamaa 2012	47	0	76	52	0.38 [0.30, 0.47]	1.00 [0.93, 1.00]
Harada 2002	49	0	52	22	0.49 [0.38, 0.59]	1.00 [0.85, 1.00]
Hornstein 1995	17	3	57	46	0.23 [0.14, 0.34]	0.94 [0.83, 0.99]
Kitawaki 2005	253	71	180	271	0.58 [0.54, 0.63]	0.79 [0.75, 0.83]
Koninckx 1996	12	4	12	27	0.50 [0.29, 0.71]	0.87 [0.70, 0.96]
Kurdoglu 2009	58	2	43	24	0.57 [0.47, 0.67]	0.92 [0.75, 0.99]
Lanzone 1991	43	5	38	33	0.53 [0.42, 0.64]	0.87 [0.72, 0.96]
Maiorana 2007	46	1	23	16	0.67 [0.54, 0.78]	0.94 [0.71, 1.00]
Martinez 2007	17	2	19	70	0.47 [0.30, 0.65]	0.97 [0.90, 1.00]
Mohamed 2013	21	5	9	25	0.70 [0.51, 0.85]	0.83 [0.65, 0.94]
Molo 1994	0	1	19	15	0.0 [0.0, 0.18]	0.94 [0.70, 1.00]
Muscatello 1992	43	5	38	33	0.53 [0.42, 0.64]	0.87 [0.72, 0.96]
Patton 1986	5	5	32	71	0.14 [0.05, 0.29]	0.93 [0.85, 0.98]
Somigliana 2004	12	1	33	34	0.27 [0.15, 0.42]	0.97 [0.85, 1.00]
Vigil 1999	20	1	25	2	0.44 [0.30, 0.60]	0.67 [0.09, 0.99]
Yang 1994	10	2	18	12	0.36 [0.19, 0.56]	0.86 [0.57, 0.98]
Zeng 2005	16	4	20	18	0.44 [0.28, 0.62]	0.82 [0.60, 0.95]

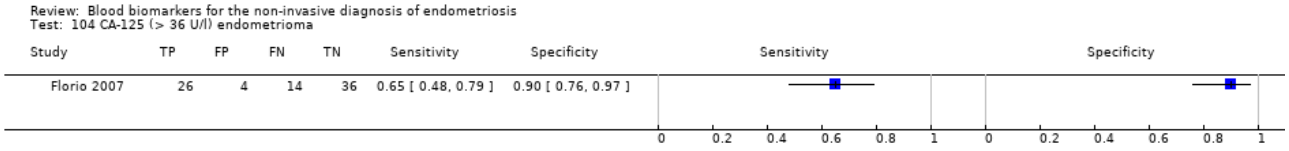


Test 103. CA-125 (> 35 U/ml), endometrioma.

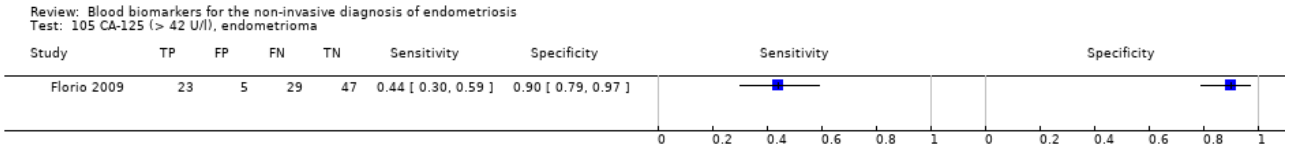
Review: Blood biomarkers for the non-invasive diagnosis of endometriosis
Test: 103 CA-125 (> 35 U/ml), endometrioma

Study	TP	FP	FN	TN	Sensitivity	Specificity
Guerriero 1996b	17	15	12	57	0.59 [0.39, 0.76]	0.79 [0.68, 0.88]

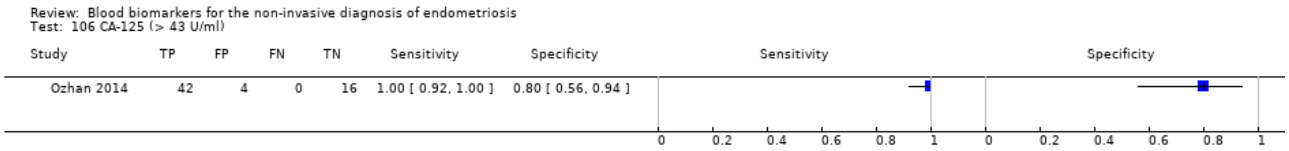
Test 104. CA-125 (> 36 U/l) endometrioma.



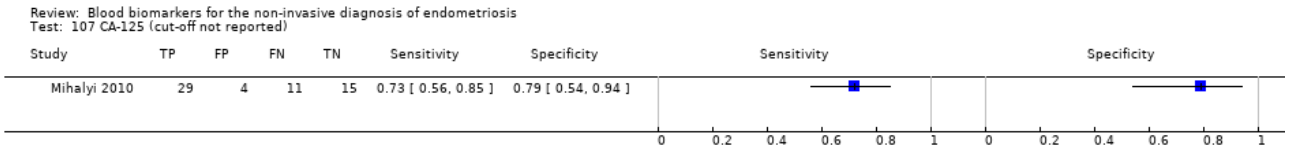
Test 105. CA-125 (> 42 U/l), endometrioma.



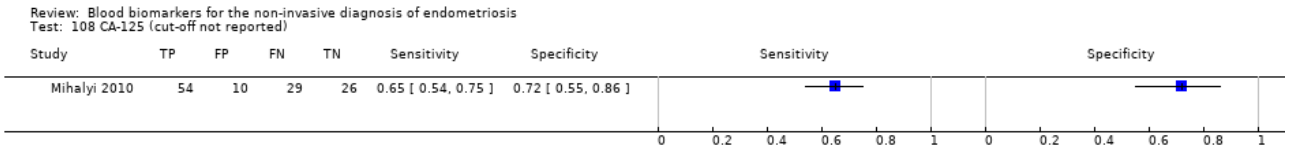
Test 106. CA-125 (> 43 U/ml).



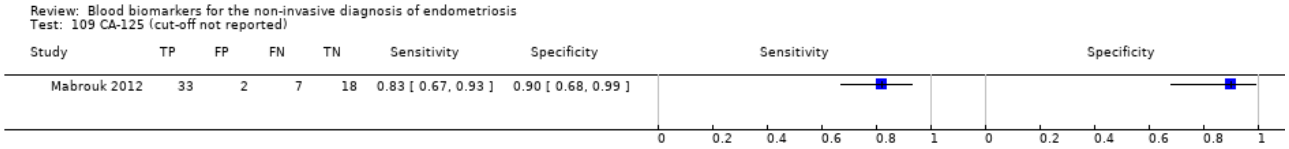
Test 107. CA-125 (cut-off not reported).



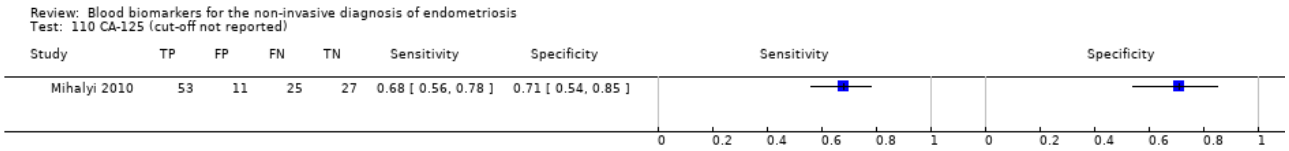
Test 108. CA-125 (cut-off not reported).



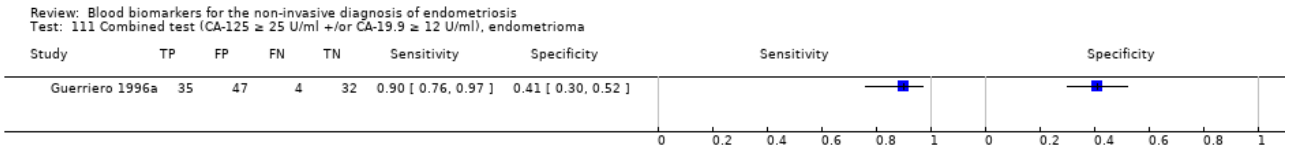
Test 109. CA-125 (cut-off not reported).



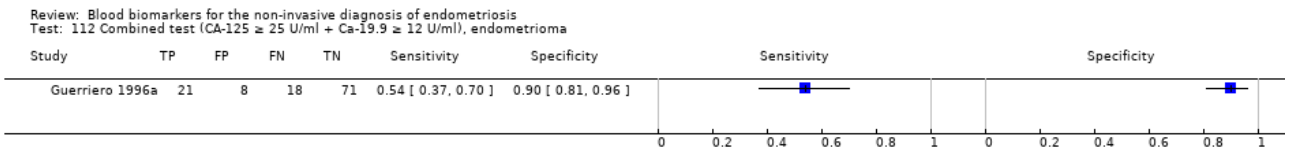
Test 110. CA-125 (cut-off not reported).



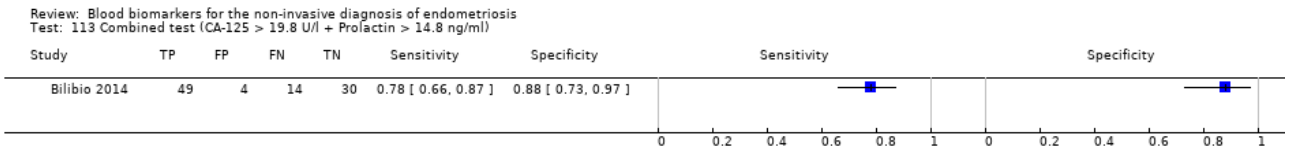
Test 111. Combined test (CA-125 ≥ 25 U/ml +/or CA-19.9 ≥ 12 U/ml), endometrioma.



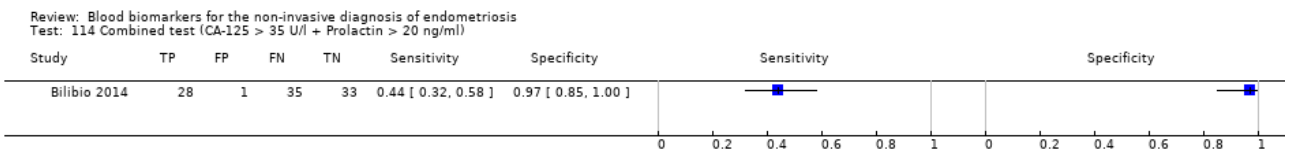
Test 112. Combined test (CA-125 ≥ 25 U/ml + Ca-19.9 ≥ 12 U/ml), endometrioma.



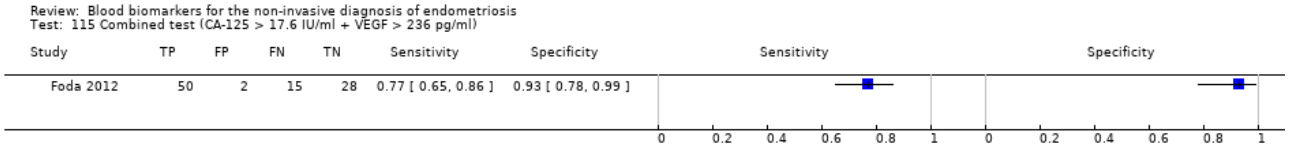
Test 113. Combined test (CA-125 > 19.8 U/l + Prolactin > 14.8 ng/ml).



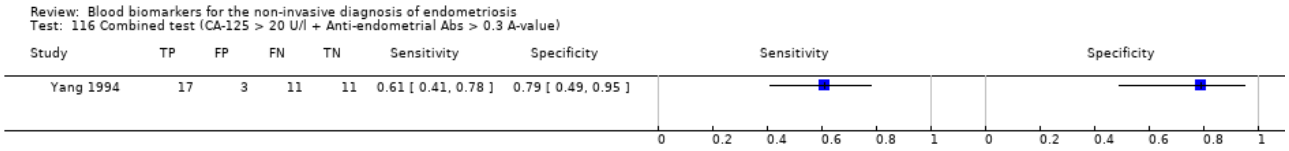
Test 114. Combined test (CA-125 > 35 U/l + Prolactin > 20 ng/ml).



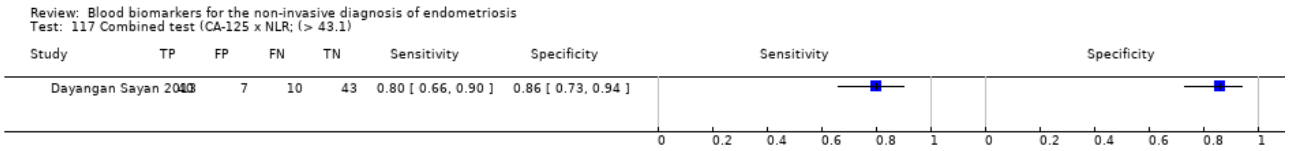
Test 115. Combined test (CA-125 > 17.6 IU/ml + VEGF > 236 pg/ml).



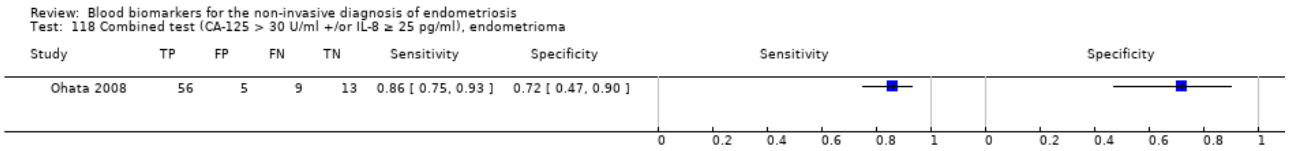
Test 116. Combined test (CA-125 > 20 U/l + Anti-endometrial Abs > 0.3 A-value).



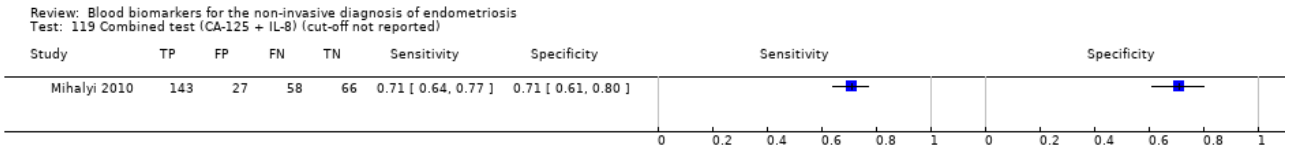
Test 117. Combined test (CA-125 x NLR; (> 43.1).



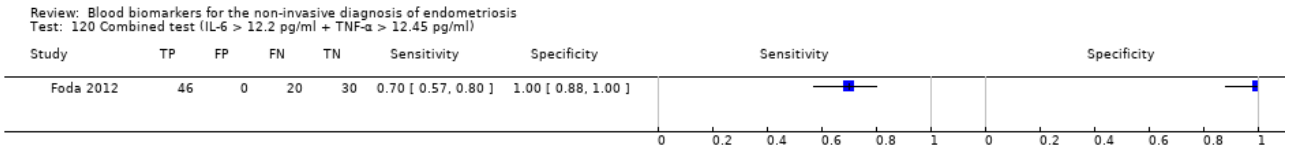
Test 118. Combined test (CA-125 > 30 U/ml +/or IL-8 ≥ 25 pg/ml), endometrioma.



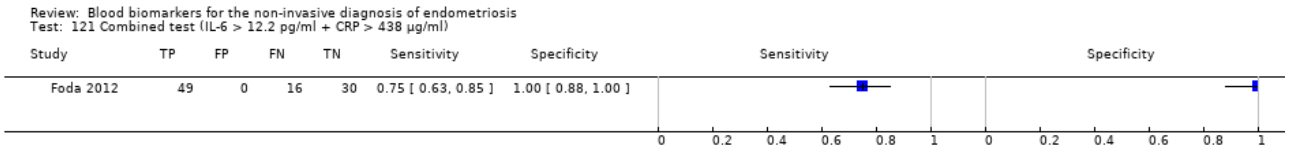
Test 119. Combined test (CA-125 + IL-8) (cut-off not reported).



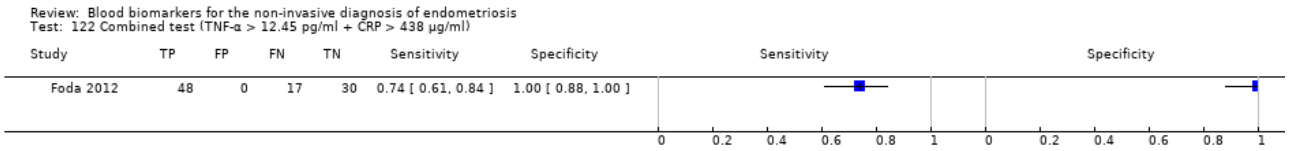
Test 120. Combined test (IL-6 > 12.2 pg/ml + TNF-α > 12.45 pg/ml).



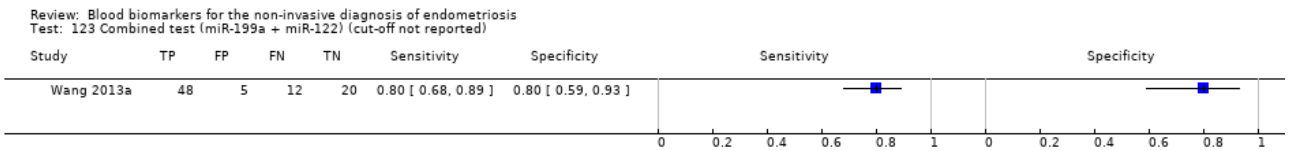
Test 121. Combined test (IL-6 > 12.2 pg/ml + CRP > 438 µg/ml).



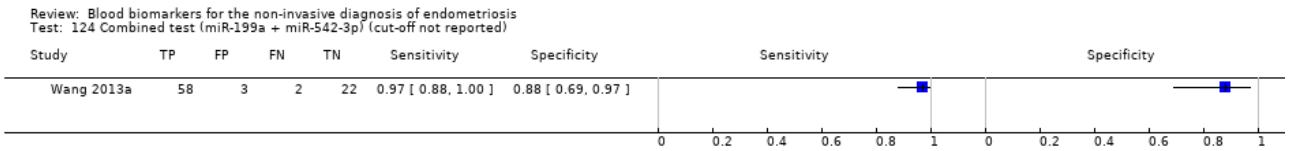
Test 122. Combined test (TNF-α > 12.45 pg/ml + CRP > 438 µg/ml).



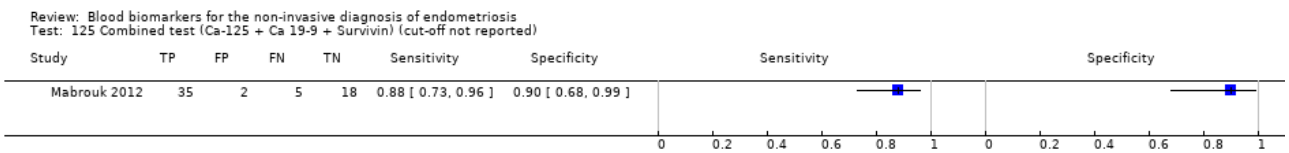
Test 123. Combined test (miR-199a + miR-122) (cut-off not reported).



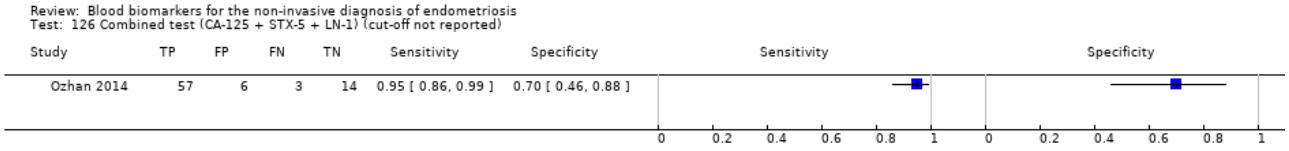
Test 124. Combined test (miR-199a + miR-542-3p) (cut-off not reported).



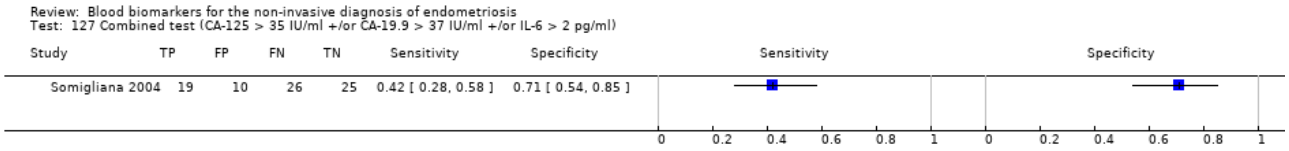
Test 125. Combined test (Ca-125 + Ca 19-9 + Survivin) (cut-off not reported).



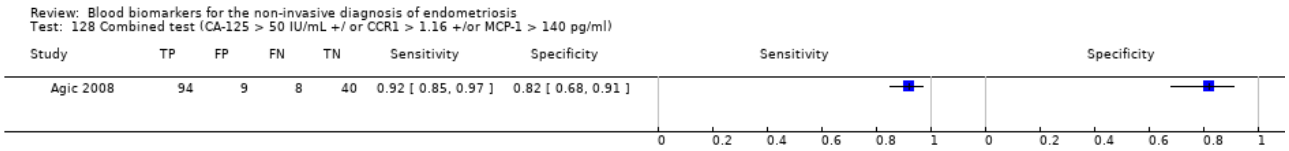
Test 126. Combined test (CA-125 + STX-5 + LN-1) (cut-off not reported).



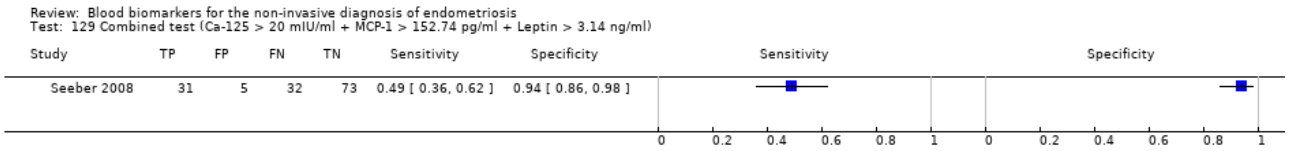
Test 127. Combined test (CA-125 > 35 IU/ml +/or CA-19.9 > 37 IU/ml +/or IL-6 > 2 pg/ml).



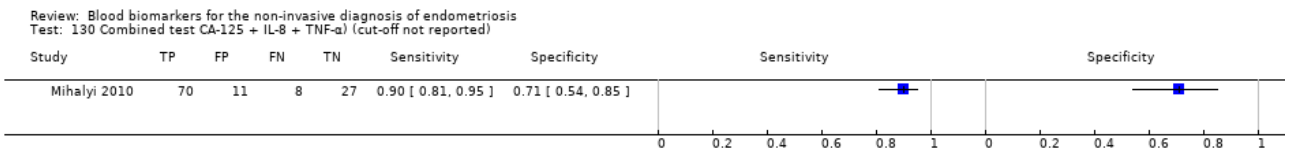
Test 128. Combined test (CA-125 > 50 IU/mL +/ or CCR1 > 1.16 +/or MCP-1 > 140 pg/ml).



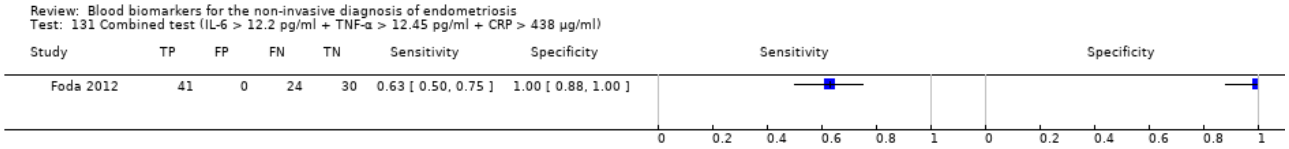
Test 129. Combined test (Ca-125 > 20 mIU/ml + MCP-1 > 152.74 pg/ml + Leptin > 3.14 ng/ml).



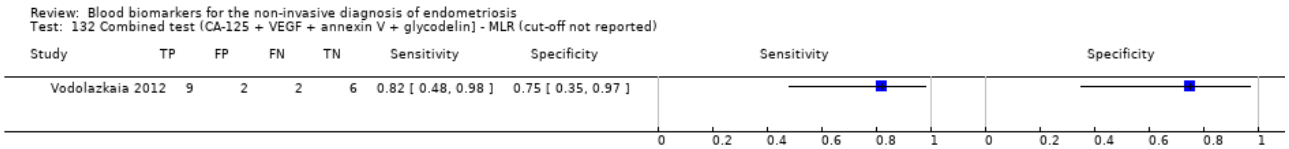
Test 130. Combined test CA-125 + IL-8 + TNF-α) (cut-off not reported).



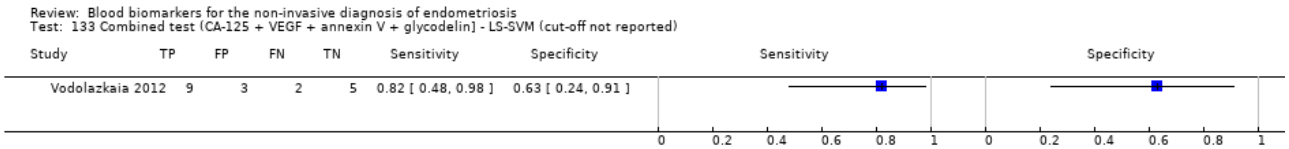
Test 131. Combined test (IL-6 > 12.2 pg/ml + TNF-α > 12.45 pg/ml + CRP > 438 µg/ml).



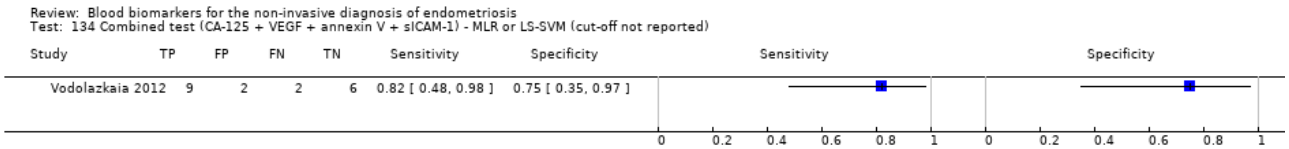
Test 132. Combined test (CA-125 + VEGF + annexin V + glycodelin] - MLR (cut-off not reported).



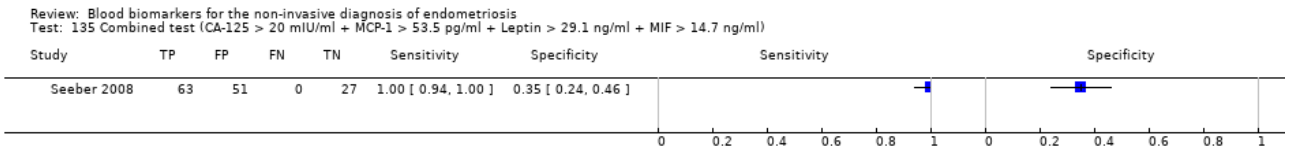
Test 133. Combined test (CA-125 + VEGF + annexin V + glycodelin] - LS-SVM (cut-off not reported).



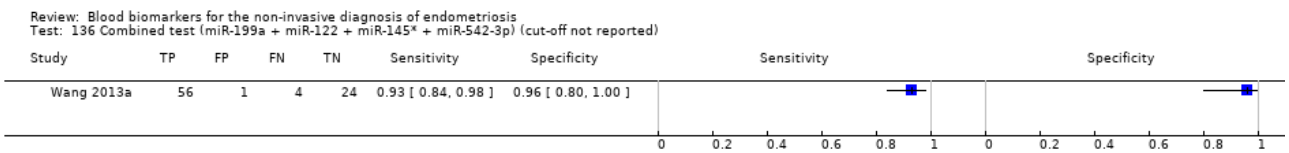
Test 134. Combined test (CA-125 + VEGF + annexin V + sICAM-1) - MLR or LS-SVM (cut-off not reported).



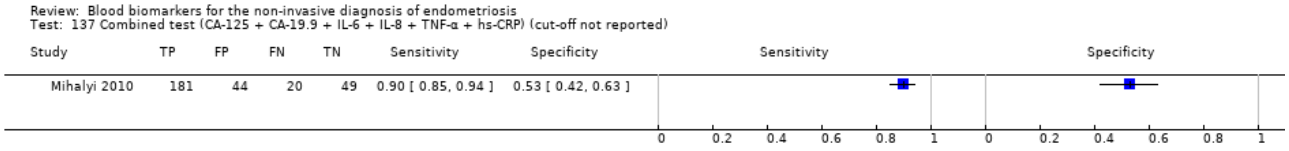
Test 135. Combined test (CA-125 > 20 mIU/ml + MCP-1 > 53.5 pg/ml + Leptin > 29.1 ng/ml + MIF > 14.7 ng/ml).



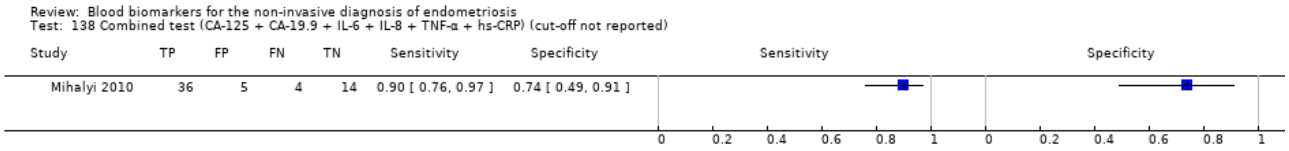
Test 136. Combined test (miR-199a + miR-122 + miR-145* + miR-542-3p) (cut-off not reported).



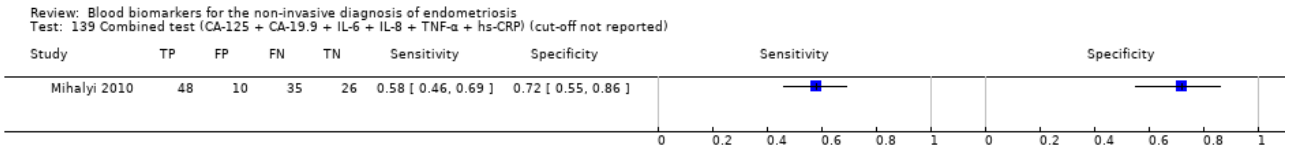
Test 137. Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported).



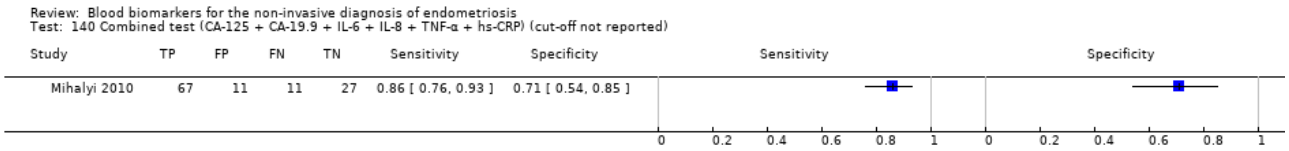
Test 138. Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported).



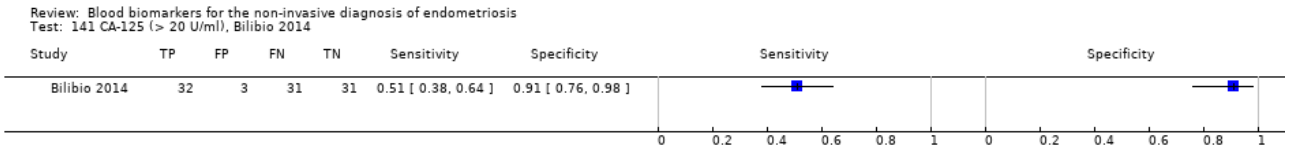
Test 139. Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported).



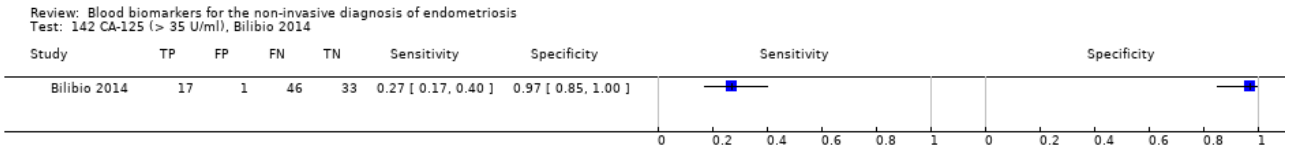
Test 140. Combined test (CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP) (cut-off not reported).



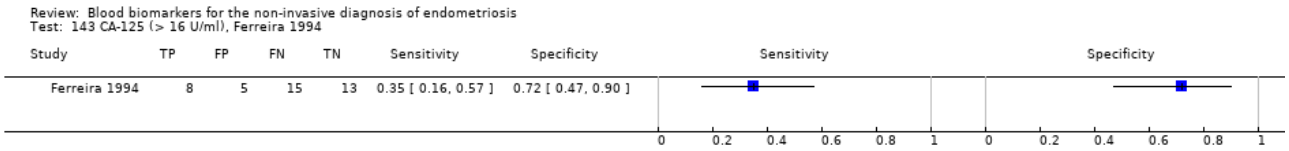
Test 141. CA-125 (> 20 U/ml), Bilibio 2014.



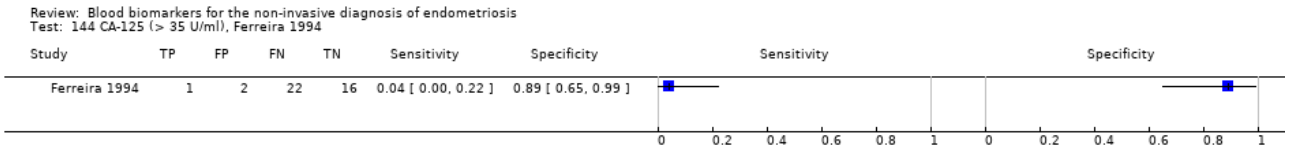
Test 142. CA-125 (> 35 U/ml), Bilibio 2014.



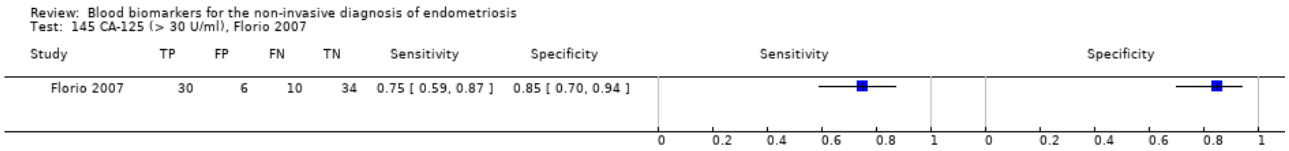
Test 143. CA-125 (> 16 U/ml), Ferreira 1994.



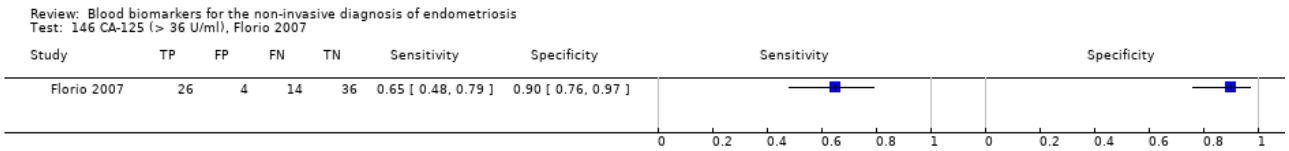
Test 144. CA-125 (> 35 U/ml), Ferreira 1994.



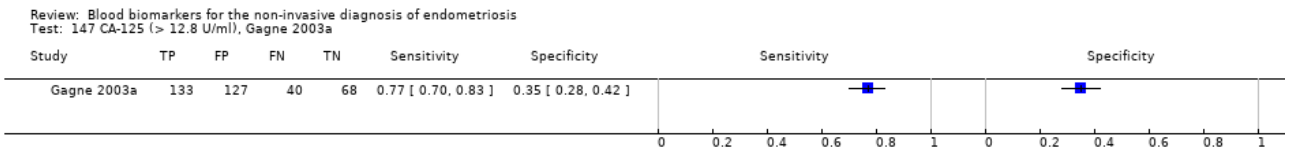
Test 145. CA-125 (> 30 U/ml), Florio 2007.



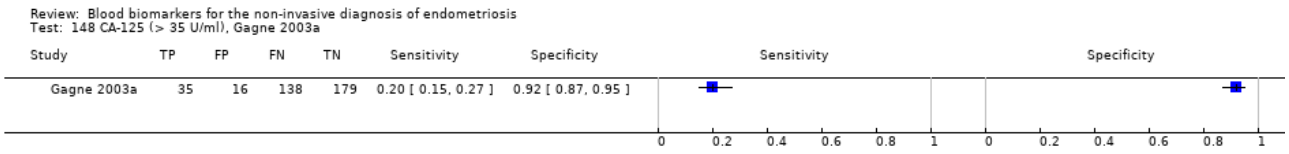
Test 146. CA-125 (> 36 U/ml), Florio 2007.



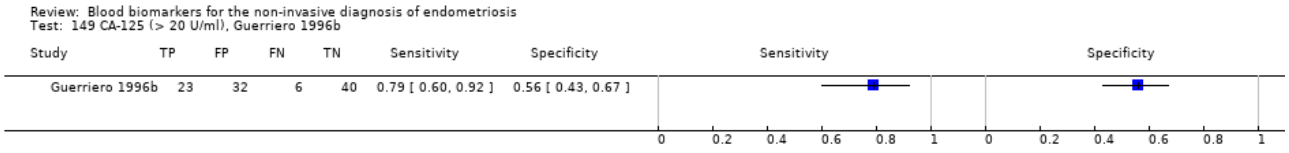
Test 147. CA-125 (> 12.8 U/ml), Gagne 2003a.



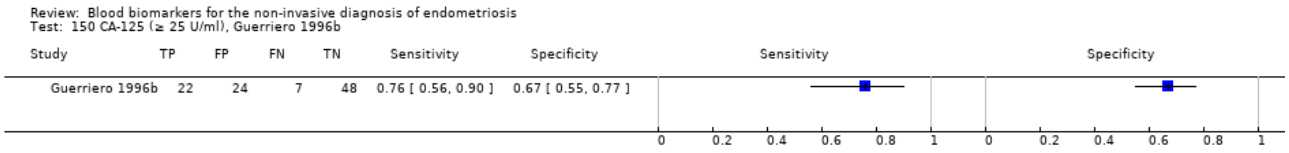
Test 148. CA-125 (> 35 U/ml), Gagne 2003a.



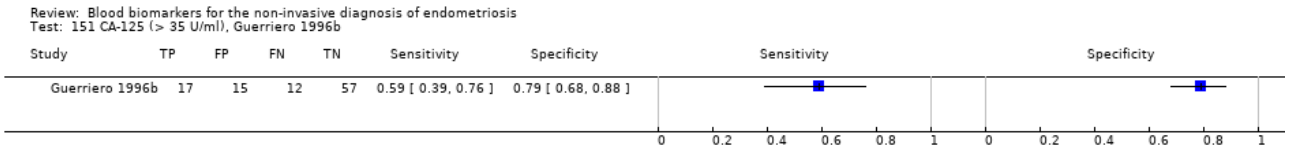
Test 149. CA-125 (> 20 U/ml), Guerriero 1996b.



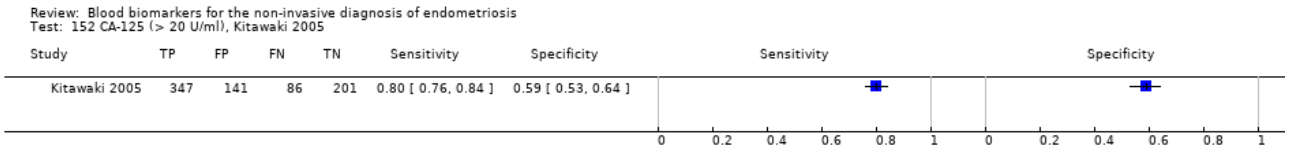
Test 150. CA-125 (≥ 25 U/ml), Guerriero 1996b.



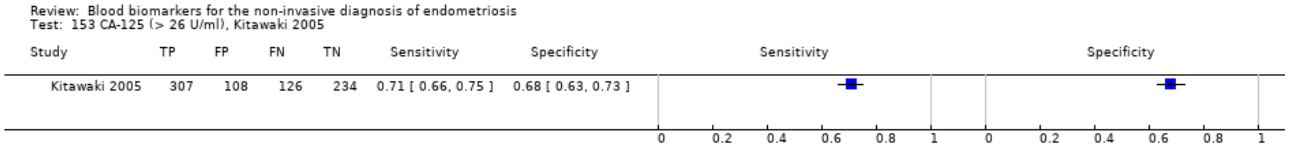
Test 151. CA-125 (> 35 U/ml), Guerriero 1996b.



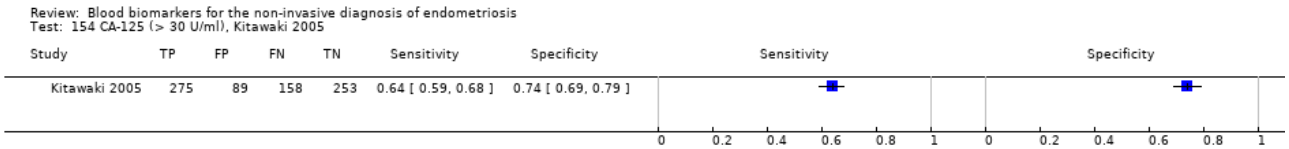
Test 152. CA-125 (> 20 U/ml), Kitawaki 2005.



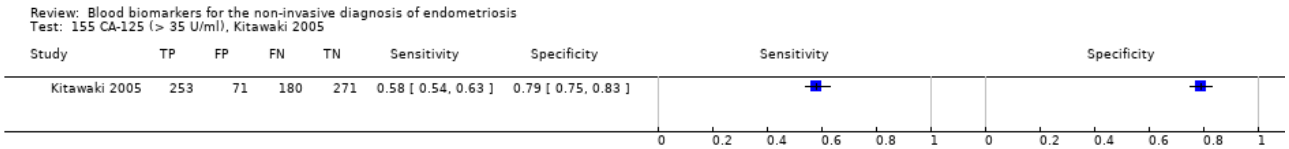
Test 153. CA-125 (> 26 U/ml), Kitawaki 2005.



Test 154. CA-125 (> 30 U/ml), Kitawaki 2005.

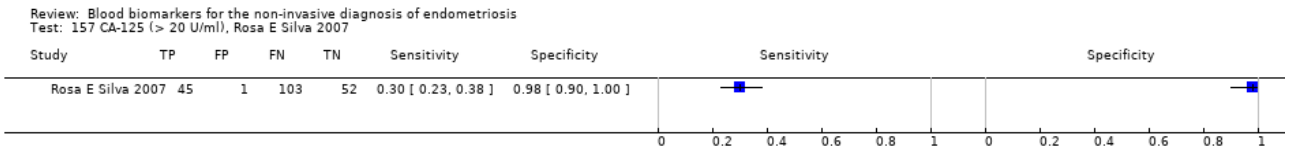


Test 155. CA-125 (> 35 U/ml), Kitawaki 2005.



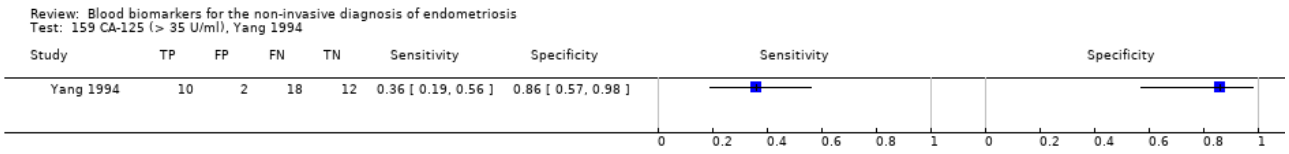
Test 156. CA-125 (> 10 U/ml), Rosa E Silva 2007.

Test 157. CA-125 (> 20 U/ml), Rosa E Silva 2007.

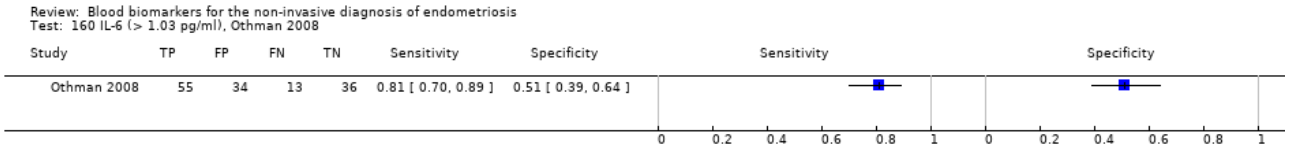


Test 158. CA-125 (> 20 U/ml), Yang 1994.

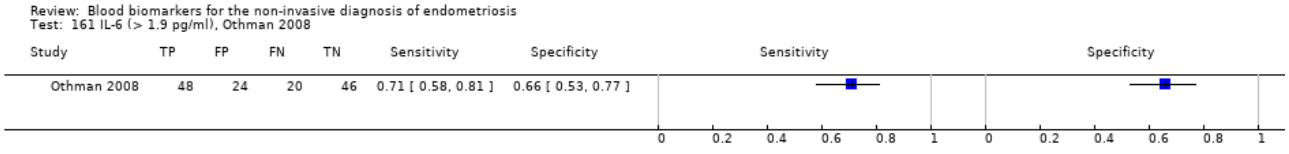
Test 159. CA-125 (> 35 U/ml), Yang 1994.



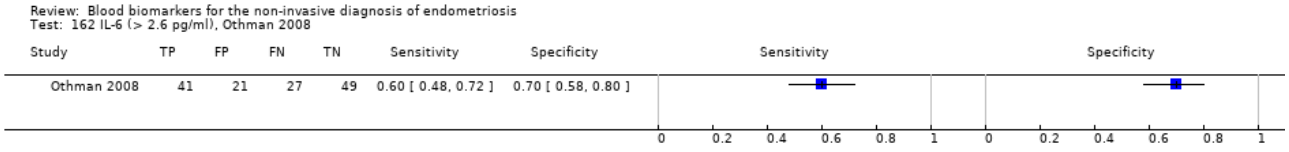
Test 160. IL-6 (> 1.03 pg/ml), Othman 2008.



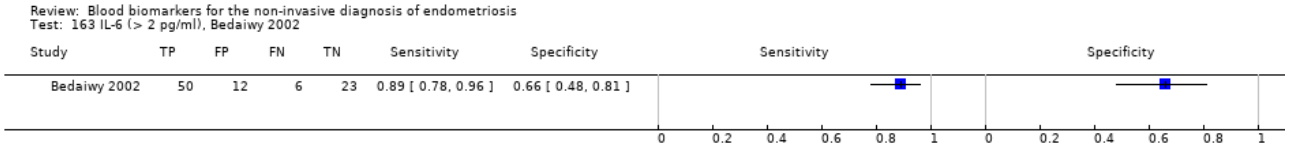
Test 161. IL-6 (> 1.9 pg/ml), Othman 2008.



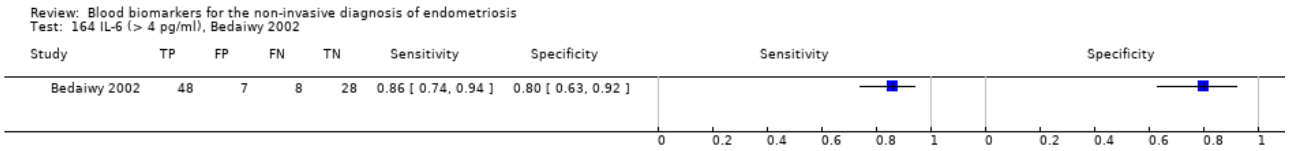
Test 162. IL-6 (> 2.6 pg/ml), Othman 2008.



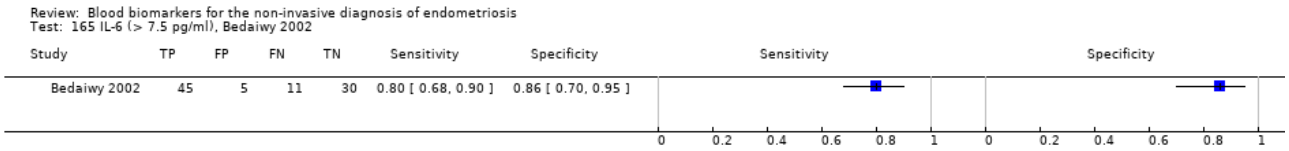
Test 163. IL-6 (> 2 pg/ml), Bedaiwy 2002.



Test 164. IL-6 (> 4 pg/ml), Bedaiwy 2002.



Test 165. IL-6 (> 7.5 pg/ml), Bedaiwy 2002.



ADDITIONAL TABLES

Table 1. Staging of endometriosis, rASRM classification

Location of endometriosis	**Extent**	**Depth**		
		< 1 cm	**1-3 cm**	**> 3 cm**
Peritoneum	Superficial	1	2	4
	Deep	2	4	6
Ovary	R Superficial	1	2	4
	Deep	4	16	20
	L Superficial	1	2	4
	Deep	4	16	20
Posterior cul-de-sac obliteration		**Partial**	**Complete**	
		4	40	
Adhesions		**< 1/3 Enclosure**	**1/3-2/3 Enclosure**	**> 2/3 Enclosure**
Ovary	R Filmy	1	2	4
	Dense	4	8	16
	L Filmy	1	2	4
	Dense	4	8	16
Tube	R Filmy	1	2	4

Table 1. Staging of endometriosis, rASRM classification *(Continued)*

Dense	4[a]	8[a]	16
L Filmy	1	2	4
Dense	4[a]	8[a]	16

Stage ·1 (Minimal) - score 1-5; Stage II (Mild) - score 6-15; Stage III (Moderate) - score 16-40; Stage IV (Severe) - score >40
[a]If the fimbriated end of the fallopian tube is completely enclosed, change the point assignment to 16 (ASRM 1997)

Table 2. Blood biomarkers evaluated in this review

Biomarker	
Angiogenesis and growth factors and their receptors	
Glycodelin-A (PP14 or PAEP) (or placental protein 14 or progestogen-associated endometrial protein)[a]	VEGF (vascular endothelial growth factor)[a]
IGFBP-3 (insulin-like growth factor-binding protein-3)[a]	Urocortin
Leptin[a]	
Apoptosis markers	
Annexin V[a]	Survivin
Cell adhesion molecules and other matrix-related proteins	
sICAM-1 (soluble form of intercellular-adhesion molecule-1)[a]	LN-1 (laminin-1)
High-throughput molecular markers	
Metabolome	Proteome
Hormonal markers	
Prolactin	
Immune system and inflammatory markers	
Autoantibodies • Anti-endometrial Abs (anti-endometrial auto antibodies)[a] • Anti-laminin-1 Abs (anti-laminin auto antibodies)	Immune cells • Neutrophils[a] • NLR (neutrophil-to-lymphocyte ratio)[a] • WBC (white blood cells)[a]
Chemokines • CCR1 (C-C motif receptor 1) • MCP-1 (monocyte chemotactic protein-1)[a]	Interleukins • IL-1β (interleukin - 1β)[a] • IL-4 (interleukin - 4)[a] • IL-6 (interleukin - 6)[a] • IL-8 (interleukin - 8)[a]
Other cytokines	Other immune/inflammatory markers

Table 2. Blood biomarkers evaluated in this review *(Continued)*

• IFN-γ (interferon-gamma)[a] • MIF (macrophage migration inhibitory factor)[a] • TNF-α (tumour necrosis factor alpha)[a]	• sCD23 (soluble CD23, low-affinity IgE receptor)[a] • Copeptin, *vasopressin surrogate* • hs-CRP (high sensitive C-reactive protein)[a]
Other peptides and proteins shown to influence key events implicated in endometriosis	
Follistatin, *activin-binding protein; involved in diverse activities from embryonic development to cell secretion*	
STX-5 (syntaxin-5), *protein belonging to syntaxin-family, a vesicular membrane fusion protein receptor in endoplasmic reticulum membrane*	
Oxidative stress markers	
Carbonyls	Thiols
PON-1 (paraoxonase-1)	
Post-transcriptional regulators of gene expression (microRNAs)	
miR-9*	miR-141*
miR-17-5	miR-145*
miR-20a	miR-199a
miR-22	miR-532-3p
miR-122	
Tumour markers	
CA-15.3 (cancer antigen-15.3)	CA-72 (TAG-72) (cancer antigen-72 or (tumour associated glycoprotein-72])
CA-19.9 (cancer antigen-19.9)[a]	CA-125 (cancer antigen-125)[a]
Blood biomarkers that did not exhibit differential expression in endometriosis and for which diagnostic performance was not assessed	
Angiogenesis and growth factors and their receptors	
Angiogenic activity of serum	IGF-1 (insulin-like growth factor-1)
CAC (circulating angiogenic cells)	IGF-2 (insulin-like growth factor-2)
EGF (epidermal growth factor)	IGFBP-3 (insulin-like growth factor binding protein-3)[a]
sEGF-R (soluble epidermal growth factor-receptor)	Leptin[a]
sFlt-1 (sVEGFR-1] (soluble fms-like tyrosine kinase or variant of VEGF receptor 1)	PDGF (platelet derived growth factor)
Glycodelin-A (PP14 or PAEP] (or placental protein 14 or progestogen-associated endometrial protein)[a]	VEGF (vascular endothelial growth factor)[a]
HGF (hepatocyte growth factor)	

Table 2. Blood biomarkers evaluated in this review *(Continued)*

Apoptosis markers	
Annexin V[a]	sFas (soluble Fas)
Apoptotic cells	anti-survivin Abs (anti-survivin antibodies)
Cell adhesion molecules and other matrix-related proteins	
Biglycan	sE-selectin (soluble E selectin)
sICAM-1 (soluble form of intercellular-adhesion molecule-1)[a]	MMP-9 (matrix metalloproteinase-9)
Cytoskeleton molecules	
CK 19 (Cytokeratin-19)	
DNA-repair and telomere maintenance molecules	
TL (telomere length)	
Hormonal markers	
E2 (oestradiol)	LH (luteinizing hormone)
FSH (follicle stimulating hormone)	Progesterone
Immune system and inflammatory markers	
Autoantibodies • Anti-endometrial Abs (anti-endometrial auto antibodies)[a] • Anti-sperm Abs (anti-sperm auto antibodies) • Anti-ZP Abs (anti-zona pellucida auto antibodies)	Interleukins • IL-1β[a] • IL-2 • IL-4 [a] • IL-6 [a] • IL-8 [a] • IL-10 • IL-12 • IL-15 • IL-16 • IL-18 • IL-13 • IL-17 • IL-23
Chemokines • MCP-1 (monocyte chemotactic protein-1)[a] other Cytokines • Epo (erythropoietin) • GM-CSF (granulocyte/macrophage-colony stimulating factor) • IFN-γ (interferon-gamma)[a] MIF (macrophage migration inhibitory factor)[a] • TNF-α (tumour necrosis factor alpha)[a]	
Immune cells	Other immune/inflammatory markers

Table 2. Blood biomarkers evaluated in this review *(Continued)*

- Peripheral blood mononuclear cells:
  - Lymphocytes (overall and subpopulations of B- and T-cells)
  - Monocytes/macrophages
  - Neutrophils[a]
  - NLR (neutrophil-to-lymphocyte ratio)[a]
  - NK (natural killer cells)
  - NKR (natural killer cells receptors)
  - Tregs (Regulatory T cells)
  - WBC (white blood cells)[a]
- Other blood cells and blood cell parameters
  - Haemoglobin
  - MPV (mean platelet volume)
  - Platelet count
  - PLR (platelet/lymphocyte ratio
- C3a (anaphylatoxin)
- sCD23 (soluble CD23, low-affinity IgE receptor)[a]
- sCD163 (soluble haemoglobin scavenger receptor)
- CRP (C-reactive protein)[a]
- sHLA-I (soluble human leukocyte class I antigens)
- Immunoglobulins: IgA, IgG
- MPO (myeloperoxidase)
- NAG (N-acetyl-b-Dglucosaminidase)
- PGE2 (prostaglandin E2)
- Phospholipid fatty acids
- PLA2G2A (phospholipase A2 group IIA)
- RANTES (regulated on activation, normal T cell expressed and secreted)

Nerve growth markers	
CNTF (ciliary Neurotrophic Factor)	NGF (nerve growth factor)
GDNF (glial cell-derived neurotrophic factor)	NT4 (neurotrophin 4)
Other peptides and proteins shown to influence key events implicated in endometriosis	
DBP (vitamin D binding protein), *component of Gc-globulin and is the major plasma carrier protein of vitamin D metabolites, responsible for the transport of fat and endotoxins, important factor in the actin scavenging system, plays an important role in the immune system*	
Enolase (phosphopyruvate hydratase], *a glycolytic enzyme, frequently associated with autoimmune diseases*	
PDPK1 (phosphoinositide dependent protein kinase 1), *a master kinase involved in the signalling pathways* *activated by several growth factors and hormones (glucose metabolism, cellular proliferation, cellular survival, and angiogenesis)*	
Oxidative stress markers	
Ascorbic acid	Nitrotyrosine
GSH (glutathione)	SOD3 (superoxide dismutase-3)
HSP70 (heat shock protein 70)	TRX (Thioredoxin)
IMA (Ischemia-modified albumin)	Vitamin E
Malondialdehyde	
Tumour markers	
AFP (alpha-fetoprotein)	c-erbB-2 (HER-2/neu] (erythroblastosis oncogene B or human epidermal growth factor receptor-2 derived from glioblastoma)
CA-19.9 (cancer antigen-19.9)[a]	HE4 (human epididymal secretory protein E4)
CA-125 (cancer antigen-125)[a]	

Table 2. Blood biomarkers evaluated in this review *(Continued)*

[a] Biomarkers that belong to both groups (evaluated as a diagnostic test for endometriosis in some studies and did not exhibit differential expression in endometriosis in the other studies).

For a comprehensive list of all biomarkers with their biological annotation, please see Appendix 1.

Table 3. Application of the QUADAS-2 tool for assessment of methodological quality of the included studies

Domain 1 - Patient selection	
Description	Describe methods of patient selection and included patients
Type of bias assessed	Selection bias, spectrum bias
Review Question	Women of reproductive age with clinically suspected endometriosis (symptoms, clinical examination ± presence of pelvic mass), scheduled for surgical exploration of pelvic/abdominal cavity for confirmation of the diagnosis ± treatment
Informaton collected	Study objectives, study population, selection (inclusion/exclusion criteria), study design, clinical presentation, age, number of enrolled and number of available for analysis, setting, place and period of the study
Signalling question 1	***Was a consecutive or random sample of patients enrolled?***
Yes	If a consecutive sample or a random sample of the eligible participants was included in the study
No	If a consecutive sample or a random sample of the eligible participants was not included in the study
Unclear	If this information was unclear
Signalling question 2	***Did the study avoid inappropriate exclusions?***
Yes	If inclusion and exclusion criteria were presented and all participants with suspected endometriosis were included, with an exception for those who either had a history of medical conditions or were on medical therapy that would have potentially interfered with interpretation of index test (e.g. malignancy, pregnancy, autoimmune disorders, infectious diseases, treatment with hormonal or immunomodulator substances); refused to participate in the study; or were unfit for surgery
No	If the study excluded the participants based on education level, psychosocial factors, genetic testing or phenotype or excluded participants with any comorbidities commonly present in general population, including a population that could have undergone a testing for endometriosis in clinical setting (hypertension, asthma, obesity, benign gastrointestinal or renal disease, etc)
Unclear	If the study did not provide clear definition of the selection (inclusion/exclusion) criteria and 'no' judgement was not applicable
Signalling question 3	***Was a 'two-gate' design avoided?***

Table 3. Application of the QUADAS-2 tool for assessment of methodological quality of the included studies *(Continued)*

Yes	If the study had a single set of inclusion criteria, defined by the clinical presentation (i.e. only participants in whom the target condition is suspected) - a single-gate design
No	If the study had more than one set of inclusion criteria in respect to clinical presentation (i.e. participants suspected of target condition and participants with alternative diagnosis in whom the target condition would not be suspected in clinical practice) - a two-gate study design
Unclear	If it was unclear whether a two-gate deign was avoided or not
Risk of bias	**Could the selection of patients have introduced bias?**
Low	If 'yes' classification for all the above 3 questions
High	If 'no' classification for any of the above 3 questions
Unclear	If 'unclear' classification for any of the above 3 questions and 'high risk' judgement was not applicable
Concerns about applicability	**Are there concerns that the included patients do not match the review question?**
Low	If the study includes only clinically relevant population that would have undergone index test in real practice and includes representative form of target condition
High	If the study population differed from the population defined in the review question in terms of demographic features and comorbidity (e.g. studies with multiple sets of inclusion criteria with respect to clinical presentation including either healthy controls or alternative diagnosis controls that would not have undergone index test in real practice). Further, if target condition diagnosed in the study population was not representative of the entire spectrum of disease, such as limited spectrum of severity (e.g. only mild forms) or limited type of endometriosis (e.g. only deep infiltrating endometriosis)
Unclear	If this information was unclear (e.g. severity of endometriosis was not reported)
Domain 2 - Index test	
Description	Describe the index test, how it was conducted and interpreted
Type of bias assessed	Test review bias, clinical review bias, interobserver variation bias
Review question	Any type of blood-based biomarker
Informaton collected	Index test name, description of positive case definition by index test as reported, threshold for positive result, examiners (number, level of expertise, blinding), interobserver variability, conflict of interests
Signalling question 1	***Were the index test results interpreted without knowledge of the results of the reference standard?***
Yes	If the operators performing/interpreting index test were unaware of the results of the reference standard

Table 3. Application of the QUADAS-2 tool for assessment of methodological quality of the included studies *(Continued)*

No	If the operators performing/interpreting index test were not blinded to the results of the reference standard
Unclear	If this information was unclear
Signalling question 2	***If a threshold was used, was it pre-specified?***
Yes	If study clearly provided a threshold for positive result and was defined before execution/interpretation of index test
No	If a threshold for positive result was not provided or not defined prior to test execution
Unclear	If it was unclear whether a threshold was pre-specified or not
Signalling question 3	***Was a menstrual cycle phase considered in interpreting the index test?***
Yes	If all the included participants were in the same phase of menstrual cycle, if the study reported subgroup analyses per cycle phase, or if study reported the pooled estimates after impact of the cycle phase on biomarker expression was not detected
No	If study included participants in different phases of menstrual cycle, but effect of cycle phase on index test was not assessed
Unclear	If the cycle phase was not reported
Risk of bias	**Could the conduct or interpretation of the index test have introduced bias?**
Low	If 'yes' classification for all the above 3 questions
High	If 'no' classification for any of the above 3 questions
Unclear	If 'unclear' classification for any of the above 3 questions and 'high risk' judgement was not applicable
Concerns about applicability	**Are there concerns that the index test, its conduct, or interpretation differ from the review question?**
Low	We considered all types of blood-based biomarkers as eligible, therefore all the included studies were classified as 'low concern', unless 'unclear' judgement was applicable
High	We did not consider the studies where index tests other than blood-based biomarkers were included (or excluded information on other index tests reported in addition to blood tests) or where index test looked at other target conditions not specified in the review (e.g. studies aimed at classifying pelvic masses as benign and malignant); therefore none of the included studies was classified as 'high concern'
Unclear	If study reported, but did not present sufficient information on any of the following: laboratory method, sample handling, reagents used or experience of the test operators
Domain 3 - Reference standard	

Table 3. Application of the QUADAS-2 tool for assessment of methodological quality of the included studies *(Continued)*

Description	Describe the reference standard, how it was conducted and interpreted
Type of bias assessed	Verification bias, bias in estimation of diagnostic accuracy due to inadequate reference standard
Review question	Target condition - pelvic endometriosis, ovarian endometriosis, deep infiltrating endometriosis. Reference standard - visualisation of endometriosis at surgery (laparoscopy or laparotomy) with or without histological confirmation
Informaton collected	Target condition, prevalence of target condition in the sample, reference standard, description of positive case definition by reference test as reported, examiners (number, level of expertise, blinding), interobserver variability, conflict of interests
Signalling question 1	***Is the reference standard likely to correctly classify the target condition?***
Yes	If the study reported at least one of the following: surgical procedure was described in sufficient detail; criteria for positive reference standard were stated; diagnosis was confirmed by histopathology; or the procedure was performed by a team with high level of expertise in diagnosis/surgical treatment of target condition, including tertiary referral centres for endometriosis
No	If reference standard did not classify target condition correctly; considering the inclusion criteria and nature of the reference standard, none of the studies were classified as 'no' for this item
Unclear	If information on execution of the reference standard, its interpretation or operators was unclear
Signalling question 2	***Were the reference standard results interpreted without knowledge of the results of the index tests?***
Yes	If operators performing the reference test were unaware of the results of the index test
No	If operators performing the reference test were aware of the results of the index test
Unclear	If this information was unclear
Risk of bias	**Could the reference standard, its conduct, or its interpretation have introduced bias?**
Low	If 'yes' classification for both of the above 2 questions
High	If 'no' classification for any of the above 2 questions
Unclear	If 'unclear' classification for any of the above 2 questions and 'high risk' judgement was not applicable
Concerns about applicability	**Are there concerns that the target condition as defined by the reference standard does not match the question?**
Low	Considering the inclusion criteria, all the studies were classified as 'low concern', unless 'unclear' judgement was applicable

Table 3. Application of the QUADAS-2 tool for assessment of methodological quality of the included studies *(Continued)*

High	We excluded the studies where participants did not undergo surgery for diagnosis of endometriosis; therefore none of the included studies were classified as 'high concern'
Unclear	Only studies where laparoscopy/laparotomy served as a reference test were included; therefore none of the included studies were classified as 'unclear concern'
Domain 4 - Flow and timing	
Description	Describe any participants who did not receive the index tests or reference standard or who were excluded from the 2 x 2 table; describe the interval and any interventions between index tests (sample collection) and the reference standard
Type of bias assessed	Disease progression bias, bias of diagnostic performance due to missing data
Review question	Less than 12-month interval between index test (sample collection) and reference standard - endometriosis may progress over the time, so we had chosen an arbitrary time interval of 12 months as an acceptable time interval between the index test and surgical confirmation of diagnosis
Informaton collected	Time interval between index test (sample collection) and reference standard, withdrawals (overall number of reported and if explanation)
Signalling question 1	***Was there an appropriate interval between index test (sample collection) and reference standard?***
Yes	If time interval was reported and was less than 12 months
No	We excluded all the studies where time interval was longer than 12 months; therefore none of the included studies were classified as 'no' for this item
Unclear	If time interval was not stated clearly, but authors description allowed us to assume that the interval was reasonably short
Signalling question 2	***Did all patients receive the same reference standard?***
Yes	If all participants underwent laparoscopy/laparotomy as a reference standard. Considering the inclusion criteria, all the studies were classified as 'yes' for this item, as anticipated
No	If all participants did not undergo surgery or had alternative reference standard or if only a subset of participants had surgery as reference standard, but the information on this population was not available in isolation
Unclear	If this information was unclear. Considering the inclusion criteria, none of the included studies were classified as 'unclear' for this item
Signalling question 3	***Were all patients included in the analysis?***
Yes	If all the participants were included in the analysis or if the participants were excluded because they did not meet inclusion criteria prior to execution of index test or if the withdrawals were less than 5% of the enrolled population (arbitrary selected cut-off)

Table 3. Application of the QUADAS-2 tool for assessment of methodological quality of the included studies *(Continued)*

No	If any participants were excluded from the analysis because of uninterpretable results, inability to undergo either index test or reference standard, or unclear reasons
Unclear	If this information was unclear
Risk of bias	**Could the patient flow have introduced bias?**
Low	If 'yes' classification for all the above 3 questions
High	If 'no' classification for any of the above 3 questions
Unclear	If 'unclear' classification for any of the above 3 questions and 'high risk' judgement was not applicable

Table 4. Blood biomarkers to be validated for their diagnostic potential in endometriosis

Blood biomarkers [1]	Replacement test	SnOUT triage test	SpIN triage test
1. Angiogenesis and growth markers			
VEGF > 680 pg/ml	±	±	+
VEGF > 236 pg/ml	±	±	
2. High-throughput markers			
Metabolome by ESI-MS/MS (SMOH C16:1 + PCaa C36:2/PCae C34:2)		±	
Proteome by SELDI-TOF-MS (6 peaks with molecular weights of 1.63, 3.05, 3.53, 3.77, 5.05 and 5,07 Da)			+
3. Immune system and inflammatory markers			
IL-6 > 12.2 pg/ml	+	+	
4. Oxidative stress markers			
PON-1 < 141.5 U/l	+	+	
Carbonyls < 14.9 µM		±	
5. Post-transcriptional regulators of gene expression (microRNAs)			
miR-9*			+
miR-141*			+
miR-145*			+
miR-20a < 0.69			±
miR-22 < 0.56	±	±	

Table 4. Blood biomarkers to be validated for their diagnostic potential in endometriosis *(Continued)*

miR-532-3p			±
6. Tumour markers			
CA-125 (cut-off value > 43 U/ml)	+	+	
7. Combined blood tests			
IL-6 > 12.2 pg/ml + TNF-α > 12.45 pg/ml			+
IL-6 > 12.2 pg/ml + CRP > 438 µg/ml			+
TNF-α > 12.45 pg/ml + CRP > 438 µg/ml			+
miR-199a + miR-542-3p	+	+	
CA-125 + STX-5 + LN-1	±	+	
IL-6 > 12.2 pg/ml + TNF-α > 12.45 pg/ml + CRP > 438 µg/ml			+
miR-199a + miR-122 + miR-145* + miR-542-3p	±	±	+
CA-125 > 17.6 IU/ml + VEGF > 236 pg/ml			±
CA-125 + CA-19-9 + survivin			±
CA-125 > 50 IU/ml +/or CCR1 > 1.16 +/or MCP-1 > 140 pg/ml	±	±	
CA-125 > 20 IU/ml + MCP-1 > 152.744 pg/ml + leptin > 3.14 ng/ml			±
CA-125 + IL-8 + TNF-α		±	
CA-125 + CA-19.9 + IL-6 + IL-8 + TNF-α + hs-CRP (in menstrual phase of the cycle)	±	±	
8. Tests that specifically differentiate endometrioma from other benign ovarian cysts in women of reproductive age			
Urocortin > 29 pg/ml	+	+	
Urocortin > 33 pg/ml			±
Follistatin > 1433 pg/ml	±	±	±
CA-125 > 30 U/ml and > 36 U/ml			+
CA-125 ≥ 25 U/ml + CA-19.9 ≥ 22 U/ml			±

Notes:

+ meets the criteria

- Replacement test: sensitivity ≥ 94 and specificity ≥ 79
- SnOUT triage test: sensitivity ≥ 95 and specificity ≥ 50
- SpIN triage test: sensitivity ≥ 50 and specificity ≥ 95

± approaches the criteria (within 5% of the pre-defined criteria)

[1] This group included: tests with an adequate diagnostic performance, but insufficient data to confidently comment on their diagnostic role (less than 3 studies with the diagnostic estimates meeting the criteria for either a replacement or triage test); and tests where the diagnostic estimates were approaching the criteria for replacement or triage tests in a small number of studies, and it is possible that they would reach this criteria if further studies were performed (less than 3 studies with the diagnostic estimates within 5% of the criteria for either replacement or triage tests).
For a comprehensive list of all biomarkers with their biological annotation, please see Appendix 1.

APPENDICES

Appendix 1. Alphabetical list of blood biomarkers

	Biomarker	Biological group	Biological sub-group 1	Biological sub-group 2
1	Angiogenic activity of serum	Angiogenesis and growth factors and their receptors		
2	Annexin V	Apoptosis markers		
3	Anti-endometrial Abs or AEA (anti-endometrial autoantibodies)	Immune system and inflammatory markers	Autoantibodies	
4	Anti-laminin-1 Abs (anti-laminin autoantibodies)	Immune system and inflammatory markers	Autoantibodies	
5	Anti-sperm Abs (anti-sperm autoantibodies)	Immune system and inflammatory markers	Autoantibodies	
6	Anti-survivin Abs (anti-survivin antibodies)	Apoptosis markers		
7	Anti-ZP Abs (anti-zona pellucida autoantibodies)	Immune system and inflammatory markers	Autoantibodies	
8	Apoptotic cells	Apoptosis markers		
9	Ascorbic acid	Oxidative stress markers		
10	Biglycan	Cell adhesion molecules and other matrix-related proteins		
11	B-lymphocytes	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)
12	C3a (anaphylatoxin)	Immune system and inflammatory markers	Other immune/inflammatory markers	
13	CA-125 (cancer antigen-125)	Tumour markers		
14	CA-15.3 (cancer antigen-15.3)	Tumour markers		
15	CA-19.9 (cancer antigen-19.9)	Tumour markers		

(Continued)

16	CA-72 (TAG-72) (cancer antigen-72 or (tumour associated glycoprotein-72))	Tumour markers	
17	CAC (circulating angiogenic cells)	Angiogenesis and growth factors and their receptors	
18	Carbonyls	Oxidative stress markers	
19	CCR1 (C-C motif receptor 1)	Immune system and inflammatory markers	Chemokines
20	c-erbB-2 (HER-2/neu) (erythroblastosis oncogene B or human epidermal growth factor receptor-2 derived from glioblastoma)	Tumour markers	
21	CK 19 (cytokeratin-19)	Cytoskeleton molecules	
22	CNTF (ciliary neurotrophic factor)	Nerve growth markers	
23	Copeptin	Immune system and inflammatory markers	Other immune/inflammatory markers
24	CRP (C-reactive protein) or hs-CRP (high sensitive C-reactive protein)	Immune system and inflammatory markers	Other immune/inflammatory markers
25	DBP (vitamin D binding protein)	Other peptides/proteins shown to influence key events implicated in endometriosis	
26	E2 (oestradiol)	Hormonal markers	
27	EGF (epidermal growth factor)	Angiogenesis and growth factors and their receptors	
28	Enolase	Other peptides/proteins shown to influence key events implicated in endometriosis	
29	Epo (erythropoietin)	Immune system and inflammatory markers	Other cytokines
30	Follistatin	Other peptides/proteins shown to influence key events implicated in endometriosis	
31	FSH (follicle stimulating hormone)	Hormonal markers	
32	GDNF (glial-derived neurotrophic factor)	Nerve growth markers	

(Continued)

33	glycodelin-A (PP14 or PAEP) (placental protein 14 or progestogen-associated endometrial protein)	Angiogenesis and growth factors and their receptors		
34	GM-CSF (granulocyte macrophage-colony stimulating factor) or sGM-CSF (soluble granulocyte macrophage-colony stimulating factor)	Immune system and inflammatory markers	Other cytokines	
35	GSH (glutathione)	Oxidative stress markers		
36	Haemoglobin	Immune system and inflammatory markers	Immune cells	Other blood cells and blood cell parameters
37	HE4 (human epididymal secretory protein E4)	Tumour markers		
38	HGF (hepatocyte growth factor)	Angiogenesis and growth factors and their receptors		
39	HSP70 (heat shock protein 70)	Oxidative stress markers		
40	IFN-γ (interferon-gamma) or sIFN-γ (soluble interferon-gamma)	Immune system and inflammatory markers	Other cytokines	
41	IGF-1 (insulin-like growth factor-1) or sIGF-1 (soluble Insulin-like growth factor-1)	Angiogenesis and growth factors and their receptors		
42	IGF-2 (insulin-like growth factor-2)	Angiogenesis and growth factors and their receptors		
43	IGFBP-3 (insulin-like growth factor binding protein-3)	Angiogenesis and growth factors and their receptors		
44	IL-1β	Immune system and inflammatory markers	Interleukins	
45	IL-2	Immune system and inflammatory markers	Interleukins	
46	IL-4	Immune system and inflammatory markers	Interleukins	
47	IL-6	Immune system and inflammatory markers	Interleukins	
48	IL-8	Immune system and inflammatory markers	Interleukins	
49	IL-10	Immune system and inflammatory markers	Interleukins	

(Continued)

50	IL-12	Immune system and inflammatory markers	Interleukins	
51	IL-13	Immune system and inflammatory markers	Interleukins	
52	IL-15	Immune system and inflammatory markers	Interleukins	
53	IL-16	Immune system and inflammatory markers	Interleukins	
54	IL-17	Immune system and inflammatory markers	Interleukins	
55	IL-18	Immune system and inflammatory markers	Interleukins	
56	IL-23	Immune system and inflammatory markers	Interleukins	
57	IMA (ischemia-modified albumin)	Oxidative stress markers		
58	Immunoglobulins IgA or IgG	Immune system and inflammatory markers	Other immune/inflammatory markers	
59	Leptin	Angiogenesis and growth factors and their receptors		
60	LH (luteinizing hormone)	Hormonal markers		
61	LN-1 (laminin-1)	Cell adhesion molecules and other matrix-related proteins		
62	Lymphocytes	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)
63	Malondialdehyde	Oxidative stress markers		
64	MCP-1 (monocyte chemotactic protein-1)	Immune system and inflammatory markers	Chemokines	
65	Metabolome	High-throughput molecular markers		
66	MIF (macrophage migration inhibitory factor)	Immune system and inflammatory markers	Other cytokines	
67	miR-122	Post-transcriptional regulators of gene expression (microRNAs)		
68	miR-141*	Post-transcriptional regulators of gene expression (microRNAs)		

(Continued)

69	miR-145*	Post-transcriptional regulators of gene expression (microRNAs)		
70	miR-17-5	Post-transcriptional regulators of gene expression (microRNAs)		
71	miR-199a	Post-transcriptional regulators of gene expression (microRNAs)		
72	miR-20a	Post-transcriptional regulators of gene expression (microRNAs)		
73	miR-22	Post-transcriptional regulators of gene expression (microRNAs)		
74	miR-532-3p	Post-transcriptional regulators of gene expression (microRNAs)		
75	miR-9*	Post-transcriptional regulators of gene expression (microRNAs)		
76	MMP-9 (matrix metalloproteinase-9)	Cell adhesion molecules and other matrix-related proteins		
77	Monocytes/macrophages	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)
78	MPO (myeloperoxidase)	Immune system and inflammatory markers	Other immune/inflammatory markers	
79	MPV (mean platelet volume)	Immune system and inflammatory markers	Immune cells	Other blood cells and blood cell parameters
80	NAG (N-acetyl-b-Dglucosaminidase)	Immune system and inflammatory markers	Other immune/inflammatory markers	
81	Neutrophils	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)
82	NGF (nerve growth factor)	Nerve growth markers		
83	Nitrotyrosine	Oxidative stress markers		
84	NK (natural killer cells)	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)
85	NKR CD158b+ (KIR2DL2+NK) (killer cell inhibitory receptor subfamily 2DL2 on NK cells)	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)

(Continued)

86	NKR CD94 + (lectin-like receptor on natural killer cells)	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)
87	NLR (neutrophil/lymphocyte ratio)	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)
88	NT4 (neurotrophin 4)	Nerve growth markers		
89	PAEP (glycodelin) (progestagen-associated endometrial protein)	Angiogenesis and growth factors and their receptors		
90	PDGF (platelet derived growth factor)	Angiogenesis and growth factors and their receptors		
91	PDPK1 (phosphoinositide dependent protein kinase 1)	Other peptides/proteins shown to influence key events implicated in endometriosis		
92	PGE2 (prostaglandin E2)	Immune system and inflammatory markers	Other immune/inflammatory markers	
93	Phospholipid fatty acids	Immune system and inflammatory markers	Other immune/inflammatory markers	
94	PLA2G2A (phospholipase A2 group IIA)	Immune system and inflammatory markers	Other immune/inflammatory markers	
95	Platelet count	Immune system and inflammatory markers	Immune cells	Other blood cells and blood cell parameters
96	PLR (platelet/lymphocyte ratio)	Immune system and inflammatory markers	Immune cells	Other blood cells and blood cell parameters
97	PON-1 (paraoxonase-1)	Oxidative stress markers		
98	Progesterone	Hormonal markers		
99	Prolactin	Hormonal markers		
100	Proteome	High-throughput molecular markers		
101	RANTES (regulated on activation, normal T cell expressed and secreted)	Immune system and inflammatory markers	Other immune/inflammatory markers	
102	sCD163 (soluble haemoglobin scavenger receptor)	Immune system and inflammatory markers	Other immune/inflammatory markers	

(Continued)

103	sCD23 (soluble CD23, low-affinity IgE receptor)	Immune system and inflammatory markers	Other immune/inflammatory markers	
104	sEGF-R (soluble epidermal growth factor-receptor)	Angiogenesis/ Growth factors and their receptors		
105	sE-selectin (soluble E selectin)	Cell adhesion molecules and other matrix-related proteins		
106	sFas (soluble Fas)	Apoptosis markers		
107	sFlt-1 (sVEGFR-1) (soluble fms-like tyrosine kinase or (variant of VEGF receptor 1)	Angiogenesis and growth factors and their receptors		
108	sHLA-I (soluble human leukocyte class I antigens)	Immune system and inflammatory markers	Other immune/inflammatory markers	
109	sICAM-1 (soluble form of intercellular adhesion molecule-1)	Cell adhesion molecules and other matrix-related proteins		
110	SOD3 (superoxide dismutase)	Oxidative stress markers		
111	STX-5 (syntaxin-5)	Other peptides/proteins shown to influence key events implicated in endometriosis		
112	Survivin	Apoptosis markers		
113	Thiols	Oxidative stress markers		
114	TL (telomere length)	DNA-repair/telomere maintenance molecules		
115	T-lymphocytes	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)
116	TNF-α (tumour necrosis factor alpha)	Immune system and inflammatory markers	Other cytokines	
117	Tregs (regulatory T cells)	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)
118	TRX (thioredoxin)	Oxidative stress markers		
119	Urocortin	Angiogenesis and growth factors and their receptors		
120	VEGF (vascular endothelial growth factor)	Angiogenesis and growth factors and their receptors		
121	Vitamin E	Oxidative stress markers		

(Continued)

122	WBC (white blood cells)	Immune system and inflammatory markers	Immune cells	Peripheral blood mononuclear cells (PBMC)

Appendix 2. Search strategy for MEDLINE (OVID platform)

Database: MEDLINE (Ovid) <1946 to February, week 2 2015 (16.2.2015)>

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R)

Search Strategy:

---

1 (biomarker$ or marker$).tw. (605002)

2 Laboratory Test$.tw. (29839)

3 growth factor$.tw. (272049)

4 scatter factor$.tw. (1287)

5 cytokine$.tw. (250618)

6 hepatocyte growth factor.tw. (8053)

7 (FGF or fibroblast growth factor$).tw. (31798)

8 (PDGF or platelet derived growth factor$).tw. (19864)

9 (EGF or epidermal growth factor$).tw. (58069)

10 (IGF-I or insulin-like growth factor$ or IGF1).tw. (43539)

11 (TGF-a or transforming growth factor alfa or TGFa).tw. (281)

12 (TGF-b or transforming growth factor beta or TGFb).tw. (28842)

13 (EGFR or epidermal growth factor receptor$).tw. (41719)

14 (VEGF or vascular endothelial growth factor$).tw. (53588)

15 exp Luteinizing Hormone/bl (Blood] (24587)

16 leptin$.tw. (24994)

17 exp Progesterone/bl (Blood] (18412)

18 Proteolytic enzyme$.tw. (9768)

19 exp matrix metalloproteinase 1/ or exp matrix metalloproteinase 2/ or exp matrix metalloproteinase 3/ or exp matrix metalloproteinase 9/ (22968)

20 matrix metalloproteinase$.tw. (34522)

21 MMP$.tw. (44439)

22 TIMP$.tw. (10777)

23 exp "tissue inhibitor of metalloproteinase-1"/ or exp "tissue inhibitor of metalloproteinase-2"/ (6146)

24 exp Glycoproteins/ (637149)

25 (Ca-125 or Ca125 or cancer antigen 125).tw. (6761)

26 (Ca-19-9 or Ca19-9 or cancer antigen 19-9).tw. (4194)

27 (PP 14 or PP14).tw. (229)

28 serum placental protein$.tw. (33)

29 exp Follistatin/ (1134)

30 Osteopontin$.tw. (6769)

31 exp intercellular adhesion molecule-1/ or exp selectins/ (25302)

32 soluble intercellular adhesion.tw. (1588)

33 Soluble adhesion molecule$.tw. (779)

34 sICAM.tw. (2258)

35 sVCAM$.tw. (1277)

36 (sEcadherin or soluble E-cadherin).tw. (95)

37 (sEselectin or soluble E-selectin).tw. (689)

38 exp t-lymphocytes/ or exp natural killer t-cells/ (272580)

39 Immune cells alteration$.tw. (1)

40 (T helper$ or T supressor$ or T helper$ T supressor$ ratio).tw. (21275)

41 Total complement level$.tw. (23)

42 Autoantibodies.tw. (33457)

43 exp Antibodies, Antiphospholipid/ (7522)

44 Anti-endometrial.tw. (23)

45 Antiphospholipid$.tw. (9974)

46 exp hla antigens/ or exp hla-a1 antigen/ or exp hla-a2 antigen/ (64462)

47 (HLA or human leucocyte antigen$).tw. (80501)

48 Anti-laminin-1.tw. (33)

49 Anti-thyroid.tw. (1414)

50 Anti-Thomsen Friedenreich antigen$.tw. (6)

51 Anti-transferrin.tw. (275)

52 Anti-LDL.tw. (181)

53 (Anti-2HSG or Heremans-Schmidt glycoprotein).tw. (3)

54 interleukin$.tw. (175195)

55 (MCP-I or monocyte chemoattractant protein-I).tw. (44)

56 (MIF or migration inhibitory factor$).tw. (4479)

57 (TNF-a or tumour necrosis factor$ alfa).tw. (1344)

58 Fas ligand$.tw. (6032)

59 Endometrial marker$.tw. (11)

60 CAMs.tw. (1756)

61 cell adhesion molecule$.tw. (20903)

62 exp Integrins/ (44414)

63 Integrin$.tw. (39960)

64 Selectin$.tw. (55426)

65 Cadherin$.tw. (20780)

66 Aromatase P450.tw. (180)

67 estrogen receptor$.tw. (38819)

68 progesterone receptor$.tw. (16623)

69 MTMMP$.tw. (7)

70 cyr61.tw. (559)

71 exp Cysteine-Rich Protein 61/ (386)

72 cysteine-rich heparin-binding protein$.tw. (9)

73 (ANXA 1 or ANXA1).tw. (313)

74 (Annexin 1 or Annexin1).tw. (339)

75 (PGP 9?5 or PGP9?5 or protein gene product$).tw. (2096)

76 serum marker$.tw. (5429)

77 neural marker$.tw. (925)

78 cell surface marker$.tw. (4456)

79 inflammatory marker$.tw. (10916)

80 microarray$.tw. (75404)

81 microRNA$.tw. (29731)

82 proteomic$.tw. (45292)

83 genomic$.tw. (190985)

84 (endometri$ adj2 biops$).tw. (3411)

85 Follistatin$.tw. (1663)

86 Vascular Endothelial Growth Factor A/ (35738)

87 Vitamin D-Binding Protein/ (1282)

88 exp Cytokines/ (547522)

89 exp interleukins/ or exp interleukin-1/ or exp interleukin-6/ or exp interleukin-8/ or exp interleukin-12/ or exp interleukin-13/ (188479)

90 exp Epidermal Growth Factor/ (21298)

91 exp Fibroblast Growth Factors/ (25075)

92 Platelet-Derived Growth Factor/ (11030)

93 Keratin-19/ (1090)

94 exp Clinical Laboratory Techniques/ (2132820)

95 (Luteinizing Hormone$ or LH).tw. (56679)

96 cytokeratin-19.tw. (1469)

97 (VDBP or vitamin D-binding protein$).tw. (1158)

98 urinary peptide$.tw. (137)

99 VDBP-Cr.tw. (1)

100 urinary VDBP corrected for creatinine expression.tw. (1)

101 urinary marker$.tw. (638)

102 or/1-101 (4086291)

103 Endometriosis/di (Diagnosis] (3354)

104 102 or 103 (4088946)

105 exp Endometriosis/ (17244)

106 Endometrio$.tw. (21492)

107 105 or 106 (24940)

108 104 and 107 (10490)

109 (animals not (humans and animals)).sh. (3892900)

110 108 not 109 (10113)

Additional search February 2015 - May 2015

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present (3.9.2015)>

Search Strategy:

---

1 (biomarker$ or marker$).tw. (652345)

2 Laboratory Test$.tw. (31389)

3 growth factor$.tw. (287701)

4 scatter factor$.tw. (1326)

5 cytokine$.tw. (267766)

6 hepatocyte growth factor.tw. (8585)

7 (FGF or fibroblast growth factor$).tw. (33674)

8 (PDGF or platelet derived growth factor$).tw. (20842)

9 (EGF or epidermal growth factor$).tw. (61625)

10 (IGF-I or insulin-like growth factor$ or IGF1).tw. (45386)

11 (TGF-a or transforming growth factor alfa or TGFa).tw. (306)

12 (TGF-b or transforming growth factor beta or TGFb).tw. (30559)

13 (EGFR or epidermal growth factor receptor$).tw. (46446)

14 (VEGF or vascular endothelial growth factor$).tw. (58203)

15 exp Luteinizing Hormone/bl (Blood] (24870)

16 leptin$.tw. (26783)

17 exp Progesterone/bl (Blood] (18699)

18 Proteolytic enzyme$.tw. (9992)

19 exp matrix metalloproteinase 1/ or exp matrix metalloproteinase 2/ or exp matrix metalloproteinase 3/ or exp matrix metalloproteinase 9/ (24504)

20 matrix metalloproteinase$.tw. (37055)

21 MMP$.tw. (47849)

22 TIMP$.tw. (11419)

23 exp "tissue inhibitor of metalloproteinase-1"/ or exp "tissue inhibitor of metalloproteinase-2"/ (6447)

24 exp Glycoproteins/ (662211)

25 (Ca-125 or Ca125 or cancer antigen 125).tw. (7058)

26 (Ca-19-9 or Ca19-9 or cancer antigen 19-9).tw. (4399)

27 (PP 14 or PP14).tw. (232)

28 serum placental protein$.tw. (34)

29 exp Follistatin/ (1180)

30 Osteopontin$.tw. (7267)

31 exp intercellular adhesion molecule-1/ or exp selectins/ (26225)

32 soluble intercellular adhesion.tw. (1663)

33 Soluble adhesion molecule$.tw. (795)

34 sICAM.tw. (2374)

35 sVCAM$.tw. (1360)

36 (sEcadherin or soluble E-cadherin).tw. (97)

37 (sEselectin or soluble E-selectin).tw. (713)

38 exp t-lymphocytes/ or exp natural killer t-cells/ (284378)

39 Immune cells alteration$.tw. (1)

40 (T helper$ or T supressor$ or T helper$ T supressor$ ratio).tw. (22494)

41 Total complement level$.tw. (24)

42 Autoantibodies.tw. (35161)

43 exp Antibodies, Antiphospholipid/ (7759)

44 Anti-endometrial.tw. (22)

45 Antiphospholipid$.tw. (10351)

46 exp hla antigens/ or exp hla-a1 antigen/ or exp hla-a2 antigen/ (66724)

47 (HLA or human leucocyte antigen$).tw. (83856)

48 Anti-laminin-1.tw. (33)

49 Anti-thyroid.tw. (1478)

50 Anti-Thomsen Friedenreich antigen$.tw. (8)

51 Anti-transferrin.tw. (284)

52 Anti-LDL.tw. (183)

53 (Anti-2HSG or Heremans-Schmidt glycoprotein).tw. (3)

54 interleukin$.tw. (184697)

55 (MCP-I or monocyte chemoattractant protein-I).tw. (46)

56 (MIF or migration inhibitory factor$).tw. (4718)

57 (TNF-a or tumour necrosis factor$ alfa).tw. (1428)

58 Fas ligand$.tw. (6204)

59 Endometrial marker$.tw. (11)

60 CAMs.tw. (1823)

61 cell adhesion molecule$.tw. (22033)

62 exp Integrins/ (46487)

63 Integrin$.tw. (42447)

64 Selectin$.tw. (58540)

65 Cadherin$.tw. (22688)

66 Aromatase P450.tw. (182)

67 estrogen receptor$.tw. (41210)

68 progesterone receptor$.tw. (17437)

69 MTMMP$.tw. (7)

70 cyr61.tw. (620)

71 exp Cysteine-Rich Protein 61/ (425)

72 cysteine-rich heparin-binding protein$.tw. (9)

73 (ANXA 1 or ANXA1).tw. (355)

74 (Annexin 1 or Annexin1).tw. (358)

75 (PGP 9?5 or PGP9?5 or protein gene product$).tw. (2190)

76 serum marker$.tw. (5721)

77 neural marker$.tw. (1026)

78 cell surface marker$.tw. (4751)

79 inflammatory marker$.tw. (12244)

80 microarray$.tw. (81764)

81 microRNA$.tw. (35967)

82 proteomic$.tw. (49911)

83 genomic$.tw. (205064)

84 (endometri$ adj2 biops$).tw. (3518)

85 Follistatin$.tw. (1762)

86 Vascular Endothelial Growth Factor A/ (38477)

87 Vitamin D-Binding Protein/ (1356)

88 exp Cytokines/ (575020)

89 exp interleukins/ or exp interleukin-1/ or exp interleukin-6/ or exp interleukin-8/ or exp interleukin-12/ or exp interleukin-13/ (197567)

90 exp Epidermal Growth Factor/ (21875)

91 exp Fibroblast Growth Factors/ (26259)

92 Platelet-Derived Growth Factor/ (11355)

93 Keratin-19/ (1179)

94 exp Clinical Laboratory Techniques/ (2203416)

95 (Luteinizing Hormone$ or LH).tw. (57796)

96 cytokeratin-19.tw. (1538)

97 (VDBP or vitamin D-binding protein$).tw. (1262)

98 urinary peptide$.tw. (148)

99 VDBP-Cr.tw. (1)

100 urinary VDBP corrected for creatinine expression.tw. (1)

101 urinary marker$.tw. (679)

102 or/1-101 (4283825)

103 Endometriosis/di (Diagnosis] (3449)

104 102 or 103 (4286552)

105 exp Endometriosis/ (17833)

106 Endometrio$.tw. (22478)

107 105 or 106 (26003)

108 104 and 107 (10936)

109 (animals not (humans and animals)).sh. (4004321)

110 108 not 109 (10539)

111 (201501$ or 201502$ or 201503$ or 201504$).ed. (322721)

112 110 and 111 (215)

Appendix 3. Search strategy for CENTRAL (OVID platform)

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <July 2015 (3.09.2015)>

Search Strategy:

---

1 (biomarker$ or marker$).tw. (23692)

2 Laboratory Test$.tw. (2793)

3 growth factor$.tw. (5448)

4 scatter factor$.tw. (8)

5 cytokine$.tw. (6264)

6 hepatocyte growth factor.tw. (111)

7 (FGF or fibroblast growth factor$).tw. (433)

8 (PDGF or platelet derived growth factor$).tw. (250)

9 (EGF or epidermal growth factor$).tw. (1077)

10 (IGF-I or insulin-like growth factor$ or IGF1).tw. (2132)

11 (TGF-a or transforming growth factor alfa or TGFa).tw. (519)

12 (TGF-b or transforming growth factor beta or TGFb).tw. (236)

13 (EGFR or epidermal growth factor receptor$).tw. (1905)

14 (VEGF or vascular endothelial growth factor$).tw. (1532)

15 exp Luteinizing Hormone/bl (Blood] (151)

16 leptin$.tw. (1399)

17 exp Progesterone/bl (Blood] (58)

18 Proteolytic enzyme$.tw. (136)

19 exp matrix metalloproteinase 1/ or exp matrix metalloproteinase 2/ or exp matrix metalloproteinase 3/ or exp matrix metalloproteinase 9/ (292)

20 matrix metalloproteinase$.tw. (676)

21 MMP$.tw. (905)

22 TIMP$.tw. (229)

23 exp "tissue inhibitor of metalloproteinase-1"/ or exp "tissue inhibitor of metalloproteinase-2"/ (101)

24 exp Glycoproteins/ (10108)

25 (Ca-125 or Ca125 or cancer antigen 125).tw. (305)

26 (Ca-19-9 or Ca19-9 or cancer antigen 19-9).tw. (71)

27 (PP 14 or PP14).tw. (23)

28 serum placental protein$.tw. (6)

29 exp Follistatin/ (13)

30 Osteopontin$.tw. (80)

31 exp intercellular adhesion molecule-1/ or exp selectins/ (929)

32 soluble intercellular adhesion.tw. (256)

33 Soluble adhesion molecule$.tw. (89)

34 sICAM.tw. (319)

35 sVCAM$.tw. (223)

36 (sEcadherin or soluble E-cadherin).tw. (4)

37 (sEselectin or soluble E-selectin).tw. (99)

38 exp t-lymphocytes/ or exp natural killer t-cells/ (2645)

39 Immune cells alteration$.tw. (1)

40 (T helper$ or T supressor$ or T helper$ T supressor$ ratio).tw. (445)

41 Total complement level$.tw. (0)

42 Autoantibodies.tw. (428)

43 exp Antibodies, Antiphospholipid/ (85)

44 Anti-endometrial.tw. (0)

45 Antiphospholipid$.tw. (152)

46 exp hla antigens/ or exp hla-a1 antigen/ or exp hla-a2 antigen/ (563)

47 (HLA or human leucocyte antigen$).tw. (1724)

48 Anti-laminin-1.tw. (0)

49 Anti-thyroid.tw. (49)

50 Anti-Thomsen Friedenreich antigen$.tw. (0)

51 Anti-transferrin.tw. (0)

52 Anti-LDL.tw. (3)

53 (Anti-2HSG or Heremans-Schmidt glycoprotein).tw. (0)

54 interleukin$.tw. (7276)

55 (MCP-I or monocyte chemoattractant protein-I).tw. (0)

56 (MIF or migration inhibitory factor$).tw. (75)

57 (TNF-a or tumour necrosis factor$ alfa).tw. (3923)

58 Fas ligand$.tw. (47)

59 Endometrial marker$.tw. (2)

60 CAMs.tw. (53)

61 cell adhesion molecule$.tw. (568)

62 exp Integrins/ (781)

63 Integrin$.tw. (248)

64 Selectin$.tw. (2183)

65 Cadherin$.tw. (71)

66 Aromatase P450.tw. (3)

67 estrogen receptor$.tw. (1252)

68 progesterone receptor$.tw. (531)

69 MTMMP$.tw. (0)

70 cyr61.tw. (1)

71 exp Cysteine-Rich Protein 61/ (1)

72 cysteine-rich heparin-binding protein$.tw. (0)

73 (ANXA 1 or ANXA1).tw. (3)

74 (Annexin 1 or Annexin1).tw. (2)

75 (PGP 9?5 or PGP9?5 or protein gene product$).tw. (18)

76 serum marker$.tw. (411)

77 neural marker$.tw. (9)

78 cell surface marker$.tw. (46)

79 inflammatory marker$.tw. (1739)

80 microarray$.tw. (501)

81 microRNA$.tw. (103)

82 proteomic$.tw. (176)

83 genomic$.tw. (526)

84 (endometri$ adj2 biops$).tw. (464)

85 Follistatin$.tw. (26)

86 Vascular Endothelial Growth Factor A/ (560)

87 Vitamin D-Binding Protein/ (18)

88 exp Cytokines/ (13960)

89 exp interleukins/ or exp interleukin-1/ or exp interleukin-6/ or exp interleukin-8/ or exp interleukin-12/ or exp interleukin-13/ (4413)

90 exp Epidermal Growth Factor/ (91)

91 exp Fibroblast Growth Factors/ (197)

92 Platelet-Derived Growth Factor/ (99)

93 Keratin-19/ (19)

94 exp Clinical Laboratory Techniques/ (35164)

95 (Luteinizing Hormone$ or LH).tw. (2935)

96 cytokeratin-19.tw. (25)

97 (VDBP or vitamin D-binding protein$).tw. (44)

98 urinary peptide$.tw. (8)

99 VDBP-Cr.tw. (0)

100 urinary VDBP corrected for creatinine expression.tw. (0)

101 urinary marker$.tw. (67)

102 or/1-101 (90390)

103 Endometriosis/di (Diagnosis] (6)

104 102 or 103 (90394)

105 exp Endometriosis/ (469)

106 Endometrio$.tw. (1026)

107 105 or 106 (1067)

108 104 and 107 (226)

109 (animals not (humans and animals)).sh. (1)

110 108 not 109 (226)

Appendix 4. Search strategy for EMBASE (OVID platform)

Database: EMBASE (Ovid) <1980 to 2015 Week 07 (16.02.2015)>

Search strategy:

---

1 Laboratory Test$.tw. (41662)

2 growth factor$.tw. (318593)

3 scatter factor$.tw. (1388)

4 cytokine$.tw. (322134)

5 hepatocyte growth factor.tw. (9594)

6 (FGF or fibroblast growth factor$).tw. (37191)

7 (PDGF or platelet derived growth factor$).tw. (23530)

8 (EGF or epidermal growth factor$).tw. (69553)

9 (IGF-I or insulin-like growth factor$ or IGF1).tw. (49806)

10 (TGF-a or transforming growth factor alfa or TGFa).tw. (542)

11 (TGF-b or transforming growth factor beta or TGFb).tw. (30820)

12 (EGFR or epidermal growth factor receptor$).tw. (64664)

13 (VEGF or vascular endothelial growth factor$).tw. (73191)

14 exp luteinizing hormone/ec (Endogenous Compound] (21924)

15 leptin$.tw. (32576)

16 exp progesterone blood level/ or exp progesterone urine level/ (6285)

17 Proteolytic enzyme$.tw. (9643)

18 exp matrix metalloproteinase/ (19364)

19 matrix metalloproteinase$.tw. (41445)

20 MMP$.tw. (58466)

21 TIMP$.tw. (14174)

22 exp "tissue inhibitor of metalloproteinase 2"/ (4824)

23 exp "tissue inhibitor of metalloproteinase 1"/ (8779)

24 exp glycoprotein/ec (Endogenous Compound] (246077)

25 (Ca-125 or Ca125 or cancer antigen 125).tw. (9536)

26 (Ca-19-9 or Ca19-9 or cancer antigen 19-9).tw. (6054)

27 (PP 14 or PP14).tw. (244)

28 serum placental protein$.tw. (43)

29 exp follistatin/ (2148)

30 Osteopontin$.tw. (8475)

31 exp intercellular adhesion molecule 1/ (32066)

32 exp selectin/ (3082)

33 soluble intercellular adhesion.tw. (1788)

34 Soluble adhesion molecule$.tw. (919)

35 sICAM.tw. (2888)

36 sVCAM$.tw. (1793)

37 (sEcadherin or soluble E-cadherin).tw. (120)

38 (sEselectin or soluble E-selectin).tw. (822)

39 exp T lymphocyte/ (374675)

40 exp natural killer T cell/ (5800)

41 Immune cells alteration$.tw. (6)

42 (T helper$ or T supressor$ or T helper$ T supressor$ ratio).tw. (24786)

43 Total complement level$.tw. (20)

44 Autoantibodies.tw. (42037)

45 exp phospholipid antibody/ (9920)

46 Anti-endometrial.tw. (23)

47 Antiphospholipid$.tw. (13777)

48 exp HLA antigen/ (81011)

49 exp HLA A1 antigen/ (597)

50 exp HLA A2 antigen/ (3288)

51 (HLA or human leucocyte antigen$).tw. (104497)

52 Anti-laminin-1.tw. (43)

53 Anti-thyroid.tw. (1873)

54 Anti-Thomsen Friedenreich antigen$.tw. (5)

55 Anti-transferrin.tw. (290)

56 Anti-LDL.tw. (186)

57 (Anti-2HSG or Heremans-Schmidt glycoprotein).tw. (4)

58 interleukin$.tw. (199692)

59 (MCP-I or monocyte chemoattractant protein-I).tw. (112)

60 (MIF or migration inhibitory factor$).tw. (5063)

61 (TNF-a or tumour necrosis factor$ alfa).tw. (5998)

62 Fas ligand$.tw. (6708)

63 Endometrial marker$.tw. (18)

64 CAMs.tw. (2100)

65 cell adhesion molecule$.tw. (24039)

66 exp integrin/ (29036)

67 Integrin$.tw. (48293)

68 Selectin$.tw. (67300)

69 Cadherin$.tw. (27150)

70 Aromatase P450.tw. (202)

71 estrogen receptor$.tw. (46656)

72 progesterone receptor$.tw. (19861)

73 MTMMP$.tw. (15)

74 cyr61.tw. (755)

75 exp cysteine rich protein 61/ (753)

76 cysteine-rich heparin-binding protein$.tw. (12)

77 (ANXA 1 or ANXA1).tw. (452)

78 (Annexin 1 or Annexin1).tw. (425)

79 (PGP 9?5 or PGP9?5 or protein gene product$).tw. (2620)

80 serum marker$.tw. (7720)

81 neural marker$.tw. (1119)

82 cell surface marker$.tw. (5851)

83 inflammatory marker$.tw. (17339)

84 microarray$.tw. (101846)

85 microRNA$.tw. (40082)

86 proteomic$.tw. (55191)

87 genomic$.tw. (217184)

88 (endometri$ adj2 biops$).tw. (4369)

89 Follistatin$.tw. (1945)

90 exp vasculotropin/ (69810)

91 Vascular Endothelial Growth Factor A.tw. (2275)

92 exp vitamin D binding protein/ (2064)

93 exp cytokine/ (1034772)

94 exp interleukin derivative/ (2790)

95 exp interleukin 1/ (48499)

96 exp interleukin 6/ (136328)

97 exp interleukin 8/ (48884)

98 exp interleukin 12/ (31842)

99 exp interleukin 13/ (13584)

100 exp epidermal growth factor/ (32130)

101 exp fibroblast growth factor/ (13858)

102 cytokeratin 19/ (3601)

103 platelet derived growth factor/ (18930)

104 cytokeratin-19.tw. (1918)

105 (VDBP or vitamin D-binding protein$).tw. (1413)

106 urinary peptide$.tw. (174)

107 VDBP-Cr.tw. (1)

108 urinary VDBP corrected for creatinine expression.tw. (1)

109 urinary marker$.tw. (830)

110 exp blood analysis/ (118854)

111 exp endometrium biopsy/ (4988)

112 exp urinalysis/ or exp biological marker/ (210153)

113 (biomarker or biomarkers).tw. (159748)

114 or/1-113 (2734501)

115 endometriosis/di (Diagnosis] (4979)

116 114 or 115 (2738583)

117 exp endometriosis/ (25923)

118 Endometriosis.tw. (22110)

119 117 or 118 (27911)

120 116 and 119 (10326)

121 Animal/ not Human/ (1204497)

122 120 not 121 (10279)

Additional search February 2015 - May 2015

Embase <1980 to 2015 Week 35 (3.09.2015)>

Search Strategy:

---

1 Laboratory Test$.tw. (44290)

2 growth factor$.tw. (335543)

3 scatter factor$.tw. (1407)

4 cytokine$.tw. (343623)

5 hepatocyte growth factor.tw. (10104)

6 (FGF or fibroblast growth factor$).tw. (39159)

7 (PDGF or platelet derived growth factor$).tw. (24591)

8 (EGF or epidermal growth factor$).tw. (73599)

9 (IGF-I or insulin-like growth factor$ or IGF1).tw. (51838)

10 (TGF-a or transforming growth factor alfa or TGFa).tw. (583)

11 (TGF-b or transforming growth factor beta or TGFb).tw. (32580)

12 (EGFR or epidermal growth factor receptor$).tw. (71526)

13 (VEGF or vascular endothelial growth factor$).tw. (79087)

14 exp luteinizing hormone/ec (Endogenous Compound] (22767)

15 leptin$.tw. (34921)

16 exp progesterone blood level/ or exp progesterone urine level/ (6534)

17 Proteolytic enzyme$.tw. (9903)

18 exp matrix metalloproteinase/ (20462)

19 matrix metalloproteinase$.tw. (44380)

20 MMP$.tw. (63208)

21 TIMP$.tw. (15146)

22 exp "tissue inhibitor of metalloproteinase 2"/ (5136)

23 exp "tissue inhibitor of metalloproteinase 1"/ (9381)

24 exp glycoprotein/ec (Endogenous Compound] (260024)

25 (Ca-125 or Ca125 or cancer antigen 125).tw. (10051)

26 (Ca-19-9 or Ca19-9 or cancer antigen 19-9).tw. (6446)

27 (PP 14 or PP14).tw. (243)

28 serum placental protein$.tw. (44)

29 exp follistatin/ (2283)

30 Osteopontin$.tw. (9173)

31 exp intercellular adhesion molecule 1/ (33492)

32 exp selectin/ (3217)

33 soluble intercellular adhesion.tw. (1865)

34 Soluble adhesion molecule$.tw. (944)

35 sICAM.tw. (3049)

36 sVCAM$.tw. (1924)

37 (sEcadherin or soluble E-cadherin).tw. (125)

38 (sEselectin or soluble E-selectin).tw. (861)

39 exp T lymphocyte/ (394405)

40 exp natural killer T cell/ (6310)

41 Immune cells alteration$.tw. (6)

42 (T helper$ or T supressor$ or T helper$ T supressor$ ratio).tw. (26082)

43 Total complement level$.tw. (20)

44 Autoantibodies.tw. (44153)

45 exp phospholipid antibody/ (10362)

46 Anti-endometrial.tw. (25)

47 Antiphospholipid$.tw. (14399)

48 exp HLA antigen/ (83748)

49 exp HLA A1 antigen/ (622)

50 exp HLA A2 antigen/ (3409)

51 (HLA or human leucocyte antigen$).tw. (109332)

52 Anti-laminin-1.tw. (43)

53 Anti-thyroid.tw. (2059)

54 Anti-Thomsen Friedenreich antigen$.tw. (7)

55 Anti-transferrin.tw. (297)

56 Anti-LDL.tw. (191)

57 (Anti-2HSG or Heremans-Schmidt glycoprotein).tw. (4)

58 interleukin$.tw. (210083)

59 (MCP-I or monocyte chemoattractant protein-I).tw. (114)

60 (MIF or migration inhibitory factor$).tw. (5342)

61 (TNF-a or tumour necrosis factor$ alfa).tw. (6488)

62 Fas ligand$.tw. (6895)

63 Endometrial marker$.tw. (18)

64 CAMs.tw. (2198)

65 cell adhesion molecule$.tw. (25207)

66 exp integrin/ (30330)

67 Integrin$.tw. (50938)

68 Selectin$.tw. (71624)

69 Cadherin$.tw. (29496)

70 Aromatase P450.tw. (207)

71 estrogen receptor$.tw. (49530)

72 progesterone receptor$.tw. (21068)

73 MTMMP$.tw. (16)

74 cyr61.tw. (822)

75 exp cysteine rich protein 61/ (829)

76 cysteine-rich heparin-binding protein$.tw. (12)

77 (ANXA 1 or ANXA1).tw. (500)

78 (Annexin 1 or Annexin1).tw. (440)

79 (PGP 9?5 or PGP9?5 or protein gene product$).tw. (2760)

80 serum marker$.tw. (8158)

81 neural marker$.tw. (1234)

82 cell surface marker$.tw. (6222)

83 inflammatory marker$.tw. (19492)

84 microarray$.tw. (110181)

85 microRNA$.tw. (47554)

86 proteomic$.tw. (60599)

87 genomic$.tw. (233444)

88 (endometri$ adj2 biops$).tw. (4589)

89 Follistatin$.tw. (2081)

90 exp vasculotropin/ (74115)

91 Vascular Endothelial Growth Factor A.tw. (2526)

92 exp vitamin D binding protein/ (2196)

93 exp cytokine/ (1094317)

94 exp interleukin derivative/ (3281)

95 exp interleukin 1/ (50850)

96 exp interleukin 6/ (147379)

97 exp interleukin 8/ (52281)

98 exp interleukin 12/ (33479)

99 exp interleukin 13/ (14685)

100 exp epidermal growth factor/ (33057)

101 exp fibroblast growth factor/ (14499)

102 cytokeratin 19/ (3886)

103 platelet derived growth factor/ (19655)

104 cytokeratin-19.tw. (2030)

105 (VDBP or vitamin D-binding protein$).tw. (1520)

106 urinary peptide$.tw. (189)

107 VDBP-Cr.tw. (1)

108 urinary VDBP corrected for creatinine expression.tw. (1)

109 urinary marker$.tw. (883)

110 exp blood analysis/ (124468)

111 exp endometrium biopsy/ (5197)

112 exp urinalysis/ or exp biological marker/ (232619)

113 (biomarker or biomarkers).tw. (182609)

114 or/1-113 (2911073)

115 endometriosis/di (Diagnosis] (5173)

116 114 or 115 (2915302)

117 exp endometriosis/ (27433)

118 Endometriosis.tw. (23449)

119 117 or 118 (29532)

120 116 and 119 (10922)

121 Animal/ not Human/ (1261620)

122 120 not 121 (10862)

123 (201501$ or 201502$ or 201503$ or 201504$).em. (49200)

124 122 and 123 (34)

Appendix 5. Search strategy for CINAHL database (EBSCO platform)

Database: CINAHL Plus with Full Text (EBSCOhost) <1980 to 20.04.2015>

Search strategy:

#	Query	Results
S97	S3 AND S96	1131
S96	S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR S62 OR S63 OR S64 OR S65 OR S66 OR S67 OR S68 OR S69 OR S70 OR S71 OR S72 OR S73 OR S74 OR S75 OR S76 OR S77 OR S78 OR S79 OR S80 OR S81 OR S82 OR S83 OR S84 OR S85 OR S86 OR S87 OR S88 OR S89 OR S90 OR S91 OR S92 OR S93 OR S94 OR S95	341775
S95	TX urinary peptide*	1598
S94	TX (VDBP or vitamin D-binding protein*)	134
S93	TX cytokeratin-19	109
S92	TX (Luteinizing Hormone* or LH)	18041
S91	(MH "Diagnosis, Laboratory+")	101773
S90	"Keratin-19"	2
S89	(MH "Platelet-Derived Growth Factor")	394
S88	(MH "Epidermal Growth Factors")	1264
S87	(MH "Interleukins")	6584

(Continued)

S86	(MH "Cytokines")	6860
S85	TX Vitamin D-Binding Protein	131
S84	(MH "Vascular Endothelial Growth Factor A")	194
S83	TX (endometri* N2 biops*)	432
S82	TX (endometri* adj2 biops*)	0
S81	TX genomic$	7487
S80	TX proteomic*	2434
S79	TX microRNA	824
S78	TX microarray	3123
S77	TX (PGP 95 or PGP95 or protein gene product*)	9925
S76	TX (Annexin 1 or Annexin1)	472
S75	TX (ANXA 1 or ANXA1)	41
S74	TX cysteine-rich heparin-binding protein*	12
S73	(MH "Protein Array Analysis")	73
S72	TX cyr61	34
S71	TX MTMMP*	0
S70	TX progesterone receptor*	1927
S69	TX estrogen receptor*	5193
S68	TX Aromatase P450	38
S67	TX Cadherin*	900
S66	TX Selectin*	28411
S65	TX Integrin*	1587
S64	TX cell adhesion molecule*	1578
S63	TX CAMs	550
S62	TX Endometrial marker*	54
S61	TX Fas ligand	338
S60	TX (TNF-a or tumour necrosis factor* alfa)	1489
S59	TX (MIF or migration inhibitory factor*)	399

(Continued)

S58	TX (MCP-I or monocyte chemoattractant protein-I)	13
S57	TX interleukin	13809
S56	TX (Anti-2HSG or Heremans-Schmidt glycoprotein)	7
S55	TX Anti-LDL	9
S54	TX Anti-transferrin	3
S53	TX Anti-Thomsen Friedenreich antigen*	1
S52	TX Anti-thyroid	109
S51	TX Anti-laminin-1	15
S50	TX (HLA or human leucocyte antigen*)	4202
S49	(MM "HLA Antigens")	638
S48	TX Antiphospholipid*	1249
S47	TX Anti-endometrial	34
S46	(MH "Antibodies/BL/DU")	1294
S45	TX Autoantibodies	4385
S43	TX Total complement level	3
S42	TX (T helper* or T supressor*)	2341
S41	TX Immune cells alteration*	24
S40	TX natural killer t-cells	669
S39	(MM "T Lymphocytes")	2404
S38	TX (sEselectin or soluble E-selectin)	91
S37	TX (sEcadherin or soluble E-cadherin)	8
S36	TX sVCAM	100
S35	TX sICAM	173
S34	TX Soluble adhesion molecule	368
S33	TX soluble intercellular adhesion	237
S32	(MM "Cell Adhesion Molecules")	52
S31	TX Osteopontin*	416
S30	TX Follistatin	74

(Continued)

S29	TX serum placental protein*	11
S28	TX (Ca-19-9 or Ca19-9 or cancer antigen 19-9)	262
S27	TX (Ca-125 or Ca125 or cancer antigen 125)	831
S26	(MM "Glycoproteins/BL/DU")	224
S25	TX tissue inhibitor of metalloproteinase	423
S24	TX TIMP*	1845
S23	TX MMP*	4244
S22	TX matrix metalloproteinase*	3325
S21	TX Proteolytic enzyme*	1461
S20	(MM "Progesterone/BL/DU")	51
S19	TX leptin*	3258
S18	(MM "Luteinizing Hormone/BL/DU")	38
S17	TX (VEGF or vascular endothelial growth factor*)	7166
S16	TX (EGFR or epidermal growth factor receptor*)	6188
S15	TX (TGF-b or transforming growth factor beta or TGFb)	2972
S14	TX (TGF-a or transforming growth factor alfa or TGFa)	464
S13	TX (IGF-I or insulin-like growth factor* or IGF1)	3588
S12	TX (EGF or epidermal growth factor*)	6250
S11	TX (PDGF or platelet derived growth factor*)	3195
S10	TX (FGF or fibroblast growth factor*)	3395
S9	TX hepatocyte growth factor*	880
S8	TX cytokine*	20821
S7	TX scatter factor*	1864
S6	TX growth factor*	76163
S5	TX Laboratory Test*	82732
S4	TX (biomarker* or marker*)	84857
S3	S1 OR S2	2841
S2	TX Endometrio*	2841

(Continued)

S1	(MM "Endometriosis")	889
S4	TX (biomarker* or marker*)	61,794
S3	S1 OR S2	2,174
S2	TX Endometrio*	2,174
S1	(MM "Endometriosis")	1,306

Appendix 6. Search strategy for other databases

Search for clinical studies

Database: Web of Science Core Collection (Thomson Reuters) <1900 to Present (20.04.2015)>

Search strategy:

1. Topic=(endometrio*) AND Topic=(diagnos* OR test* OR marker* OR biomarker*); Timespan=All Years (7425)

Database: PsycINFO (Ovid) <1806 to April Week 2 2015 (20.04.2015)>

Search strategy:

1. endometriosis.tw. (174)

Database: LILACS <20.04.2015>

Search strategy:

1. (tw:(endometriosis)) AND (tw:(diagnos*)) (420)

Database: OAIster (WorldCat.org) <20.04.2015>

Search strategy:

1. endometriosis and (marker* or biomarker*) (11)

2. endometriosis and diagnos* (446)

Database: TRIP <20.04.2015>

Search strategy:

1. (endometriosis and diagnos*) (1648)

Searches of trial registers for ongoing and registered trials

Database: 'ClinicalTrials.gov', a service of the US national Institute of Health

Search strategy:

1. endometriosis (220)

2. endometriosis AND diagnosis (22)

Database: WHO International Clinical Trials Registry Platform (ICTRP) <20.04.2015>

Search strategy:

1. endometriosis (523)

Searches for the reviews as potential source of references

Database: MEDION <10.01.2014>

Search strategy:

ICP Code female genital system (including breast), Signssymp medical imaging, laboratory tests, histology and cytology, endoscopy and laparoscopy. Filter: systematic reviews of diagnostic studies (2)

Database: DARE (CRD) <20.04.2015>

Search strategy:

1. endometriosis (99)

PubMed, a 'Systematic Review' search under the 'Clinical Queries' link <20.04.2015>

Search strategy:

(endometriosis) AND systematic(sb] (418)

Category: Diagnosis; Scope: Broad

Searches for the papers recently published and not yet indexed in the major databases

Search engine: PubMed <20.10.2014 to 20.04.2015>

Search strategy:

1. marker (14979) 2. test (61151) 3. diagnos* (69743) 4. biomarker (10806) 5. or/1-4 (7943) Filters: Publication date from 2014/10/20 to 2015/04/20	Index test(s) set
6. Endometriosis (584) Filters: Publication date from 2014/10/20 to 2015/04/20	Target condition set
7. 5 and 6 (267) Filters: Publication date from 2014/10/20 to 2015/04/20	Combined sets

Appendix 7. Summary of findings table 2: Blood biomarkers that do not distinguish between women with and without endometriosis

Review question	Which blood biomarkers are unlikely to serve a basis of the diagnostic test for endometriosis?	
Importance	Biomarkers that do not show differential expression in women with and without endometriosis, are unlikely to be diagnostically useful. Information regarding negative trials can focus research on better diagnostic targets.The biomarkers that display conflicting results (distinguish women with and without endometriosis in some but not all studies) can be identified and reported on. Studies that did not show differential expression of a biomarker in endometriosis but were adequately designed and that met inclusion criteria for this review were included	
Patients	Reproductive-aged women with suspected endometriosis or persistent ovarian mass, or women undergoing infertility work-up/gynaecological laparoscopy	
Settings	Hospitals (public or private of any level), outpatient clinics (general gynaecology, reproductive medicine, pelvic pain) or research laboratory	
Reference standard	Visualisation of endometriosis at surgery (laparoscopy or laparotomy) with or without histological confirmation	
Study design	Cross-sectional of a single-gate design (N = 39) or two-gate design (N = 41); unable to determine if single- or two-gate design for 2 studies; prospective enrolment; a single study could assess more than one test	
Risk of bias	Overall judgement	Poor quality (no studies had 'low risk' assessment in all 4 domains)
	Patient selection bias	High risk: 50 studies; unclear risk: 25 studies; low risk: 7 studies
	Index test interpretation bias	High risk: 80 studies; unclear risk: 2 studies; low risk: 0 studies
	Reference standard interpretation bias	High risk: 0 studies; unclear risk: 29 studies; low risk: 53 studies
	Flow and timing selection bias	High risk: 12 studies; unclear risk: 7 studies; low risk: 63 studies
Applicability concerns	Concerns regarding patient selection	High concern: 45 studies, unclear concern: 5 studies; low concern: 32 studies
	Concerns regarding index test	High concern: 0 studies, unclear concern: 1 study; low concern: 81 studies
	Concerns regarding reference standard	High concern: 0 studies; unclear concern: 0 studies; low concern: 82 studies

(Continued)

Biomarker	Number of participants		Units	Outcome measures			rASRM stage	Menstrual cycle phase	Reference
	Endometriosis	Controls		Expression in endometriosis	Expression in controls	P value			
1. Angiogenesis and growth factors and their receptors									
1.1. angiogenic activity of serum	52	32	mean ± SD, number of newly formed blood vessels	13.57 ± 1.68	13.43 ± 1.29	NS	I-IV	follicular	Barcz 2002
1.2. CAC (circulating angiogenic cells)	42	22	mean ± SEM, %	0.084 ± 0.007	0.072 ± 0.009	0.32	I-IV	any	Webster 2013
1.3. EGF (epidermal growth factor)	36	36	mean ± SD, pg/ml	NS					Philippoussis 2004
				497.8 ± 99.1	490.5 ± 191.2				
				NS					
				493.4 ± 180.4	494.4 ± 169.8				
1.4. sEGF-R (soluble epidermal growth factor-receptor)	28	20		below detection limit of assay	below detection limit of assay		I-IV	n/a	Matalliotakis 2003a
1.5. sFlt-1 (sVEGFR-1] (soluble fms-like tyrosine kinase or (variant of VEGF receptor 1])	46	24	mean ± SEM, pg/ml	119.89 ± 5.43	112.30 ± 5.23	NS	I-IV	follicular/ luteal	Cho 2007
1.6.a.glycodelin A (PP14 or PAEP] ((placental protein 14 or progestogen-associated endometrial protein])	33	17	mean ± SD, ng/ml	44.21 ± 45.67	34.55 ± 58.13	0.19	I-IV	follicular	Drosdzol-Cop 2012a
1.6.b. PAEP (glycodelin] (Progestagen-associated endometrial protein)	36	19	mean ± SE, U/ml	follicular cycle phase:	follicular cycle phase:	NS	I-IV	follicular/ luteal	Joshi 1986
				5±2 (rASRM I-II);	9 ± 3;				
				10±2 (rASRM III-IV);	luteal cycle phase: 23 ± 6				

(Continued)

				luteal cycle phase: 18±4 (rASRM I-II); 23±4 (rASRM III-IV)					
1.6.c. glycodelin A (PP14 or PAEP] ((placental protein 14 or progestogen-associated endometrial protein])	69	32		below detection limit of assay	below detection limit of assay		I-IV	n/a	Paiva 2014
1.7. HGF (hepatocyte growth factor)	37	21	mean ± SD, pg/ml	6879 ± 53.3	675.19 ± 40.9	NS	I-IV	follicular/ luteal	Khan 2006
1.8.a. sIGF-1 (soluble Insulin-like growth factor-1)	28	20	%OD increase over background ± S.E.M	275 ± 50	300 ± 33.3	NS	I-IV	n/a	Matalliotakis 2003a
1.8.b. IGF-1 (insulin-like growth factor-1)	77	71	mean ± SD, ng/ml	crude values		NS	I-IV	follicular/ luteal	Steff 2004b
				follicular cycle phase: 269.1 ± 90.3; luteal cycle phase: 290.2 ± 93.3	follicular cycle phase: 270.1 ± 91.7; luteal cycle phase: 271.6 ± 76.8				
				adjusted values (age, BMI using a univariate general linear model]		NS			
				follicular cycle phase: 274.8 ± 87.3; luteal cycle phase: 296.8 ± 82.9	follicular cycle phase: 264.0 ± 87.4; luteal cycle phase: 264.4 ± 83.0				

(Continued)

1.9. IGF-2 (insulin-like growth factor-2)	29	15	mean ± SEM, ng/ml	406.2 ± 27.5 (rASRM I-II); 430.6±30.2 (rASRM III-IV)	442.6 ± 30.5	NS	I-IV	follicular/ luteal	Gurgan 1999
1.10.a. IGFBP-3 (insulin-like growth factor binding protein-3)	29	15	mean ± SEM, ng/ml	1256.2 ± 31 (rASRM I-II); 1210.7 ± 51.9 (rASRM III-IV)	1250.6 ± 33.4	NS	I-IV	follicular/ luteal	Gurgan 1999
1.10.b. IGFBP-3 (insulin-like growth factor binding protein-3)	36	36	mean ± SD, ng/ml	crude values		NS	I-IV	luteal	Philippoussis 2004
				48.2 ± 8.8	46.8 ± 7.9				
				adjusted values (indication for surgery, BMI, and presence of uterine leiomyoma using a univariate general linear model]		NS			
				48.2 ± 7.8	45.7 ± 7.9				
1.11.a. leptin	60	20	median (IQR), ng/ml	4.3 (5.1)	5.2 (11.2)	NS	I-IV	n/a	Ozhan 2014
1.11.b. leptin	69	32	median (IQR), pg/ml	1.48 (0.1 - 16.7)	1.03 (0.3 - 67.3)	0.95	I-IV	n/a	Paiva 2014
1.11.c. leptin	42	25	mean ± SEM, ng/ml	12.5 ± 8.4 (rASRM I-II); 11.8 ± 7.7 (rASRM III-IV)	12.5 ± 9.4	NS	I-IV	any	Vigano 2002
1.11.d. leptin	33	30	mean ± SD, µg/L	3.145 ± 0.389	2.088 ± 0.373 (tubal factor infertility) 1.963 ± 0.410 (benign ovarian cyst)	NS	I-IV	any	Wei 2005
1.12. PDGF (platelet derived growth factor)	17	23	median (IQR), pg/ml	98.7 (82.7 - 149.5)	99.7 (80.1 - 145.6)	0.682	I-II	luteal	Kalu 2007

(Continued)

1.13.a. VEGF (vascular endothelial growth factor)	46	24	mean ± SEM, pg/ml	240.92 ± 19.77	222.37 ± 26.72	NS	I-IV	follicular/ luteal	Cho 2007
1.13.b. VEGF (vascular endothelial growth factor)	10	7	median (IQR), pg/ml	0.135 (0.109 - 0.624)	0.107 (0.097 - 0.124)	0.093	II-IV	follicular	Da Silva 2014
1.13.c. VEGF (vascular endothelial growth factor)	131	146	mean ± SD, pg/ml	crude values		NS	I-IV	luteal	Gagne 2003b
				241 ± 164	221 ± 128				
				adjusted (indication for surgery, infertility, BMI, gravidity, pelvic pain and length of menses, using a univariate general linear model]		NS			
				230 ± 149	222 ± 149				
1.13.d. VEGF (vascular endothelial growth factor)	90	89	mean ± SEM, ng/l	46.7 ± 10	53.3 ± 9.3	NS	n/a	follicular/ luteal	Kianpour 2013
1.13.e. VEGF–A (vascular endothelial growth factor A)	40	20	mRNA, relative quantification	1.04 (0.6 - 1.9) DIE; 1.12 (0.5 - 1.9) endometrioma	1 (0.1 - 1.9)	0.581	n/a	follicular	Mabrouk 2012
1.13.f. VEGF (vascular endothelial growth factor)	68	70	median (IQR), pg/ml	26.32 (3.18 - 63.36)	31.80 (7.28 - 79.35)	0.22	I-IV	follicular/ luteal	Othman 2008
1.13.g. VEGF (vascular endothelial growth factor)	69	32	median (IQR), pg/ml	6.0 (0.0 - 37.0)	7.2 (0.0 - 35.7)	0.25	I-IV	n/a	Paiva 2014
2. Apoptosis markers									
2.1. annexin V	69	32	median (IQR), ng/ml	3.59 (2.50 - 13.6)	3.69 (2.6 - 5.0)	0.46	I-IV	n/a	Paiva 2014
2.2. apoptotic cells	32	30	mean ± SD, %	6.34 ± 1.94	5.68 ± 2.14	NS	I-II	peri ovulatory	Mier-Cabrera 2011
2.3. sFas (soluble Fas)	17	23	median (IQR), pg/ml	450.0 (345.9 - 723.1)	484.6 (366.4 - 557.2)	0.827	I-II	luteal	Kalu 2007

(Continued)

2.4. anti-survivin Abs (anti-survivin antibodies)	98	47	median, OD	0.078	0.119	NS	I-IV	n/a	Lamp 2012
3. Cell adhesion molecules and other matrix-related proteins									
3.1. biglycan	56	40	mean ± SD, ng/ml	13.8 ± 7.0	14.5 ± 11.8 (benign ovarian cyst); 12.9 ± 9.1 (healthy women)	0.7487	I-IV	follicular/ luteal	Kocbek 2014b
3.2.a. sICAM-1 (soluble form of intercellular adhesion molecule-1)	15	15	mean ± SD, OD	0.43 ± 0.1	0.44 ± 0.1	NS	I-IV	follicular/ luteal	De Placido 1998
3.2.b. sICAM-1 (soluble form of intercellular adhesion molecule-1)	11	9	mean, ng/ml	8.31	10.3	NS	I-II	luteal	Goluda 1998
3.2.c. sICAM-1 (soluble form of intercellular adhesion molecule-1)	69	32	median (IQR), pg/ml	181.8 (115.3 - 338.3)	176.0 (120.2 - 337.8)	0.6	I-IV	n/a	Paiva 2014
3.2.d. sICAM-1 (soluble form of intercellular adhesion molecule-1)	71	49	mean ± SD, ng/ml	257.7 ± 72.9	240.7 ± 70.7	0.21	I-IV	n/a	Somigliana 2002
3.3. sE-selectin (soluble E selectin)	11	9	mean, ng/ml	1.21	1.39	NS	I-II	luteal	Goluda 1998
3.4. MMP-9 (matrix metalloproteinase-9)	40	20	mRNA, relative quantification	0.76 (0.1 - 5.6) DIE; 1.14 (0.1 - 5.6) endometrioma	1 (0.1 - 1.9)	0.676	n/a	follicular	Mabrouk 2012
4. Cytoskeleton molecules									
4.1. CK 19 (cytokeratin-19)	44	32	mean ± SD, ng/ml	1.1 ± 1.1	1.0 ± 1.3	0.77	n/a	follicular/ luteal	Kuessel 2014
5. DNA-repair/telomere maintenance molecules									

(Continued)

5.1. TL (telomere length)	25	25	ANOVA	(F(1,33) = 284 642]		0.36	I-IV	luteal	Hapanga-ma 2008
6. Hormonal markers									
6.1. E2 (oestradiol)	25	25	mean, pg/l	window of implantation: 423; late luteal cycle phase: 217	window of implantation: 393; late luteal cycle phase: 341	NS	I-IV	luteal	Hapanga-ma 2008
6.2. FSH (follicle stimulating hormone)	28	21	mean ± SD, IU/ml	3.5 ± 0.3 - 4.5 ±0.4	4.1 ± 0.4	NS	I-IV	luteal	Lima 2006
6.3. LH (luteinizing hormone)	28	21	mean ± SD, IU/ml	2.7 ± 0.3 - 3.9 ± 0.3	2.9 ± 0.3	NS	I-IV	luteal	Lima 2006
6.4. progesterone	25	25	mean, ng/ml	window of implantation: 8.61; late luteal cycle phase: 2.54	window of implantation: 7.38; late luteal cycle phase: 4.20	NS	I-IV	luteal	Hapanga-ma 2008
7. Immune system and inflammatory markers									
7.1. Autoantibodies									
7.1.1.a. anti-endometrial Abs, MW 28 kd (anti-endometrial auto antibodies with molecular weight of 28 kilodalton	18	18	n (%)	6 (33.3%)	5 (27.8%)	NS	I-IV	n/a	Gorai 1993
7.1.1.b. anti-endometrial Abs, MW 38 kd (anti-endometrial auto antibodies with molecular weight of 38 kilodalton	18	18	n (%)	5 (27.8%)	1 (5.6%)	NS	I-IV	n/a	Gorai 1993
7.1.1.c. anti-endometrial Abs, MW 64 kd (anti-endometrial auto antibodies with molecular weight of 64 kilodalton	18	18	n (%)	7 (38.9%)	4 (22.2%)	NS	I-IV	n/a	Gorai 1993

(Continued)

7.1.1.d. AEA (anti-endometrial auto antibodies)	60	20	median (IQR), OD	0.033 (0.046)	0.036 (0.041)	NS	I-IV	n/a	Ozhan 2014
7.1.2. anti-sperm Abs (anti-sperm auto antibodies)	50	48	median ± mean deviation, fg/sperm	6.43 ± 6.98 (infertile participants); 10.16 ± 7.24 (fertile participants)	8.57 ± 17.05 (infertile participants); 8.56 ± 8.38 (fertile participants)	NS	I	luteal	Szczepanska 2001a
7.1.3. anti-ZP Abs (anti-zona pellucida auto antibodies)	50	48	median ± mean deviation, ng/oocyte	3.65 ± 5.20 (infertile participants); 4.08 ± 5.70 (fertile participants)	3.60 ± 10.44 (infertile participants); 3.98 ± 9.00 (fertile participants)	NS	I	luteal	Szczepanska 2001a
7.2. Chemokines									
7.2.1.a. MCP-1 (monocyte chemotactic protein-1)	33	17	mean ± SD, pg/ml	97.34 ± 107.36	79.14 ± 53.43	0.76	I-IV	follicular	Drosdzol-Cop 2012b
7.2.1.b. MCP-1 (monocyte chemotactic protein-1)	94	76	mean ± SE, pg/ml	321.0 ± 14.7	348.6 ± 21.4	NS	I-IV	follicular	Kim 2008
7.2.1.c. MCP-1 (monocyte chemotactic protein-1)	69	32	median (IQR), pg/ml	25.2 (14.2 - 73.9)	27.5 (14.9 - 65.9)	0.27	I-IV	n/a	Paiva 2014
7.2.1.d. MCP-1 (monocyte chemotactic protein-1)[1]	63	78	range, pg/ml	25 - 320; AUC (CIs) 0.597 (0.503 - 0.691)	10 - 350	NS	II-IV	follicular/luteal/unknown	Seeber 2008
7.3. other Cytokines									
7.3.1. Epo (erythropoietin)	33	22	mean ± SD, IU/ml	60.22 ± 9.11	30.32 ± 7.94	0.099	I-IV	n/a	Yagmur 2013
7.3.2.a. sGM-CSF (soluble granulocyte macrophage-colony stimulating factor)	28	20	%OD increase over background ± S.E.M	2.63 ± 0.25	2.75 ± 0.19	NS	I-IV	n/a	Matalliotakis 2003a

(Continued)

7.3.2.b. GM-CSF (granulocyte macrophage-colony stimulating factor)	68	70	median (IQR), pg/ml	20.51 (14.5 - 29.68)	11.97 (8.8 - 20.19)	0.51	I-IV	follicular/ luteal	Othman 2008
7.3.2.c. GM-CSF (granulocyte macrophage-colony stimulating factor)	69	32	median (IQR), pg/ml	0.58 (0.25 - 3.02)	0.59 (0.3 - 3.6)	0.84	I-IV	n/a	Paiva 2014
7.3.3.a. IFN-γ (interferon-gamma)	60: 42 - rASRM I-II; 18 - rASRM III-IV)	37	mean rank values	47.98 (rASRM I-II); 50.75 (rASRM III-IV)	49.31	0.13	I-IV	follicular	Hassa 2009
7.3.3.b. sIFN-γ (soluble interferon-gamma)	28	20	%OD increase over background ± S.E.M	1.1 ± 0.13	1.0 ± 0.12	NS	I-IV	n/a	Matalliotakis 2003a
7.3.3.c. IFN-γ (interferon-gamma)	65	33	median (range), pg/ml	1.6 (0 - 11.7)	2.1 (0 - 6.6)	0.571	I-IV	follicular/ luteal	Podgaec 2007
7.3.3.d. IFN-γ (interferon-gamma)	63	78		below detection limit of assay	below detection limit of assay		II-IV	follicular/ luteal/unknown	Seeber 2008
7.3.3.e. INF-γ (interferon-gamma)	36	35	mean ± SD, pg/ml	1.91 ± 0.18	2.05 ± 0.24	0.07	I-IV	n/a	Wu 1998
7.3.4.a. MIF (macrophage migration inhibitory factor)	60	20	median (IQR), pg/ml	901.5 (556.1)	585.0 (434.0)	NS	I-IV	n/a	Ozhan 2014
7.3.4.b. MIF (macrophage migration inhibitory factor)	69	32	median (IQR), pg/ml	544 (237 - 2354)	583 (196 - 3791)	0.51	I-IV	n/a	Paiva 2014
7.3.4.c. MIF (macrophage migration inhibitory factor)[1]	63	78	range, ng/ml	5 - 100	3 - 100	NS	II-IV	follicular/ luteal/unknown	Seeber 2008
				AUC (CIs) 0.539 (.443 - .634)					
7.3.5.a. TNF-α (tumour necrosis factor alpha)	10	7	median (IQR), pg/ml	1765,2 (425,6 - 2583,8)	221,3 (0,00 - 1910,58)	0.243	II-IV	follicular	Da Silva 2014

(Continued)

7.3.5.b. TNF-α (tumour necrosis factor alpha)	33	17	mean ± SD, pg/ml	7.40 ± 12.55	4.87 ± 1.56	0.26	I-IV	follicular	Drosd-zol-Cop 2012a
7.3.5.c. TNF-α (tumour necrosis factor alpha)	15	20	median (IQR), pg/ml	4.00 (4.0 - 4.0)	4.00 (4.0 - 4.5)	0.638	I-II	luteal	Kalu 2007
7.3.5.d. TNF-α (tumour necrosis factor alpha)	68	70	median (IQR), pg/ml	1.04 (0.84 - 1.36)	1.07 (0.89 - 1.47)	0.6	I-IV	follicular/ luteal	Othman 2008
7.3.5.e. TNF-α (tumour necrosis factor alpha)	65	33	median (range), pg/ml	2.3 (0 - 9.6)	3.7 (0 - 10.4)	0.188	I-IV	follicular/ luteal	Podgaec 2007
7.3.5.f. TNF-α (tumour necrosis factor alpha)	63	78	mean (range), pg/ml	0 - 125	0 (0 - 125)	NS	II-IV	follicular/ luteal/un-known	Seeber 2008
				AUC (CIs) 0.549 (0.453–0.645)					
7.3.5.g. TNF-α (tumour necrosis factor alpha)	46	48	mean ± SD, pg/ml	23.2 ± 43.6	17.0 ± 27.1	NS	I-IV	n/a	Vercellini 1993
7.3.5.h. TNF-α (tumour necrosis factor alpha)	33	22	mean ± SD, pg/ml	26.26 ± 9.31	65.40 ± 9.86	0.051	I-IV	n/a	Yagmur 2013
7.4. Immune cells									
7.4.1.Peripheral blood mononuclear cells (PBMC)									
7.4.1.1.a. activated lympho-cytes	60: 42 - rASRM I-II; 18 - rASRM III-IV)	37	mean rank values	49.31 (rASRM I-II); 47.94 (rASRM III-IV)	47.76	0.93	I-IV	follicular	Hassa 2009
7.4.1.1.b. lymphocytes	22	20	mean ± SD, x10^3 cells/ml	1.63 ± 0.44	1.7 ± 0.3	NS	I-IV	follicular	Gogacz 2014
7.4.1.1.c. lymphocytes	62	57	mean ± SD, 10^9/l	1.869 ± 0.5	2.090 ± 0.5	NS	I-II	follicular/ luteal	Matveeva 1990
7.4.1.1.d. lymphocytes	61: 33 (en-dometri-oma); 28 (non-en-	33	mean ± SD, 10^3/µl	2.12 ± 0.87 (en-dometrioma); 2.02 ± 0.68 (non-endometrioma)	2.25 ± 0.66	0.463	III-IV	n/a	Yavuzcan 2013

(Continued)

	dometri-oma)								
7.4.1.2.a. B-lymphocytes, CD3-/CD19+	19	26	mean ± SD, % of CD45+ cells	10.1 ± 7.9	7.3 ± 5.1	NS	I-IV	follicular	Iwasaki 1993
7.4.1.2.b. B-lymphocytes, CD19+	28	26	mean ± SD, %	11.6 ± 3.2 (rASRM I-II); 9.8 ± 4.1 (rASRM III-IV)	8.4 ± 3.6	NS	I-IV	follicular	Maeda 2002a
7.4.1.2.c. B-lymphocytes, CD19+	56	68	mean ± SD, % among Lym-phocytes	11.6 ± 4.8	10.5 ± 4.3	NS	I-IV	n/a	Zhang 2006a
7.4.1.3. monocytes/ macrophages, CD14+	28	26	mean ± SD, %	17.9 ± 10.2 (rASRM I-II); 17.1 ± 8.6 (rASRM III-IV)	16.6 ± 10.3	NS	I-IV	follicular	Maeda 2002a
7.4.1.4. neutrophils	61: 33 (en-dometri-oma); 28 (non-en-dometri-oma)	33	mean ± SD, 10^3/µl	4.14 ± 1.73 (en-dometrioma); 4.68 ± 2.18 (non-endometrioma)	4.50 ± 1.57	0.501	III-IV	n/a	Yavuzcan 2013
7.4.1.5. NLR (neutrophil/ lym-phocyte ratio)	61: 33 (en-dometri-oma); 28 (non-en-dometri-oma)	33	mean ± SD	2.40 ± 2.04 (en-dometrioma); 2.51 ± 1.37 (non-endometrioma)	2.11 ± 0.86	0.555	III-IV	n/a	Yavuzcan 2013
7.4.1.6.a. NK (natural killer cells)	60: 42 - rASRM I-II; 18 - rASRM III-IV)	37	mean rank values	41.36 (rASRM I-II); 50.75 (rASRM III-IV)	45.98	0.67	I-IV	follicular	Hassa 2009
7.4.1.6.b. NK (natural killer cells), CD3-/CD16+ or CD56+	19	26	mean ± SD, % of CD45+ cells	18.6 ± 8.9	23.3 ± 9.6	NS	I-IV	follicular	Iwasaki 1993
7.4.1.6.c. NK (natural killer cells), CD16	28	26	mean ± SD, %	28.8 ± 13.2 (rASRM I-II); 24.9	26.9 ± 11.2	NS	I-IV	follicular	Maeda 2002a

(Continued)

				± 11.5 (rASRM III-IV)					
7.4.1.6.d. NK (natural killer cells), CD56+	56	68	mean ± SD, % among lymphocytes	15.1 ± 9.2	13.6 ± 6.1	NS	I-IV	n/a	Zhang 2006a
7.4.1.7.a. NKR CD158b+ (KIR2DL2+NK] (killer cell inhibitory receptor subfamily 2DL2 on NK cells)	42	40	mean ± SD, %	32.2 ± 16.8	33.9 ± 14.3	NS	I-IV	n/a	Maeda 2002b
7.4.1.7.b. NKR CD158b+ (killer immunoglobulin-like receptor on natural killer cells)	56	68	mean ± SD, % among CD56+ NK cells	38.1 ± 14.5	36.0 ± 13.9	NS	I-IV	n/a	Zhang 2006a
7.4.1.8.a. NKR CD94+ (lectin-like receptor on natural killer cells)	42	40	mean ± SD, %	52.6 ± 17.9	56.2 ± 15.7	NS	I-IV	n/a	Maeda 2002b
7.4.1.8.b. NKR CD94+ (lectin-like receptor on natural killer cells)	56	68	mean ± SD, % among CD56+ NK cells	54.5 ± 14.9	50.0 ± 16.9	NS	I-IV	n/a	Zhang 2006a
7.4.1.9.a. T-lymphocytes, CD3+/CD19-	19	26	mean ± SD, % of CD45+ cells	68.3 ± 10.6	66.4 ± 10	NS	I-IV	follicular	Iwasaki 1993
7.4.1.9.b. T-lymphocytes, non MHC restricted, CD3+/CD16+ or CD56+	19	26	mean ± SD, % of CD45+ cells	4.5 ± 3.0	6.6 ± 5.6	NS	I-IV	follicular	Iwasaki 1993
7.4.1.9.c. T-lymphocytes, CD3+	28	26	mean ± SD, %	49.4 ± 15.0 (rASRM I-II); 47.3 ± 15.9 (rASRM III-IV)	52.5 ± 11.2	NS	I-IV	follicular	Maeda 2002a
7.4.1.9.d. T-lymphocytes, CD3+	62	57	mean ± SD, %	60.1 ± 8.4	60.4 ± 6.6	NS	I-II	follicular/luteal	Matveeva 1990
7.4.1.9.e. T-lymphocytes, CD2+	62	57	mean ± SD, %	67.4 ± 9.0	68.2 ± 8.3	NS	I-II	follicular/luteal	Matveeva 1990

(Continued)

7.4.1.9.f. T-lymphocytes, CD3+	56	68	mean ± SD, % among lymphocytes	61.8 ± 13.4	64.0 ± 11.0	NS	I-IV	n/a	Zhang 2006a
7.4.1.10. T-lymphocytes (inducer-T cells), CD4+/Leu 8+	19	26	mean ± SD, % of CD45+ cells	31.5 ± 11.1	28.3 ± 9.4	NS	I-IV	follicular	Iwasaki 1993
7.4.1.11.a. T-lymphocytes (T-helper cells)	60: 42 - rASRM I-II; 18 - rASRM III-IV)	37	mean rank values	46.76 (rASRM I-II); 46.33 (rASRM III-IV)	50.07	0.76	I-IV	follicular	Hassa 2009
7.4.1.11.b. T-lymphocytes (T-helper cells), CD4+/Leu 8-	19	26	mean ± SD, % of CD45+ cells	10.0 ± 4.1	10.2 ± 5.9	NS	I-IV	follicular	Iwasaki 1993
7.4.1.11.c. T-lymphocytes (T-helper cells), CD4+	28	26	mean ± SD, %	38.8 ± 9.8 (rASRM I-II);36.7 ± 13.1 (rASRM III-IV)	42.8 ± 6.5	NS	I-IV	follicular	Maeda 2002a
7.4.1.11.d. T-lymphocytes, CD4+	62	57	mean ± SD, %	40.7 ± 6.8	43.7 ± 7.4	NS	I-II	follicular/ luteal	Matveeva 1990
7.4.1.11.e. T-lymphocytes producing IL-2, CD4+/IL-2	32	30	mean ± SD (median; minimum–maximum), %	4.74 ± 2.51 (4; 1–12)	5.28 ± 2.24 (5; 1–10)	NS	I-II	periovulatory	Mier-Cabrera 2011
7.4.1.11.f. T-lymphocytes, CD4+	56	68	mean ± SD, % among lymphocytes	36.3 ± 10.5	36.3 ± 6.0	NS	I-IV	n/a	Zhang 2006a
7.4.1.12.a. T-lymphocytes (cytotoxic T-cells), CD8+	28	26	mean ± SD, %	33.2 ± 6.8 (rASRM I-II); 30.1 ± 9.6 (rASRM III-IV)	34.2 ± 7.8	NS	I-IV	follicular	Maeda 2002a
7.4.1.12.b. T-lymphocytes (T-supressor cells), CD8+	60: 42 - rASRM I-II; 18 - rASRM III-IV)	37	mean rank values	51.04 (rASRM I-II); 47.53 (rASRM III-IV)	49.46	0.62	I-IV	follicular	Hassa 2009
7.4.1.12.c. T-lymphocytes, CD8+	62	57	mean ± SD, %	21.9 ± 5.6	23.7 ± 6.8	NS	I-II	follicular/ luteal	Matveeva 1990

(Continued)

7.4.1.12.d. T-lymphocytes, CD8+	56	68	mean ± SD, % among lymphocytes	39.1 ± 7.4	40.7 ± 7.4	NS	I-IV	n/a	Zhang 2006a
7.4.1.12.e. T-lymphocytes producing IL-2, CD8+/IL-2	32	30	mean ± SD (median; minimum–maximum), %	4.03 ± 2.24 (4; 1–9)	4.68 ± 2.1 (5; 1–11)	NS	I-II	periovulatory	Mier-Cabrera 2011
7.4.1.12.f. T-lymphocytes producing interferon-gamma, CD8+/IFN-γ	32	30	mean ± SD (median; minimum–maximum), %	6.17 ± 1.85 (6; 2–10)	6.59 ± 1.69 (7; 3–9)	NS	I-II	periovulatory	Mier-Cabrera 2011
7.4.1.13.a. Tregs (regulatory T cells)	22	20	mean ± SD, %CD4+	6.5 ± 3.2	6.5 ± 3.7	NS	I-IV	follicular	Gogacz 2014
7.4.1.13.b. Treg cells (regulatory T cells), CD25+ FOXP3+	17	15	median (IQR), %	4.4 (3.11 - 5.5)	5.2 (4.1 - 5.71)	0.4&	III-IV	follicular	Olkowska-Truchanowicz 2013 [2]
7.4.1.13.c. Treg cells (regulatory T cells), $CD25^{low}$ FOXP3+	17	15	median (IQR), %	3.3 (2.1 - 4.9)	3.7 (2.4 - 4.5)	0.95	III-IV	follicular	Olkowska-Truchanowicz 2013 [2]
7.4.1.14.a. WBC (white blood cells)	22	20	mean ± SD, x10^3 cells/ml	7.53 ± 2.31	6.8 ± 1.8	NS	I-IV	follicular	Gogacz 2014
7.4.1.14.b. WBC (white blood cells)	50	36	mean ± SD, n/ul	7.299 ± 1.622	6.743 ± 1.632	0.118	I-IV	n/a	Tuten 2014a
7.4.1.14.c. WBC (white blood cells)	61: 33 (endometrioma); 28 (non-endometrioma)	33	mean ± SD, 10^3/µl	7081.5 ± 2170.6 (endometrioma); 7268.9 ± 2321.7 (non-endometrioma)	7311.2 ± 2027.2	0.902	III-IV	n/a	Yavuzcan 2013
7.4.2. Other blood cells and blood cell parameters									
7.4.2.1. Haemoglobin	61: 33 (endometrioma); 28 (non-en-	33	mean ± SD, g/dl	11.9 ± 1.6 (endometrioma);	12.0 ± 1.8	0.97	III-IV	n/a	Yavuzcan 2013

(Continued)

	dometri-oma)			12.0 ± 1.4 (non-endometrioma)					
7.4.2.2. MPV (mean platelet volume)	61: 33 (endometrioma); 28 (non-endometrioma)	33	mean ± SD, fl	8.75 ± 1.52 (endometrioma); 8.56 ± 1.27 (non-endometrioma)	8.56 ± 1.27	0.836	III-IV	n/a	Yavuzcan 2013
7.4.2.3. Platelet count	61: 33 (endometrioma); 28 (non-endometrioma)	33	mean ± SD, 10^3/µl	269848 ± 65202 (endometrioma); 298964 ± 107813 (non-endometrioma)	286484 ± 67636	0.373	III-IV	n/a	Yavuzcan 2013
7.4.2.4. PLR (Platelet/ Lymphocyte ratio)	61: 33 (endometrioma); 28 (non-endometrioma)	33	mean ± SD	162.84 ± 141.28 (endometrioma); 159.14 ± 61.20 (non-endometrioma)	132.45 ± 35.74	0.358	III-IV	n/a	Yavuzcan 2013
7.5. Interleukins									
7.5.1.a. IL-1β	22	30	median (IQR), pg/ml	10.98 (8.65, 18.52)	9.7 (3.97, 15.46)	NS	I-IV	follicular/ luteal	Bedaiwy 2002
7.5.1.b. IL-1β	15	20	median (IQR), pg/ml	5.0 (5.0 - 5.0)	5.0 (5.0 - 5.0)	0.625	I-II	luteal	Kalu 2007
7.5.1.c. IL-1β	39	19	mean ± SE, pg/ml	0.12 ± 0.09	0.17 ± 0.11	NS	I-IV	follicular	Oku 2004
7.5.1.d. IL-1β	58	27	median (IQR), pg/ml	0.00 (0.00 - 0.33)	0.08 (0.00 - 4.75)	0.054	I-IV	follicular	Szubert 2014
7.5.2.a. IL-2	33	17	mean ± SD, pg/ml	124.19 ± 336.39	247.65 ± 486.15	0.86	I-IV	follicular	Drosdzol-Cop 2012b
7.5.2.b. IL-2	60: 42 - rASRM I-II;	37	mean rank values	47.88 (rASRM I-II);	53.59	0.15	I-IV	follicular	Hassa 2009

(Continued)

	18 - rASRM III-IV			42.17 (rASRM III-IV)					
7.5.2.c. IL-2	30	20	mean, ng/L	3.6	3.6	NS	I-IV	n/a	Li 2005
7.5.2.d. IL-2	68	70		below detection limit of assay	below detection limit of assay		I-IV	follicular/ luteal	Othman 2008
7.5.2.e. IL-2	65	33	median (range), pg/ml	7.4 (0 - 34.1)	8.3 (0 - 26)	0.447	I-IV	follicular/ luteal	Podgaec 2007
7.5.3.a. IL-4	60: 42 - rASRM I-II; 18 - rASRM III-IV)	37	mean rank values	47.76 (rASRM I-II); 48.25 (rASRM III-IV)	50.77	0.5	I-IV	follicular	Hassa 2009
7.5.3.b. IL-4	65	33	median (range), pg/ml	1.9 (0 - 6.3)	2.0 (0 - 4.1)	0.731	I-IV	follicular/ luteal	Podgaec 2007
7.5.4.a. IL-6	33	17	mean ± SD, pg/ml	23.59 ± 44.17	12.63 ± 15.75	0.16	I-IV	follicular	Drosd-zol-Cop 2012a
7.5.4.b. IL-6	44	51	mean ± SEM, pg/ml	5.3 ± 0.9	12.9 ± 4.0	0.295	III-IV	n/a	Jee 2008
7.5.4.c. IL-6	15	20	median (IQR), pg/ml	5.0 (5.0 - 5.0)	5.0 (5.0 - 5.0)	0.946	I-II	luteal	Kalu 2007
7.5.4.d. IL-6	63	78	mean (range), pg/ml	0 - 160	0 - 160	NS	II-IV	follicular/ luteal/un-known	Seeber 2008
				AUC (CIs) 0.556 (0.454–0.650)					
7.5.4.e. IL-6	45	35	median (IQR), pg/ml	0.6 (0 - 1.4)	1.0 (0.4 - 1.9)	0.09	I-IV	any	Somigliana 2004
7.5.4.f. IL-6	41	26	mean, pg/ml	20	10	NS	III-IV	n/a	Suen 2014
7.5.5.a. IL-8	47	22	mean ± SD, pg/ml	10.17 ± 7.98	9.81 ± 8.11	NS	I-IV	follicular	Barcz 2002

(Continued)

7.5.5.b. IL-8	20	10	mean ± SD, ng/ml	0.1 ± 0.096	0.08 ± 0.04	0.396	III-IV	follicular/ luteal	Calienno 2008
7.5.5.c. IL-8	25	22	median (IQR), ng/ml	2.5 (1.1 - 4.1)	1.5 (1 - 1.9)	0.27	I-IV	follicular/ luteal	Gazvani 1998
7.5.5.d. IL-8	15	20	median (IQR), pg/ml	9.4 (5.4 - 13.8)	5.7 (5.0 - 8.4)	0.074	I-II	luteal	Kalu 2007
7.5.5.e. IL-8	68	70		below detection limit of assay	below detection limit of assay		I-IV	follicular/ luteal	Othman 2008
7.5.5.f. IL-8	60	20	median (IQR), ng/ml	150.1 (1650.9)	120.6(1049.8)	NS	I-IV	n/a	Ozhan 2014
7.5.6.a. IL-10	40	40	mean ± SD, pg/ml	13.05 ± 29.55	10.43 ±7.56	0.604	I-II	follicular	Andreoli 2011
7.5.6.b. IL-10	20	10	mean ± SEM, pg/ml	9.2 ± 7.0	8.6 ± 5.0	NS	n/a	luteal	Braun 1996
7.5.6.c. IL-10	60: 42 - rASRM I-II; 18 - rASRM III-IV)	37	mean rank values	48.95 (rASRM I-II); 53.56 (rASRM III-IV)	46.84	3.43	I-IV	follicular	Hassa 2009
7.5.6.d. IL-10	65	33	median (range), pg/ml	3.2 (0 - 12.9)	3.1 (0 - 7.5)	0.904	I-IV	follicular/ luteal	Podgaec 2007
7.5.7.a. IL-12	40	40	mean ± SD, pg/ml	7.95 ± 3.14	14.39 ± 11.20	0.203	I-II	follicular	Andreoli 2011
7.5.7.b. IL-12	22	32	median (IQR), pg/ml	0.00 (0.00, 0.00)	0.00 (0.00, 31.32)	NS	I-IV	follicular/ luteal	Bedaiwy 2002
7.5.7.c. IL-12	72	33	mean ± SD, pg/ml	152.14 ± 22.59	97.1 ± 19.00	NS	I-IV	follicular/ luteal	Fairbanks 2009
7.5.7.d. IL-12	61	12	mean ± SD, pg/ml	80.1 ± 49.6 (pelvic endometriosis);	69.6 ± 35.4	0.74	I-IV	follicular	Kubatova 2013

(Continued)

				76.5 ± 32.1 (endometrioma)					
7.5.7.e. IL-12	41	26		below detection limit of assay	below detection limit of assay		III-IV	n/a	Suen 2014
7.5.7.f. IL-12	53	11	median (range), pg.ml	120 (86.5 - 355) (rASRM I-II); 110 (20 - 460) (rASRM III-IV)	175 (45 - 380)	NS	I-IV	luteal	Szczepanska 2001b
7.5.8. IL-13	21	32	median (IQR), pg/ml	44.57 (44.57, 49.87)	44.57 (44.57, 44.57)	NS	I-IV	follicular/ luteal	Bedaiwy 2002
7.5.9. IL-15	68	70		below detection limit of assay	below detection limit of assay		I-IV	follicular/ luteal	Othman 2008
7.5.10.a. IL-16	22	22	median (IQR), pg/ml	539.4	778.1	NS	I-IV	n/a	Lin 2005
7.5.10.b. IL-16	22	22	median (range), pg/ml	290.5 (89.4 - 2181.2	296.8 (88.3 - 1513.6)	NS	I-IV	follicular/ luteal	Zhang 2005a
7.5.11.a. IL-17	40	40	mean ± SD, pg/ml	4.83 ± 8.60	2.35 ± 2.40	0.325	I-II	follicular	Andreoli 2011
7.5.11.b. IL-17	69	32		below detection limit of assay	below detection limit of assay		I-IV	n/a	Paiva 2014
7.5.12.a. IL-18	72	33	mean ± SD, pg/ml	70.52 ± 11.53	62.07 ± 8.08	NS	I-IV	follicular/ luteal	Fairbanks 2009
7.5.12.b. IL-18	56	22	mean ± SD, pg/ml	391.07 ± 119.71	373.42 ± 129.11	NS	I-II	follicular	Glitz 2009
7.5.12.c. IL-18	39	19	mean ± SE, pg/ml	177.17 ± 28.37	174.14 ± 27.48	0.945	I-IV	follicular	Oku 2004 [3]
7.5.12.d. IL-18	39	21	mean ± SD, pg/ml	8 1. 86 ± 18. 22	78. 99 ± 28. 58	NS	I-IV	n/a	Zhang 2005b

(Continued)

7.5.13. IL-23	40	40	mean ± SD, pg/ml	6.49 ± 4.71	10.12 ± 9.87	0.209	I-II	follicular	Andreoli 2011
7.6. Other immune/ inflammatory markers									
7.6.1. C3a (anaphylatoxin)	109	51	median (range) ng/ml	102 (27 - 2213)	105 (32 - 2340)	0.84	I-IV	any	Fassbender 2009
7.6.2. sCD23 (soluble CD23, low-affinity IgE receptor)	44	58	mean ± SD, U/ml	menstrual cycle phase: 42.85 ± 3.93 late follicular cycle phase: 52.98 ± 10.58	menstrual cycle phase: 54.47 ± 7.21late follicular cycle phase: 58.08 ± 8.09	0.132 0.697	I-IV	menstrual and follicular	Ramos 2012
7.6.3. sCD163 (soluble haemoglobin scavenger receptor)	44	51	mean ± SEM, ng/ml	3431.7 ± 343.9	3,231.0 ± 391.7	0.212	III-IV	n/a	Jee 2008
7.6.4.a. CRP (C-reactive protein)	50	50	mean (range)	3.57 (0.3 - 27.66)	1.79 (0.21 - 10.85)	0.101	I-IV	early follicular	Dayangan Sayan 2013
7.6.4.b. CRP (C-reactive protein)	90	89	mean ± SEM, µg/ml	7.6 ± 1.7	6.9 ± 2.1	NS	n/a	follicular/ luteal	Kianpour 2012
7.6.4.c. CRP (C-reactive protein)	70	32	median (IQR), mg/l	1.90 (1.50 - 2.70)	2.00 (1.60 - 2.80)	0.556	I-IV	follicular	Szubert 2014
7.6.4.d. hs-CRP (high sensitive C-reactive protein)	370	464	median (range), ng.ml	0.82 (0.04 - 42.89]	0.9 (0.03 - 43.73]	0.599	I-IV	follicular/ luteal/unknown	Thubert 2014
7.6.4.e. CRP (C-reactive protein)	50	36	mean ± SD, mg/ml	3.7 ± 4.4	2.3 ± 2.2	0.062	I-IV	n/a	Tuten 2014a
7.6.4.f. CRP (C-reactive protein)	18	14	median (95% CI), mg/l	1 (1 - 2)	2 (1 - 3)	0.18	I-IV	follicular/ luteal	Riley 2007
7.6.5. sHLA-I (soluble human leukocyte class I antigens)	15	15	mean ± SD, OD	0.55 ± 0.3	0.35 ± 0.1	0.06	I-IV	follicular/ luteal	De Placido 1998

(Continued)

7.6.6. Immunoglobulins IgG	62	57	mean ± SD, mg%	1260.3 ± 378.7	1170.9 ± 342.4	NS	I-II	follicular/ luteal	Matveeva 1990
7.6.7. ImmunoglobulinsIgA	62	57	mean ± SD, mg%	196.6 ± 71.2	181.0 ± 78.3	NS	I-II	follicular/ luteal	Matveeva 1990
7.6.8. MPO (myeloperoxidase)	10	7	median (IQR), OD	0,168 (0,139 - 0,491)	0,211 (0,164 - 0,351)	0.757	II-IV	follicular	Da Silva 2014
7.6.9. NAG (N-acetyl-b-Dglu-cosaminidase)	10	7	median (IQR), OD	85,53 (43,52 - 286,56)	57,66 (38,17 - 101,15)	0.4079	II-IV	follicular	Da Silva 2014
7.6.10. PGE2 (prostaglandin E2)	58	28	median (IQR), ng/ml	3.75 (3 - 6.5)	4 (2 - 7)	NS	I-IV	any	Khan 2012
7.6.11. PLA2G2A (phospholi-pase A2 group IIA)	53	38	mean ± SD, ng/ml	2.9 ± 2.1	3.1 ± 2.2	0.7989	I-IV	follicular/ luteal	Kocbek 2014a
7.6.12.a. RANTES (regulated on activation, normal T cell ex-pressed and secreted)	17	23	median (IQR), pg/ml	789.4 (550.8 - 1009.5)	662.5 (422.8 - 960.4)	0.35	I-II	luteal	Kalu 2007
7.6.12.b. RANTES (regulated on activation, normal T cell ex-pressed and secreted)	23	9	range, pg/ml	5,200 - 57,800	3,875 - 35,100	NS	I-IV	n/a	Markham 1997a
7.6.13. Phospholipid fatty acids	64	74	mean ± SD, %				I-IV	follicular/ luteal	Khanaki 2012
• 14:0 (myristic acid)				0.29 ± 0.21	0.26 ± 0.11	0.24			
• 16:0 (palmitic acid)				49.14 ± 5.28	48.53 ± 7.43	0.57			
• 16:1(palmitoleic acid)				0.34 ± 0.20	0.39 ± 0.18	0.15			
• 18:1 n-9 (oleic acid)				6.13 ± 1.31	6.28 ± 1.72	0.57			
• 18:2 n-6 (linoleic acid)				20.31 ± 3.19	19.66 ± 3.59	0.26			
• 18:3 n-3 (α-linolenic acid)				0.37 ± 0.19	0.35 ± 0.14	0.44			
• 20:4 n-6 (AA - arachidonic acid)				6.93 ± 1.92	7.23 ± 2.31	0.41			

(Continued)

• 20:5 n-3 (eicosapentaenoic acid)				0.35 ± 0.35	0.36 ± 0.24	0.801			
• 22:6 n-3 (docosahexaenoic acid)				0.79 ± 0.82	0.70 ± 0.75	0.5			
• SFA (saturated fatty acids)				61.89 ± 4.57	62.14 ± 6.32	0.78			
• MUFA (mono unsaturated fatty acids)				6.47 ± 1.33	6.66 ± 1.73	0.46			
• Omega-3 fatty acids				1.51 ± 1.03	1.41 ± 0.79	0.52			
• Omega-6 fatty acids				27.25 ± 3.41	26.89 ± 4.48	0.6			
• SFA/UFA (saturated fatty acids to unsaturated fatty acids)				1.80 ± 0.33	1.84 ± 0.39	0.45			
• Omega-3/Omega-6				0.05 ± 0.03	0.05 ± 0.02	0.57			
• EPA/AA (eicosapentaenoic acid to arachidonic acid)				0.05 ± 0.05	0.05 ± 0.04	0.74			
8. Nerve growth markers									
8.1. CNTF (ciliary neurotrophic factor)	69	32	median (IQR), pg/ml	897.7 (29.8 - 2709.1)	450.5 930.5 - 2999.7)	0.14	I-IV	n/a	Paiva 2014
8.2. GDNF (glial-derived neurotrophic factor)	69	32	median (IQR), pg/ml	58.2 (0 - 168.1)	41.4 (0 - 268.3)	0.43	I-IV	n/a	Paiva 2014
8.1. CNTF (ciliary neurotrophic factor)	69	32	median (IQR), pg/ml	897.7 (29.8 - 2709.1)	450.5 930.5 - 2999.7)	0.14	I-IV	n/a	Paiva 2014
									Paiva 2014
9.1.a. DBP (vitamin D binding protein)	26	17	mean ± SE, µg/ml	449.4 ± 24.4	424.5 ± 23.5	0.4911	I-IV	follicular	Borkowski 2008

(Continued)

9.1.b. DBP (vitamin D binding protein)	88	40	mean ± SEM, %Vo	0.568 ± 0.034	0.563 ± 0.047	NS	I-IV	follicular/ luteal	Ferrero 2005a
9.2. enolase	60	20	median (IQR), ng/ml	6.9 (49.8)	4.0 (19.0)	NS	I-IV	n/a	Ozhan 2014
9.3. PDPK1 (phosphoinositide dependent protein kinase 1)	60	20	median (IQR), OD	0.265 (0.053)	0.287 (0.075)	NS	I-IV	n/a	Ozhan 2014
10. Oxidative stress markers									
10.1. ascorbic acid	32	30	mean ± SD (median; minimum–maximum), µmol/l	57.17 ± 12.43 (122–24)	53.42 ± 13.29 (106–28)	NS	I-II	peri ovulatory	Mier-Cabrera 2011
10.2. GSH (glutathione)	69	32	median (IQR), µg/ml	10517 (1494 - 26945)	8741 (2267 - 46420)	0.36	I-IV	n/a	Paiva 2014
10.3. HSP70 (heat shock protein 70)	30	20	median (IQR), ng/ml	2.2 (1.5 - 3)	2 (1 - 3)	NS	I-IV	any	Khan 2013
10.4. HSP70 (heat shock protein 70)	45	21	mean ± SD, ng/ml	1.240 ± 1.279	0.875 ± 1.336	0.634	I-IV	n/a	Lambrinoudaki 2009
10.5. IMA (Ischemia-modified albumin)	45	21	mean ± SD, U/ml	85.9 ± 11.9	86.4 ± 16.4	0.887	I-IV	n/a	Lambrinoudaki 2009
10.6. malondialdehyde	32	30	mean ± SD (median; minimum–maximum), µmol/l	27.17 ± 8.67 (42–14)	23.75 ± 6.46 (33–8)	NS	I-II	peri ovulatory	Mier-Cabrera 2011
10.7. nitrotyrosine	69	32		below detection limit of assay	below detection limit of assay		I-IV	n/a	Paiva 2014
10.8. SOD3 (superoxide dismutase)	69	32	median (IQR), (x10^5), pg/ml	2.14 (0.96 - 4.8)	2.03 (0.7 - 7.1)	0.95	I-IV	n/a	Paiva 2014

(Continued)

10.9. TRX (Thioredoxin)	45	21	mean ± SD, ng/ml	55.7 ± 45.1	54.6 ± 45.7	0.932	I-IV	n/a	Lambrinoudaki 2009
10.10. vitamin E	69	32	median (IQR), µmol/ml	1.07 (0.04 - 3.1)	0.98 (0.4 - 2.40	0.82	I-IV	n/a	Paiva 2014
11. Tumour markers									
11.1. AFP (alpha-fetoprotein)	36	36	mean ± SD, ng/ml	crude values		NS	I-IV	luteal	Philippoussis 2004
				1.9 ± 0.9	1.8 ± 1.5				
				adjusted values (indication for surgery, BMI, and presence of uterine leiomyoma using a univariate general linear model]		NS			
				1.9 ± 1.3	1.8 ± 1.3				
11.2. CA-19.9 (cancer antigen-19.9)	45	35	median (IQR), IU/ml	9.8 (4.5-20.8)	7.4 (2.8-11.5)	0.11	I-IV	any	Somigliana 2004
11.3.a. CA-125 (cancer antigen-125)	13	67	mean, U/ml	26.9	28.3	0.6389	I-II	follicular	Barbosa 2009
11.3.b. CA-125 (cancer antigen-125)	18	14	median (95% CI), kU/l	25 (15 - 46)	15 (11 - 19)	0.06	I-IV	follicular/ luteal	Riley 2007
11.4. c-erbB-2 (HER-2/neu] (erythroblastosis oncogene B or human epidermal growth factor receptor-2 derived from glioblastoma)	36	36	mean ± SD, ng/ml	crude values		NS	I-IV	luteal	Philippoussis 2004
				2.8 ± 1.7	2.6 ± 1.3				
				adjusted values (indication for surgery, BMI, and presence of uterine leiomyoma using a univariate general linear model]		NS			
				2.7 ± 1.6	2.9 ± 1.6				
11.5. HE4 (human epididymal secretory protein E4)	123	52	median, pM	43.5	41.2	NS	I-IV	any	Hallamaa 2012

(Continued)

Notes:

[1] The biomarker was assessed within a diagnostic model of combined biomarkers in this study.

[2] The authors also report the negative findings for CD4+ and CD4+ CD25+ Treg cells, but these are not presented in the review as data were not shown Olkowska-Truchanowicz 2013.

[3] The authors also report the negative findings for IL-2, IL-4, IL-6, IL-8, IL-10, TNF-a, GM-CSF and IFN-γ, but these are not presented in the review as data were not shown Oku 2004.

For a comprehensive list of all biomarkers with their biological annotation, please see Appendix 1.

Footnotes

Appendix 8. Blood biomarkers of limited diagnostic value in endometriosis

Biological group	Blood biomarker [1]
1. Angiogenesis and growth markers	Glycodelin
	IGFBP-3
	Leptin
2. Cell adhesion molecules	sICAM-1
3. Immune system and inflammatory markers	Anti-endometrial antibodies
	hs-CRP
	sGM-CSF
	IL-1β
	IL-2
	IL-4
	IL-6 (cut-off > 1.9-2.0 pg/ml)
	IL-8
	IL-10
	IL-12
	IL-18
	IFN-γ
	Immune cells and cell parameters (lymphocytes and lymphocyte subsets, white blood cell, platelets, haemoglobin)
	MCP-1
	MIF
	TNF-α
4. Tumour markers	CA-19.9 (cut-off > 37 U/ml)
	CA-125 (cut-off > 10-14.7 U/ml; > 16-17.6 U/ml; > 20 U/ml; > 25-26 U/ml; > 30-33 U/ml; > 35-36 U/ml)

[1] Limited diagnostic value was defined when at least 3 studies demonstrated low diagnostic estimates that do not meet or approach the criteria for either replacement or triage test and/or negative findings; we advise against further evaluation of these biomarkers in the diagnosis of endometriosis.

For a comprehensive list of all biomarkers with their biological annotation, please see Appendix 1.

Appendix 9. Blood biomarkers that possibly have limited diagnostic value in endometriosis

Biological group	Blood biomarker [1]
1. Angiogenesis and growth factors	Angiogenic activity of serum
	CAC
	EGF
	sEGF-R
	sFlt-1 (sVEGFR-1]
	HGF
	IGF-1
	IGF-2
	PDGF
2. Apoptosis markers	Annexin-V
	Anti-Survivin Abs
	Apoptotic cells
	sFas
	Survivin
3. Cell adhesion molecules	Biglycan
	sE-selectin
	LN-1
	MMP-9
4. Cytoskeleton molecules	CK 19
5. DNA-repair/telomere maintenance molecules	Telomere length
6. Hormonal markers	Prolactin
7. Immune system and inflammatory markers	Anti-laminin-1 auto Abs
	Anti-sperm and anti-zona pellucida auto Abs
	C3a
	sCD23

(Continued)

	CCR1
	Copeptin
	Epo
	sHLA-I
	IL-6 (except for the cut-off values reported in Table 4; Appendix 8)
	IL-13
	IL-15
	IL-16
	IL-17
	IL-23
	Immunoglobulins IgA and IgG
	Immune cells and cell parameters (monocytes, macrophages, neutrophils, NLR, NKR CD158b+, NKR CD94+, Treg cells)
	MPO
	NAG
	PGE2
	Phospholipid fatty acids
	PLA2G2A
	RANTES
8. Nerve growth markers	CNTF
	GDNF
	NGF
	NT4
9. Other peptides/proteins	DBP
	Enolase
	PDPK1
	STX-5
10. Oxidative stress markers	Ascorbic acid
	GSH

(Continued)

	HSP70
	IMA
	Malondialdehyde
	Nitrotyrosine
	SOD3
	Thiols
	TRX
	Vitamin E
11. Post-transcriptional regulators of gene expression (microRNAs)	miR-17-5
	miR-122
	miR-199a
12. Tumour markers	AFP
	CA-15.3
	CA-19.9 (cut-off > 37 U/ml)
	CA-72 (TAG-72)
	c-erbB-2 (HER-2/neu)
	HE4
13. Combined markers	All the reported combinations, excluding the tests presented in Table 4 as 'promising tests'

Notes:

[1] Tests that appear to have limited diagnostic value, but there is insufficient data to confidently comment on their diagnostic role (less than 3 studies with low diagnostic estimates and/or negative findings); we advise considering further investigation with a focus of specific phases of menstrual cycle, specific types of endometriosis, by implementing different cut-off values or by utilising different laboratory methods.

For a comprehensive list of all biomarkers with their biological annotation, please see Appendix 1.

CONTRIBUTIONS OF AUTHORS

Vicki Nisenblat and Louise Hull co-ordinated the production of the protocol and the review series; were involved in literature search, quality appraisal and data extraction for the included studies; and produced the first draft of the review. Patrick Bossuyt provided advice on the statistical methods for the review and performed the analyses. Rabia Shaikh participated in literature search, study selection, quality appraisal and data extraction for the included studies. Cindy Farquhar critically reviewed the methodological aspects and participated in the study design.Vanessa Jordan and Carola S Scheffers were involved in quality appraisal and data extraction for the included studies. Neil Johnson and Ben Willem Mol contributed to the design of the review and critically reviewed the review content. All the authors contributed to the revision and drafting of the review.

DECLARATIONS OF INTEREST

Vicki Nisenblat: none known.
Patrick MM Bossuyt: none known.
Rabia Shaikh: none known.
Cindy Farquhar: Cindy Farquhar is a director/shareholder of a small day stay surgical unit and fertility/gynaecology clinic and undertakes private practice within those facilities.
Vanessa Jordan: none known.
Carola S Scheffers: none known.
Ben Willem J Mol: none known.
Neil Johnson: Professor Neil Johnson is involved in research funded by Abb-Vie. He has received support to attend conferences from MSD, Merck-Serono and Bayer. He has been on an advisory board for Vifor Pharma.
M Louise Hull: Dr M. L. Hull obtained a grant of $10,000 to carry out a prevalence study of ultrasonographically diagnosed endometriosis in a fertility population.

SOURCES OF SUPPORT

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

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

General scope: this review is a part of the review series arising from the same generic protocol. The following sections were adjusted to the main topic of the review as described below.

'Background': the section on the index test was modified, and we removed all the information irrelevant to blood testing. We updated the 'Rationale' section to include a clearer definition of triage diagnostic tests.

'Objectives':

- Substantial numbers of studies revealed biomarkers with expression levels that were not altered by the presence of endometriosis (there was no statistically significant difference between women with and without the disease). We included these data from the adequately designed studies, justifying our decision in the Background section under 'Rationale', in the Methods section under 'Criteria for considering studies for this review', 'Index tests' and added to 'Objectives' as a secondary objective: 'To assess the biomarkers that were not affected by endometriosis and hence were unlikely to discriminate between women with and without the disease'.
- We updated the list of the sources of heterogeneity.

Methods:

- We updated 'Criteria for considering studies for this review' as follows.
- ◦ 'Types of studies': We removed the cohort and case-control classifications and introduced the concepts of single-gate design and two-gate design. We defined this as the presence of a single or multiple set of inclusion criteria by clinical condition or by reference standard. We found this classification more informative in the description of diagnostic studies, all of which are cross-sectional in nature. We limited the inclusion criteria to the studies with a single set of inclusion criteria by reference standard (i.e. all women who underwent abdominal surgery), but included single or multiple sets of inclusion criteria by clinical presentation (i.e. women with suspected endometriosis or other indications for abdominal surgery), referring to these as single-gate design and two-gate design, respectively.
- ◦ Likewise, we removed the terminology 'prospective studies' and introduced 'studies performed on prospectively collected samples'. This decision was guided by the fact that most diagnostic studies are retrospective in nature, as they aim to compare the result of index test with the result of reference standard in the same group of participants, where the groups are classified by the outcome of reference standard. Also, the analysis of the index test could have been performed retrospectively in a single batch on stored samples after the prospective collection of samples. The timing of sample collection (before or after surgical treatment of the disease) from a preoperatively recruited population has more impact on the test result than the timing of the laboratory assay. Therefore, we included only studies that collected blood before the reference surgical procedure, (i.e. prospectively collected), irrespective

of the actual timing of test performance. We refrained from labelling studies as prospective or retrospective to avoid confusion. This allowed us to include the studies from well-established high quality tissue banks using well-characterised archived samples, as omitting these studies would have resulted in the loss of potentially valuable data.
    - We modified the index tests to pertain only to blood biomarkers and updated the table listing the tests of interest (Table 2) accordingly.
    - Target conditions also included deep infiltrating pelvic endometriosis in view of the growing body of literature on this condition as a separate entity and its diagnostic importance to optimise the surgical approach.
    - Spectrum of disease: following an ad hoc observation, we included the studies that involved only a selected population of women with endometriosis (i.e. specific rASRM stages) in view of the emerging evidence on the poor correlation of this classification with infertility and pain symptoms. Exclusion of such studies could result in the loss of potentially important diagnostic information from otherwise eligible publications. Where possible we aimed to address the impact of the inclusion of these studies in investigations of heterogeneity.
- Search methods for identification of studies:
    - In the protocol, we stated that we would identify the grey literature (unpublished studies including conference proceedings and reports) and define specific search strategies. In practice, the paucity of relevant data that was available from abstracts made it impossible to apply the selection criteria and methodological quality judgement to these studies. We anticipated that identification of this type of study and attempts to obtain the necessary information directly from the study investigators would increase the already labour-intensive work involved in preparation of this review. Therefore, by consensus among the key authors, we removed already identified unpublished studies and did not complete an intended search for unpublished material.
    - We updated the search strings for all biomarkers excluding imaging (searched separately), applying the same principles as presented in the protocol.
- Assessment of methodological quality: We tailored the QUADAS-2 tool for the topic of the review. The differences between the original QUADAS-2 tool and the one designed for this review are outlined in the relevant section in the Methods.

Analysis:

- The section on statistical methods was amended and tailored to the types of tests included in the review.
- We performed no sensitivity analyses and no assessment of heterogeneity due to insufficient data for most tests, except for CA-125 at a single threshold.
- When a test performance was judged against the predetermined diagnostic criteria, we only considered the point estimates of sensitivity and specificity, as we believe that presenting these metrics of test performance is the most helpful and informative way to summarise the diagnostic data. We acknowledge that the choice of the most helpful summary is subjective. There are tests where the point estimates did not reach the predetermined criteria, but the confidence intervals (CIs) contain the values above the thresholds for replacement tests, triage tests or both. These tests could have diagnostic value if the point values underestimated their diagnostic potential. For the tests where the point estimates reached the criteria for a replacement or triage tests but the CIs contained values below the thresholds, point values could have overestimated the diagnostic performance of the test. If the range of the CIs rather than the point estimates of the data are used, the predetermined cut-off becomes meaningless. We did not consider CIs in qualifying the test performance; however, we used the CIs in interpreting the reliability of the obtained data.

The authors list and order changed to accurately reflect their contributions to the review.

NOTES

We split the initially planned single review on the non-invasive tests for diagnosis of endometriosis into several smaller reviews in order to facilitate data handling and interpretation, due to abundance and diversity of the suggested tests. The review was generated from a generic protocol, which we designed for all the reviews in this series. The other reviews from the series include: 'Endometrial biomarkers for the non-invasive diagnosis of endometriosis', 'Urinary biomarkers for the non-invasive diagnosis of endometriosis', 'Imaging modalities for the non-invasive diagnosis of endometriosis', and 'Combined biomarkers for the non-invasive diagnosis of endometriosis', which is also a summary review of the series.

INDEX TERMS

Medical Subject Headings (MeSH)

Autoantibodies [blood]; Biomarkers [*blood]; CA-125 Antigen [blood]; CA-19-9 Antigen [blood]; Endometriosis [*diagnosis]; Endometrium [immunology]; Interleukin-6 [blood]; Ovarian Diseases [*diagnosis]; Pelvis; Peritoneal Diseases [*diagnosis]; Randomized Controlled Trials as Topic

MeSH check words

Adult; Female; Humans