CA 125 AS A DIAGNOSTIC AND PROGNOSTIC MARKER FOR ENDOMETRIOSIS: SYSTEMATIC REVIEW AND META-ANALYSIS

SavvasArgyridis

BACKGROUND: Endometriosis affects approximately 6-10% of reproductive aged women, with significantly higher rates among infertile or chronic pelvic pain sufferers. It is a disease that often remains undiagnosed because it requires histological confirmation. According to its manifestations (adhesions, endometriomas) is classified as minimal, mild moderate and severe (stages I-IV). Various modalities have been proposed over the years in order to diagnose endometriosis, but histological confirmation remains the gold standard. Several serum biomarkers have also been evaluated as possible diagnostic and prognostic examinations in order to avoid surgical confirmation. CA 125 is the most studied and promising one as several studies show, despite the fact that several benign gynecological conditions may affect its serum levels.

PURPOSE: The identification of serum CA 125 sensitivity, specificity and positive predictive value for endometriosis diagnosis and stage correlation, as well as the correlation of other biomarkers such as CA 19-9 to CA 125 levels.

METHODS:A systematicliterature search was conducted of electronic databases in order to identify published or accepted for publication articles relevant to the subject.

RESULTS: The systematic review yielded 423 citations and after screening the title and abstract, 56 were assessed on full text. Out of these articles, after excluding 43 due to various issues, 13 were included in the meta-analysis. The mean sensitivity rate is at 62.83% for all stages of disease, somewhat higher from older relevant meta-analyses, while mean specificity rate is at 85.26% which is consistent with what other researchers reported in the past. Positive predictive value is at 90.14%, correlating well with specificity values, as expected. CA 125 correlates with disease stage, as mean control group value is at 14.05 U/ml, and mean case group value is at 43.19 U/ml. Mean values for stage I-II are at 25.05 U/ml, while stage III-IV are at 63.98 U/ml, showing a significant statistical difference. CA 125 also shows a positive correlation to the presence of endometriomas as levels are significantly higher than in the absence of it. Correlation of other biomarkers to CA 125 for endometriosis diagnosis show conflicting results, as some suggest that a positive one exists while others suggest not.

CONCLUSIONS: Sensitivity rates vary significantly among different groups and with a mean rate of 62.83% for a cut-off limit of 35 U/ml, is not considered high enough to allow the biomarker to be used as a screening tool for the disease. Specificity rates are more consistent and are high enough (85.26%) to allow its' use as a diagnostic tool. Positive predictive value also correlates well with specificity and may also be used for this. Positive correlation to disease stage also enables use as a diagnostic modality. Overall, further studies are needed in order to assess the sensitivity rates of the assay in order for serum CA 125 to be considered as a single diagnostic or screening modality for endometriosis.

TABLE OF CONTENTS

INTRODUCTION

Endometriosis is defined as the presence of endometrial glands and/or stroma in positions outside the endometrial cavity which induces a chronic inflammatory process (Giudice, 2010). It affects approximately 6-10% of reproductive aged women (Guidice and Kao, 2004),but with a much higher prevalence of about 35-50%among infertile women (Missmer and Cramer, 2003) and about 60-70% among women with chronic pelvic pain (Amaral*et al*. 2006). The true prevalence of the disease is believed to be unknown, partly due to the fact that approximately 20% of women that have endometriosis remain asymptomatic (Missmer and Cramer, 2003) and partly because disease diagnosis requires laparoscopic visual and histological confirmation (Kennedy *et al*. 2005) . Another contributing factor to this problem is the fact that minimal or mild endometriosis are usually difficult to diagnoseuntil they have progressed to moderate or severe disease due to immunological alterations in some women (Abrao*et al*., 1997). Several theories have been proposed over the years in order to explain the pathophysiological mechanisms that are implicated in endometriosis development. Retrograde menstruation as it was first proposed by Sampson in 1927- the "Sampson hypothesis", (Sampson 1927), is a common physiological even that causes intraperitoneal spill of endometrial cells which are then implanted on the peritoneum, thus causing pelvic endometriosis. Observations over the years have confirmed that the majority of women (90%) experience retrograde menstruation(Koninckx*et al*., 1980c; van der Linden *et al*., 1995) especially if they have patent fallopian tubes and/or congenital uterine or outflow abnormalities (Liu and Hitchcock, 1986).This was confirmed by numerous animal studies in which laparoscopy was conducted during the perimenstrual period (Scott *et al*., 1953).Some studies suggest that the prevalence rates of thissituation is the same amongendometriotic and healthy

individuals (Halmee*t al*., 1984), while others suggest an increase of this among endometriotic women (Liu and Hitchcock, 1986; D'Hooghe*et al*., 1986). The reason why some women develop endometriosis and others don't may be attributed to the amount of blood that is spilled into the peritoneal cavity. Short cycle length and longer menstrual flow are associated with increased prevalence of the disease (Cramer *et al*., 1986).The immunological hypothesis has been proposed in combination with the retrograde menstruation hypothesis in order to explain why some women develop endometriosis and others don't. Animal studies suggest that there is increase of inflammatory parameters in the peritoneal fluid of endometriotic women such as white blood cell count, tumor necrosis factor-a (TNF-a), tumor growth factor β-1 (TGF-β-1) and interleukins (Debrock*et al*., 2000; D'Hooghe*etal.*, 2001).A decreased endometrial cell clearance from the peritoneal fluid is attributed to a decreased natural killer (NK) cell or macrophage activity (Oosterlynch*et al*., 1991) while others support that there is insufficient evidence to support this theory, since there is no in vitro evidence that endometrial cells are being cleared by macrophages or NK cells (D'Hooghe*et al*., 1997). Evidence to support it originates from observational studies that immunosuppressed women do not experience higher prevalence rates of the disease than normal women (Armenti*et al.,* 1994). The angiogenesis theory supports that endometrial cells of women with certain gene mutations such as the cell adhesion molecules genes, have increased proliferation properties, that enable endometriosis development (Healy *et al*., 1998; Somogliana*et al*., 1996).An increase of vascular endothelial growth factor (VEGF) secretion by activated macrophages support this theory (McLaren *et al*., 1996b).Studies conducted among first-degree relatives of women with endometriosis indicate that there is increased prevalence of the disease (dos Reis *et al*., 1999) suggesting genetic predisposition as a contributing factor. An

increase inaromatase expression which in turn increases androgen to estrogen production, a predisposing factor of endometriosis, is considered as a possible genetic mechanism (Matsuzaki*et al*., 2004; Acien*et al*., 2007).The metaplasia theory suggests that there is transformation of coelomic epithelial cells into endometrial cells, responsive to a stimuli, thus fur unknown. It explains extrapelvicendometriotic findings as well as post-hysterectomy lesions (Suginami, 1991; Fujii 1991).

The histological features of the disease are variable, with the most characteristic macroscopic features being the "powder burn" or "gunshot" black-brown lesions on the peritoneal surface and ovaries. Atypical features include red petechial or polypoid hemorrhagic lesions or white-yellow plaques. Deep infiltrating endometriosis is the presence of nodules on the uterosacral ligaments or bowel which extent more than 5mm beneath the peritoneal surface. Evidence suggest that superficial disease is a natural phenomenon with little impact on a woman's health, since complications are rare (infertility or chronic pain), while deep disease displays a more "aggressive" behavior, considered responsible for more severe complications (Koninckx and Martin, 1992). Adhesions that may be filmy or dense (peritubal, periovarian, pouch of Douglas) are another feature of the disease as well as endometriomas which are thick walled, hemorrhagic fluid containing ovarian cysts (Kennedy *et al*., 2005). The disease is primarily located in pelvic sites such as the uterine serosa, pouch of Douglas, ovaries, tubes, urinary bladder and bowel, while extrapelvic sites include postoperative scars (Meigs 1960), umbilicus (Chatzikokkinou*et al.,* 2008), lungs and the nervous system.

Disease staging is a rather controversial issue because it requires surgical macroscopic observations and it does not necessarily correlate to symptom severity, except in cases of deep infiltrating endometriosis (Vercellini*et al.,*

2007). According to the American Society of Reproductive Medicine revised classification, there are four stages based on intra-operative findings (ASRM 1996).Minimal endometriosis (score 1-5) describes cases with a few, superficial implants, while mild (score 6-15) those with deeper implants. Moderate disease (score 16-40) also includes small ovarian endometriomas and filmy adhesions while severe (score >40) includes large, bilateral endometriomas and dense adhesions.

table 1. ASRM revised classification of endometriosis (ASRM, 1996).

location	**implants**	**<1cm**	**1-3cm**	**>3cm**
peritoneum	superficial	1	2	4
	deep	2	4	6
ovary	rt superficial	1	2	4
	rt deep	4	16	20
	lt superficial	1	2	4
	lt deep	4	16	20
location	**obliteration**	**partial**	**complete**	
p. cul de sac		4	40	
location	**adhesions**	**<1/3**	**1/3-2/3**	**>2/3**
ovary	rt filmy	1	2	4
	rt dense	4	8	16
	lt filmy	1	2	4
	lt dense	4	8	16
tube	rt filmy	1	2	4
	rt dense	4	8	16
	lt filmy	1	2	4
	lt dense	4	8	16

The commonest clinical symptoms are dysmenorrhea and dyspareunia (ACOG 1993). Dysmenorrhoea usually develops after years of pain-less menstruation and is attributed to micro-bleeding and inflammatory process within endometriotic implants (Brosens, 1997).There is no clear association between site, size, disease severity and pain severity (GISE, 2001). The only exception is found in cases of deep infiltrating endometriosis where pain severity is strongly associated to disease severity (Vercellini, 1997).The presence of dense adhesions may contribute to pain severity due to the fact that they contain

inflammatory and endometrial cells as well (Jirasek*et al*., 1998).Rare clinical symptoms according to unusual implant location include cyclical rectal bleeding and painful defacation, cyclical hematuria and hemoptysis.

A major long-term complication is secondary infertility due to a hostile inflammatory environment to ovum and sperm, tubal blockage due to adhesions and interference of ovulation or ovum pickup by the tubes due to the presence of endoemtriomas (Wellbery, 1999).Another possible major complication is the increased risk of ovarian cancer (endometrioid, clear cell), especially in women that develop ovarian cancer in their reproductive age (Melin*et al.,* 2006; Sayasneh*et al*., 2011).

Various diagnostic modalities have been proposed over the years for endometriosis diagnosis. Transvaginal ultrasound remains probably the most sensitive (95-98%) and specific (98-100%) imaging technique, regarding deep infiltrating endometriosis and ovarian endometriomas (Abrao*et al*., 2007).It has no role in diagnosing peritoneal implants, therefore cannot exclude minimal or mild disease (Moore *et al.,* 2002).Magnetic resonance imaging (MRI) has little value as a diagnostic modality (Ang*et al*., 2015). Laparoscopy remains the gold standard technique for diagnosis as visual inspection and histologic confirmation of the lesions (Chapron*et al.,* 2003b).Evidence though suggest an overestimation of diagnosis by using visual criteria alone and histological confirmation (Janssen *et al.,* 2013).The concurrent diagnosis and treatment of endometriotic lesions (ablation, excision) was up to recently standard of care since it was related with pain reduction and quality of life improvement (Abbott *et al.,* 2003). Recent data though suggest that endometrioma excision may cause a reduction of ovarian reserve and therefore a reduced response to gonadotropin stimulation (Ruiz-Flores and Garcia-Valesco, 2012). Others suggest that unless a serious distortion of anatomical relations is responsible for infertility, there is no

improvement in IVF results after endometrioma excision (Tsoumpou et al., 2009).

Several serum biomarkers have been proposed over the years in order to be used as diagnostic tests for endometriosis, in the effort for a easy, cheap, quick and non invasive test .A biomarker is a measureable indicator of a normal, pathological process or pharmacological response to a therapeutic intervention (Biomarkers Definitions Working Group 2001). The most widely used and assessed biomarker is CA 125.CA 125 is a high molecular weight glycoprotein encoded by the MUC16 gene (O'Brien *et al*., 2001),which is produced by coelomic cells during embryonic life and later found in epithelial cells (Koninckx PR, 1994).Several etiologies are known to effect CA 125 levels, which causes diagnostic issues in regards to endometriosis (Mol et al., 1998). Benign etiologies such as menstruation may be responsible for a three times higher level (Pittaway and Fayez, 1987).Fluctuations during menstrual cycle have also been reported in other studies, but suggest that besides the peri-menstrual time, the changes are minor in other parts of the cycle (Kafali*et al*., 2004).Other physiologic and pathologic processes include cirrhosis, congestive heart failure, pelvic inflammatory disease, adenomyosis, leiomyoma, hormone therapy, hysterectomy as well as various malignancies such as ovary, breast, colon, lung, liver cancer (Pauler*et al*., 2001; Hermsen*et al.,* 2007).Age and race variations may be considered as well (Koper *et al.,* 1997).The first association study between endometriosis and CA 125, was carried out in 1984 by Niloff et al., and since then numerous others followed (Niloff*et al.,*1984).The pathophysiological mechanisms responsible for the apparent increase of this marker in endometriotic women are believed to be related with a higher concentration in ectopic endometrial cells and an inflammatory reaction which allows higher endothelial permeability, thus allowing it to reach systematic circulation

(Barbieri *et al*., 1986).It is believed that the increase is strongly related to the stage of the disease and presence of ovarian endometriomas, but no to adhesions or small peritoneal implants (Hornstein*et al*., 1995). Superficial endometriosis is believed to show an increase of peritoneal CA 125 levels, while deep infiltrating to serum levels (Koninckx, 1994). Further data suggest that it has a strong positive correlation to ovarian endometriomas (47.7% vs 15.9%), deep infiltrating endometriosis (99.6% vs 78.6%) and increased AFS score (34 vs 6), after surgical confirmation of findings (Zomer*et al*., 2013). It has a significant negative correlation to medical treatment used such as GnRHanalogues or danazol and may therefore be used as a sign of effectiveness or recurrence (Fraser *et al.,* 1989; Franssen*et al*., 1992). After the 3-6 month treatment period, pretreatment levels rebound within 6 months (Dawood*et al.*, 1988).

The assay used is based on a combination of two monoclonal antibodies which bind to the antigen (O'Brien *et al.,* 1991). This assay has replaced previous ones as studies showed increased sensitivity and sensitivity rates (Hornstein*et al.,* 1995). There have been various cut-off points suggested, but most studies and laboratories have introduced the 30-35 IU/ml limit for considering it increased.

There is disagreement whether CA 125 could be used as a diagnostic or screening test for endometriosis. A diagnostic test is an application to patients that seek help in order to identify a cause for their symptoms, while a screening test is an examination that identifies individuals among the general population at increased risk for a disease (Wald, 2001).A screening test needs to satisfy a list of criteria in order to be considered suitable for this purpose. These criteria include (Peters *et al*., 1996; Massad 2008): health relevance (health problem), acceptability of disease (socially acceptable), natural cost (latent or asymptomatic phase), test acceptability (acceptable to women), effectiveness of treatment (early treatment effective), consensus (management options of

abnormalities), complication balance (risk of test lower than disease risk) and cost-benefit balance (test cost acceptable).Out of the eight above criteria, only three to four are partially or fully satisfied (health relevance, acceptability of test, effectiveness of treatment and complication balance, while the other four are not satisfied (Massad, 2008). Evidence therefore do not suggest use of any screening programs for endometriosis (Somiglianaet *al*., 2010).

Use of the assay as a diagnostic tool for endometriosis, especially for stages III-IV is well established, especially in the presence of endometriomas (Mol*et al*., 1998). Correlation with other biomarkers such as CA 19-9, IL6, IL8, TNF-a, VEGF has been also been studied in order to identify possible statistical significance as with CA 125 (May *et al.,* 2010). Furthermore, studies of CA 125 levels in other biological fluids such as peritoneal fluid and endometrial fluid have been made, in order to establish a correlation or not to the disease.

Limitations of the test include presence of other benign gynaecological conditions such as pelvic inflammatory disease and fibromas, fluctuations during menstrual cycle, a lack of a clear cut-off limit and necessity to have an invasive procedure such as laparoscopy in order to stage and correlate the disease.

The goal of this study is a systematic review and meta-analysis of available data of the included studies, regarding sensitivity and specificity of CA 125 in diagnosing endometriosis, correlation to disease stage and possible applications as a diagnostic and predictive tool.

AIMS-OBJECTIVES

This systematic review and meta-analysis aims at studying non invasive diagnostic examinations that will complement or even replace current diagnostic tests for endometriosis diagnosis. Due to the necessity for surgical and histological confirmation of the presence and stage of the disease, the need for such modalities is high. Given the large number of modalities being proposed and tested so far, this review will focus on serum biomarkers that have shown high sensitivity and specificity values, are cost-effective, universally accepted, non-invasive and are standardized. Despite the fact that the criteria for implementing any examination as a screening test for this particular disease are not fulfilled, use of it as diagnostic and predictive examination is possible. Serum biomarkers that have been studied include CA 125, CA 19-9, CA 15-3, CEA, ICAM, VCAM, VEGF, IL and TNFa, the majority of which have shown poor sensitivity, specificity and correlation to disease severity. The only biomarker that studies suggest has demonstrated high sensitivity and specificity rates as well as positive correlation to disease stage is CA 125.

The objectives of this study include a literature review of all recent, relevant published studies that evaluate serum CA 125 in diagnosing endometriosis as well as correlation to disease stage. Another objective is to assess studies that measure CA 125 out of other biological fluids such as peritoneal and endometrial in women diagnosed with endometriosis. The meta-analysis will assess sensitivity, specificity and positive predictive value as well as correlation to disease stage. Comparison of these values with other biomarker values studied in the same population (especially CA 19-9) will be performed as evaluation of peritoneal and endometrial CA 125 sensitivity and specificity rates. Monitoring serum CA 125 after treatment and correlation with recurrence is another objective. All of the above will enable the possible application of CA 125 as a non

invasive diagnostic and predictive marker for endometriosis.

METHODOLOGY

Identification of studies and eligibility criteria

The databases that were searched included Medline, PubMed, Embase, Google Scholar and Cochrane Central Registry of Controlled Trials. The search terms included endometriosis, endometrioma, biomarkers, and CA 125. The search period was up to March 2015 with no time period constrictions and included articles published in English or accepted for publication. Unpublished work was not included.

Selection criteria

Research questions that were considered for the systematic review of the literature included: (i) assessment of specificity, sensitivity and positive predictive value of serum CA 125 in diagnosing endometriosis, (ii) correlation (positive, negative, none) of CA 125 value to endometriosis staging according to ASRM classification, (iii) correlation (positive, negative, none) of CA 125 value to presence of endometrioma, (iv) assessment of specificity, sensitivity and positive predictive value of peritoneal or endometrial CA 125 in diagnosing endometriosis and comparison to results of serum CA 125, (v) correlation of serum CA 125 with other biomarkers such as CA 19-9.

The included studies diagnosed endometriosis by laparoscopy and visual as well as histological confirmation of lesions and staged the disease according to the ASRM or the AFS classification systems described above. Due to the great similarities of the two systems the included studies were not divided according to the system used, since it will not affect the results. All studies were case-control studies, which involved a case group of women diagnosed with endometriosis was compared to a control group of endometriosis-free or normal women. The majority of studies included women with all four stages of the disease (I-IV) and presented results in all case group, as well as in the I-II and III-

IV subgroup of women. The presence or not of an endometrioma was noted in all studies. The sampling criteria was in all cases the etiology for laparoscopy (pelvic pain, infertility, tubal ligation) but dividing the case group into subgroups according to etiology was not performed as in many studies, the exact number of each was not provided. The age limit of the participants was 18-55 years, thus excluding pre-menarchal girls and menopausal women. Racial criteria were not mentioned as race does not play a role in CA 125 level. Sampling of serum CA 125 was performed during menstruation (usually between 3rd to 5th day of cycle) and the assay used to calculate the biomarker was similar in all cases. The evaluation of other biomarkers such as CA 19-9 was noted as comparison to CA 125 is one of the research questions. Studies that measured peritoneal or endometrial CA 125 levels were also included in a separate category and compared to the ones that measured serum CA 125.

Exclusion criteria included presence of other benign gynaecological conditions such pelvic inflammatory disease, leiomyomas, adenomyosis, pregnancy, cancer or hormonal therapy (oral contraceptives, levonorgestrel releasing device, ovulation induction agents, androgens, hormone replacement therapy) for the past 6 months. All these factors are known to influence CA 125 levels.

Data extraction

Extracted data included (i) general characteristics such as authors, year, location, time frame, (ii) study design such as case-control, cohort, prospective, (iii) population characteristics such as sampling method, age, origin, number of participants, (iv) disease staging subgroups (I-II, III-IV) and presence or not of endometriomas, (v) CA 125 values (cut-off limit, day of sampling, mean, median, assay, biological fluid sampled), (vi) other biomarkers measured such as CA 19-9.

Selection bias issues included the sampling method used as all patients were referrals to a tertiary centre due to a condition that required laparoscopic

investigation (pelvic pain, infertility, tubal ligation). Verification bias included the staging system which is considered subjective, the cut-off limit for the assay as it varies slightly among studies and the day of cycle that sampling was performed as during menstruation, levels vary.

Quality assessment

The Newcastle-Ottawa scale (NOS) for assessment of non-randomized studies (case-control or cohort studies) was used (table 1). It is based on three main categories:

(i) study group selection (definition, representation, selection criteria),

(ii) comparability (controls studied to most important factor, compared to other factors), (iii) ascertainment of exposure or outcome (use of patient records, use of a structured interview, non-response rate). Three independent reviewers have assessed all studies included.

table I. Newcastle-Ottawa scale(Wells *et al*., 2000)

selection	**comparability**	**exposure**
1. case definition adequate?	1. basis of design of analysis	1. ascertainment of exposure
a. yes-independent validation⋆	a. most important factor⋆	a. secure record⋆
b.yes-self report	b. additional factor	b. structured interview⋆
c. no description		c. interview not blinded
2. representativeness of cases		d. self report, medical record
a. obvious representation⋆		e. no description
b. potential bias, not stated		2. same method ascertainment
3. selection of controls		a. yes⋆
a. community controls⋆		b. no
b. hospital controls		3. non-response rate
c. no description		a. same rate for both groups⋆
4. definition of controls		b. non respondents described
a. no history of disease⋆		c. rate different, no designate
b. no description of source		

Outcomes of interest

The primary *a priori* end point is the assessment of sensitivity, specificity and positive predictive value of CA 125 (serum, peritoneal, endometrial) in diagnosing endometriosis, as well as correlation (positive, negative, none) of

serum CA 125 to disease staging and presence of endometrioma. The secondary *a priori* end point is the correlation (positive, negative, none) of serum CA 125 and CA 19-9 in diagnosing the disease.

Data analysis

The outcome measures are (i) sensitivity: proportion of actual positives correctly identified as such (true positives / true positives + false negative), (ii) specificity: proportion of actual negatives correctly identified as such (true negatives / true negatives + false positives), (iii) positive predictive value: proportion of actual positives to the total positives (true positives / true positives + false positives) (Fawcelt 2006; Fletcher *et al*., 2005). Correlation refers to statistical relationships between two random variables or sets of data. Correlation coefficients such as the Pearson coefficient (r) measure the degree of correlation (Cohen *et al*., 2002). The data extracted were in numerical form from all studies included and confounding factors recognized included disease staging and presence of endometriomas, for which a detailed separate analysis was performed. Other confounding factors such as menstrual period day, benign gynecological conditions or malignancy were excluded. There is no need for standardization as all studies reported values of CA 125 in U/L. The cut-off value for considering a CA 125 as abnormal is 30-35 U/L in all studies.

Data were assessed for heterogenicity as odds ratio (OR) with confidence interval (CI) of 95% and a level of significance at $p < 0$. Odds ratio refers to a measurement of association between exposure and outcome, with an OR >1 (exposure associated with higher odds of outcome), OR <1 (exposure associated with lower odds of outcome) and OR =1 (no association). Confidence interval (CI) estimates how precise OR is. A large CI means that there is low precision, whereas a small CI means that there is high precision (Szumilas 2010).

Publication bias which assesses the probability of publishing a positive result

than a negative or inconclusive one, is evaluated by the construction of funnel plots of sensitivity, specificity and study size (Song *et al*., 2010).Selection bias refers to the selection of individuals, groups or data for which randomization is not achieved either due to sampling bias (non-randomization), susceptibility bias (one disease predisposes for another disease) or attrition bias (loss of participants). Assessment of this, is achieved by studying the group selection criteria such as recruitment method, confounding factors, etc (Cortes *et al*., 2008).

RESULTS

Characteristics of included studies and quality assessment

The systematic review of electronic databases yielded 423 citations. Screening of the title and abstract which was based on the selection criteria such as endometriosis diagnosis, CA 125 as a diagnostic marker for endometriosis, 367 studies were excluded. The remaining 56 studies were assessed on full text. Exclusion of 43 more articles was based on data duplication (9), inconsistency of selection criteria for inclusion to study (12) and lack of data such as sensitivity, specificity and positive predictive value of CA 125 (22). The remaining 13 full articles were included in the meta-analysis.

All of the studies were prospective case-control (majority) or cohort studies (minority). The case and control groups were selected by patients that were referrals to a tertiary centre in order to have laparoscopy for etiologies such as infertility, pelvic pain, ovarian cyst or requesting tubal ligation. Selection bias is therefore significant as there was non-randomizes selection methodology applied. The case groups were all selected after visual and histological diagnosis of endometriosis and staged (I-IV) according to the ASRM or the AFS classification system. Due to the close association of the two systems, studies were not excluded based on the classification method followed, as both systems are very similar in the criteria used. The control groups were considered normal or endometriosis free as visual and histological criteria for the disease were negative. In case of any other gynecological benign etiology such as fibroids, pelvic inflammatory disease, adenomyosisor cancer being diagnosed during surgery, candidates were excluded from both groups. Other exclusion criteria included previous treatment (hormonal, surgical) for endometriosis in the previous 6 months. Women taking hormonal contraception were also excluded. In all studies there was not ethnic selection or exclusion process as evidence

suggest, there is no relationship to ethnic origin of the disease or CA 125 levels. Most studies were performed among Caucasian and Asian descent women.
All studies included women between 18-50 (reproductive age group) and no study included premenarchal or menopausal women, since the disease is predominantly affecting women of reproductive age.

Cut-off values are calculated in U/mL and are almost identical in all studies (30-35 U/mL), while in studies that provide various cut-off points (20, 30, 35), the results of the 35U/mL cut-off point have been recorded. The sample collection time was during menstrual period in all studies. The assay used for analysis was performed by the chemiluminescence method (CLEIA) by using the same kit in a single assay for each study included. CA 125 binds to anti-CA 125 monoclonal antibody (ALP labeled) which are then immobilized on particles. The particles in turn are washed and a substrate solution is added in order to achieve a cleavage reaction and generation of a luminescent signal. All measurements are presented as mean +/-SD.

Study size was not considered as a selection criterion, therefore small and large studies were included in the meta-analysis. A funnel plot has also been constructed in order to assess for publication bias issues.

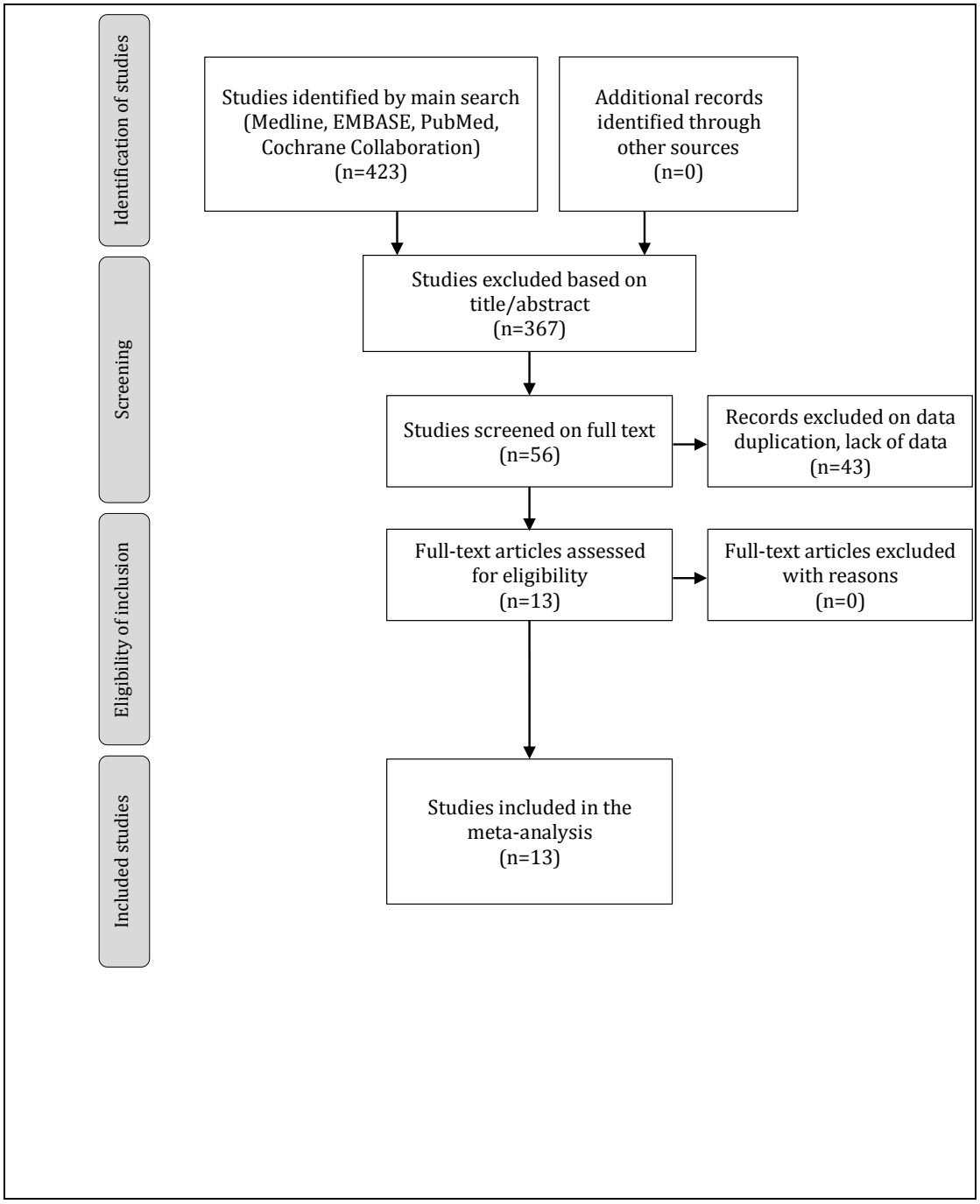
Performance bias was detected in 4 studies which did not reported if a transvaginal ultrasound was performed when reporting on endometriomas or other benign gynecological pathology, and in 3 studies that failed to report the mean age of the case and control group, although stating that were in the reproductive age group.

Attrition bias was reported in 5 studies as they failed to provide complete data regarding specificity, sensitivity, positive predictive value or area under curve. Detailed analysis of all studies reports, follows in the next segment.

Withdrawals or loss to follow up is not reported in the majority of studies, but

since there were all referral cases for surgery, the withdrawal rate is reported as small in the majority of them.

table II. Flow chart for systematic review and meta-analysis



Studies included

table III: characteristics of included studies

author	type	Case (N)	mean	sensitivity	specificity	PPV	AUC
Anastasi et al., 2013	Case control	57	46.1	-	-	-	-
de LRamos et al 2012	Case control	44	51.98	-	-	-	-
Vodolazkaia et al., 2012	Case control	232	20	86	63	-	0.85
Prayudhana et al., 2012	cohort	40	24	67	90	-	0.870
Mihalyi et al., 2010	Case control	201	22	67.9	71.1	82.8	0.772
Kurdoglu et al., 2009	Case control	101	52.56	86	61	-	-
Martinez et al., 2007	Case control	47	33.3	47	97	89.0	0.759
Kitawaki et al., 2005	Case control	247	72.8	44.3	97	92.9	0.935
Somigliana et al., 2004	Case control	45	23.4	27	94	86	
Kafali et al., 2004	Case control	-	35.8	93	92	-	0.830
Amaral et al., 2006	Case control	35	39.1	-	-	-	-
Harada et al., 2002	Case control	101	69.8	49	100	100	-
Chen et al., 1998	cohort	131	56	61.1	87.5	-	-

table IV: critical appraisal of included studies (Newcastle Ottawa scale)

study	selection	comparability	exposure
Anastasi et al 2013	⋆		⋆⋆
de L Ramos et al 2012	⋆⋆	⋆	⋆⋆⋆
Vodolazkaia et al., 2012	⋆⋆		⋆⋆
Prayudhana et al., 2012	⋆		⋆⋆
Mihalyi et al., 2010	⋆⋆	⋆	⋆⋆⋆
Kurdoglu et al, 2009	⋆⋆	⋆	⋆⋆⋆
Martinez et al., 2007	⋆⋆⋆	⋆	⋆⋆⋆
Kitawaki et al., 2005	⋆⋆⋆	⋆	⋆⋆⋆
Somigliana et al., 2004	⋆⋆⋆	⋆	⋆⋆⋆
Kafali et al., 2004	⋆⋆	⋆	⋆⋆
Amaral et al., 2004	⋆⋆		⋆⋆
Harada et al., 2002	⋆	⋆	⋆⋆
Chen et al., 1998	⋆⋆		⋆⋆

table V: Mean CA 125 values of case and control groups

study	Mean control CA 125	Mean case CA 125
Anastasi et al., 2013	16.8	46.1
de Luna Ramos et al., 2012	21.20	42.23
Vodolazkaia et al., 2012	12	20
Mihalyi et al., 2010	13	22
Kurdoglu et al., 2009	15.92	52.56
Martinez et al., 2007	15.5	51.3
Kitawaki et al., 2005	14.8	72.8
Somigliana et al., 2004	11.4	23.4
Kafali et al., 2004	12.2	35.8
Amaral et al., 2006	10.5	39.1
Harada et al., 2002	11.3	69.8

table VI: Mean CA 125 values in stages I-II and III-IV

study	Mean stage I-II CA 125	Mean stage III-IV CA 125
De Luna Ramos et al., 2012	41.79	59.73
Prayudhana et al., 2012	12	36.3
Mihalyi et al., 2010	17	32
Kurdoglu et al., 2009	32.29	72.84
Martinez et al., 2007	15.3	51.3
Somigliana et al., 2004	13.3	25.9
Amaral et al., 2006	14.8	115.3
Harada et al., 2002	52.2	95.8
Chen et al., 1998	26.8	86.7

table VII: Mean CA 125 values of case subgroup with endometriomas

study	Presence of endometrioma CA 125
Anastasi et al., 2013	46.1+/-34
Kitawaki et al., 2005	50.8+/-68.6
Somigliana et al., 2004	25.9

table VIII: Mean peritoneal CA 125 values in stages I, II, III, IV

study	Stage I	Stage II	Stage III	Stage IV
Amaral 2006	808.75	981.60	1554.30	2817.00

Critical appraisal of included studies

Anastasi et al., performed a case-control study of 50 healthy women and 57 women diagnosed with ovarian endometrioma on ultrasound and were subsequently diagnosed and staged with laparoscopy according to the rASRM criteria. Selection bias was evident as the case group was constructed out of referral cases for ovarian endometrioma, while there is no report in regards to the control group (selection criteria), but only referred to as negative for ovarian mass. This is a weakness of the study as endometriosis may be present without an ovarian cyst, and therefore some women that were recruited as controls may have minimal or mild disease. Exclusion criteria included hormonal therapy, pregnancy, chronic conditions and malignancy. It is not clarified though, what chronic conditions include (fibroids, adenomyosis, pelvic inflammatory disease or not) because many benign conditions affect CA 125 levels. Disease staging as mentioned above was performed according to the rASRM criteria and results presented indicate an overrepresentation of stage III-IV (51 out of 57 cases) and an underrepresentation of stage I-II (6 out of 57). As a result of this, the CA 125 measurements will probably be higher than anticipated as they seem to correlate well with disease stage. This will be a problem in including the results of this study in the meta-analysis. The mean age of participants was reported at 30 (21-57) for the control group and 36 (23-48) years for the case group, which is similar to that reported by the majority of studies. CA 125 assay used was the same as in all other studies, but this study failed to report the day of sample retrieval (menstruation or not) which is important as it affects the levels of the biomarker. The cut-off limit for considering abnormal is set at 35 U/ml as in the majority of studies. Results from the case group were reported as mean with SD (46.1 +/- 34) and a percentage above cut-off limit (56.1% of case group). There was not a breakdown of CA 125 levels among stage I-II and III-IV subgroups, nor

a sensitivity, specificity and positive predictive value for endometriosis was provided. Inclusion of this study to the meta-analysis will be problematic and will be avoided as it lacks data and the data provided are considered insufficient for the scope of this review. It is important to note though that this study also evaluated the use of other biomarkers such as human epididymis protein 4 (HE4) and cancer antigen 72-4 (CA 72-4) in order to differentiate between benign and malignant ovarian cysts. The results of this study show and agree with results of other studies (Moore *et al.,* 2012) that HE4 and CA72-4 are increased in ovarian cancer but not in endometriosis. This may enable differentiating benign endometriomas from ovarian carcinomas based on ultrasound and serum biomarker examination alone.

De Luna Ramos et al., conducted a case control study in order to evaluate serum CA 125 and CD-23 in endometriotic and normal women. The case and control groups were referral cases to a tertiary centre for laparoscopy indications such as infertility pelvic pain and tubal ligation. Therefore selection bias is evident. The study size was rather small as the case group was composed of 44 women and the control group of 58. Non responders were 1 in each group. Inclusion criteria for the case group was histological confirmation of endometriosis and staging according to the revised ASRM criteria. The control group consisted with women that had no evidence of endometriosis. Exclusion criteria for both groups were hormonal therapy in the previous 3 months and autoimmune disease. Other criteria such as benign gynecological conditions were not reported as exclusion criteria. Staging of the 44 case group patients was reported as I-II (n=19) and III-IV (n=25) which is representative of all stages of the disease. The mean of both groups was not reported but was stated that all participants were among 18-45 years of age. CA 125 determination was conducted by the ELISA kit and samples were drawn during the menstrual period in all subjects. A second

sample was taken after menstruation in order to estimate the effect of menstruation on CA 125 levels. The cut-off limit is considered as a range 25-35 U/ml , a fact that may cause problems in interpretation of results. Results of the study was the mean CA 125 in both case and control group and a subgroup (I-II, III-IV) mean CA 125 result. Sensitivity, specificity and positive predictive value were not reported, a fact that will exclude this study from the part of meta-analysis that will evaluate these figures. Other reported findings included a serum CD-23 measurement in order to estimate a possible statistical significance among case and control groups. In conclusion, this study suggests that there is significant difference among case and control groups and among stage I-II and III-IV subgroups regarding CA 125 but no significant difference regarding CD-23. Vodolazkaia et al., conducted a case-control study of 121 controls and 232 case women, who were referral cases at a tertiary laparoscopic centre for infertility or pelvic pain. The study size is large, probably one of the largest of included studies. Due to lack of community recruitment, selection bias is present. Non-responders rate was not reported. Inclusion criteria was histologic confirmation of endometriosis, which was also staged according to the ASRM criteria to mild-minimal and moderate-severe, something that will enable comparisons of CA 125 values among the two subgroups. Both subgroups were adequately represented (I-II 148, III-IV 84). The inclusion of the subgroup “US-negative endometriosis” (175 out of 232 endometriotic) which includes sonographic normal cases with histological diagnosis of the disease causing somewhat confusion in terms of staging as there is not such a category in the ASRM classification. Exclusion criteria included hormonal medication or related surgery for the past 6 months. There was no exclusion of other benign gynecological conditions that may affect CA 125 levels, although documented and presented post-laparoscopy. Mean age of the control and case groups was

similar (31.7 to 31.2 years respectively). The CA 125 assay used was based on automated immunoassays and plasma samples were taken on the day of surgery, which varied (follicular, luteal), a fact that will influence the values. A separate analysis was performed according to the day of the cycle the sampling was made. The cut-off limit that this study used was lower compared to the rest (11.5 U/ml compared to 30-35 U/ml), a fact that will not allow data to be included in the meta-analysis of sensitivity, specificity and positive predictive value. The other disadvantage is that there is no report of the values in the two subgroups (I-II, III-IV). Results show that CA125 has statistical significant increase in endometriosis patients than controls, even in those with US-negative for the disease. Sensitivity is 86% and specificity 63% for an AUC of 0.85 for US-negative endometriosis. Data for all endometriotic patients are not available. This study also studied a panel of 27 other biomarkers and concluded that a combination of CA 125, annexin V, VEGF, glycodelin and slCAM-1 may improve sensitivity and specificity in diagnosing the disease.

Prayudhana et al., reported on a cohort of 40 Indonesian women of reproductive age that were diagnosed with endometriosis following laparoscopy for various etiologies. The study size was rather small but representation of all disease stages was adequate (stages I-II 26, stages III-IV 14). Classification followed the revised ASRM criteria and histological confirmation of the disease was required in order to be included. There was no report on non-responder rate, exclusion criteria or cycle day of surgery or CA 125 blood sampling. The mean age of the two subgroups was 34.1 and 30.3 respectively. Results reported by this study suggest there is a statistically significant difference of CA 125 levels between stages I-II and III-IV (12.0 and 36.3 U/ml) and by using a cut-off limit of 35 U/ml, sensitivity and specificity are 67% and 90% for an AUC of 0.87. This study suggests that by using a lower cut-off limit of 16.9 U/ml, is better in

differentiating among stages I-II and III-IV. Sensitivity reaches 83% but for a lower specificity rate of 81%. This study reported in using biomarkers such as CRP and ESR in disease diagnosis but failed to prove an association to disease stage.

Mihalyi et al., conducted a case control study of women referred to a tertiary centre for laparoscopy with an indication of infertility. They were assigned either to the control group if there were no visual or histological evidence of endometriosis or to the case group if there were histological evidence of endometriosis. The control group consisted of 93 women and the case group of 201, sufficient numbers to consist a large study. Due to the lack of community recruits for the control group, selection bias is noted. The exclusion criteria for both groups included hormonal or surgical treatment in the past 6 months, or the presence of a benign gynecological condition such as PID or leiomyomas. Non responders were reported for both groups. The case group was further staged according to the ASRM criteria into minimal-mild and moderate-severe subgroups. The minimal-mild subgroup consisted of 132 women, while the moderate-severe of 69 women, numbers that indicate that there is adequate representation of both categories. The mean age of each group was not reported but it was noted that all participants were reproductive aged women. CA 125 was collected on the day of surgery, irrespectively of cycle day, but detailed analysis of each category according to cycle day (menstrual, proliferative, secretory) was conducted. The collection protocol as well as the assay used was the standard as for all the included studies. No cut-off limit was reported as there was not in the study design to compare percentages of case or control group that was above or below a certain level. Levels for CA 125 for both case subgroups were presented. This study concluded that there was significant statistical difference of CA125 levels between controls and endometriosis women

regardless of disease stage and cycle day sample collection [13 U/ml (4.0-47.0) vs 22 U/ml (6-969.0) ,p<0.0001]. It also reported statistical significant difference between minimal-mild and moderate-severe subgroups [17.0 U/ml (6-969.0) vs 32 U/ml (9-746), p<0.0001]. By using a stepwise logistic regression model performance of CA 125 was measured. Sensitivity was 67.9%, specificity 71.1%, positive predictive value 82.8% for an AUC 0.772. These data will be included in the meta-analysis. Plasma levels of Interleukin 6 (IL-6) and Interleukin 8 (IL-8) were also found to be significantly higher in endometriosis than in controls. Kurdoglu et al., performed a case control study of 101 endometriotic women and 29 normal women. These women were all referrals to a tertiary centre for various indications for laparoscopy (infertility, sterilization, pelvic pain), and therefore a substantial selection bias was noted. All women underwent laparoscopy and the diagnosis was confirmed by histological criteria. Disease staging was conducted according to ASRM guidelines. There was under-representation of stages I-II (26) to stages III-IV (75), a fact that may affect mean CA 125 levels. Exclusion criteria included all benign and malignant conditions that may cause CA 125 increase (myoma, pelvic inflammatory disease, etc), but it was not mentioned if hormonal therapy was one of them. Although the mean age of the two groups was not reported, there was note that both groups included reproductive aged women. The CA 125 protocol used was the same as in all other included studies, but there was no sampling on a particular cycle day, instead it was performed on the day of operation. This may interfere with results as menstrual period days affect CA 125 levels. The cut off limit for the assay was 35 U/ml. This study reported that CA 125 levels in stage III-IV women was significantly higher that stages I-II and controls. Sensitivity and specificity was 86% and 61%, but there was no data on positive or negative predictive value. This study also assessed CA 19-9 sensitivity and specificity

rates which were similar to that of CA 125 (89% and 52% respectively).

The study by Martinez et al., is a small case control study of 47 and 38 women respectively, all of which were referral patients for laparoscopy for various etiologies. The cases were selected after histological confirmation of endometriotic lesions and staged according to the ASRM scoring system. There seems to be an overrepresentation of stage III-IV cases (36) to stage I-II (11), a fact that may alter the mean CA 125 levels for the case group. Exclusion criteria were more strict than other studies, as they excluded women that had taken hormonal treatment for the past 2 years and not for the past 6 months as all other studies suggest. The mean age of the case and control groups were 33.2 and 35.5, with all participants belonging to the reproductive age group. CA 125 was obtained in all women during the early follicular phase, something that eliminates the effect of menstrual cycle day on results. The cut-off limit used was similar as other studies (35U/ml) which enables comparison of the results of this study to others included. This study suggests that CA 125 is significantly elevated in stages III-IV, with a sensitivity and specificity rate of 47% and 97.5% respectively. Positive predictive value was 89.0% with an AUC of 0.759 (95% CI 0.670-0.833). This study also assessed the use of IL-6 as a non-invasive marker and concluded that has a sensitivity and specificity of 75% and 83% for stages I-II.

Kitawaki et al., conducted a case control study of normal and endometriotic women who were all referred to a tertiary laparoscopy centre in Japan. The study numbers were adequate as 101 women enrolled to the control group and 249 to the case group. In all endometriotic women, histological confirmation of the disease was performed prior to inclusion, while exclusion criteria for both groups included hormonal treatment, pregnancy, malignancy and presence of leiomyomas. Disease classification did not follow the ASRM revised criteria, but

focused primarily on the presence or absence of an endometrioma. Although one might argue that the presence of an endometrioma is a sign of stage III-IV disease, a decision was made not to arbitrarily consider it as such and therefore not include related results in the meta-analysis. The relevant results for the total population of endometriotic women will be included. All included women were of reproductive age, but specific mean age of both groups was not provided. The CA 125 assay used and collection methodology was similar as the one in all other studies but the day of collection was not specific, but avoided menstrual cycle days only. A specific cut-off limit was not available but researches used three cut-off limits (20, 26 and 30 U/ml) in order to assess sensitivity, specificity and positive predictive value for each. The included numbers in the meta-analysis will be that of 30 U/ml as a cut-off limit which is closer to that used by other included studies. Results show that endometriotic women have significant higher CA 125 than normal women (72.8 U/ml and 14.8 U/ml) respectively, while endometriotic women have significant lower CA 125 level than those with no endometriomas (50.8 U/ml and 84.8 U/ml). For a cut-off limit of 30 U/ml, sensitivity was reported 44.3% for women without endometriomas and 78.9% for those with endometriomas, while specificity was 97% and positive predictive value 92.9%.

The study by Somigliana et al., is a case-control study that was performed among women of reproductive age (18-45y), with a mean age of 32 years for both groups. Both groups consisted of women that were referrals to a tertiary centre for laparoscopy due to various etiologies such as pelvic pain or infertility. Significant selection bias is obvious as community controls were not selected. Cases were women that had a histological confirmation of pelvic endometriosis. Staging of the disease was carried out according to revised ASRM guidelines. Out of 45 cases, stage I-II consisted of 14 patients and stage III-IV of 31 patients,

which indicate an over-representation of stages III-IV. Also, due to the small number of the study, the impact of this may produce higher than expected CA 125 values. Exclusion criteria were noted and included malignancy, pregnancy and hormonal use for the previous six months. There was no mention for other benign gynaecological conditions, which may influence CA 125 levels. The CA 125 assay used was similar to the rest of included studies and the cut-off limit for assessing sensitivity and specificity values was set at 31U/ml. The day of sampling varied among patients and was performed on the day of surgery, therefore affecting CA 125 levels. As shown in the relevant table, endometriotic women had significant higher levels of CA 125 than normal women. The biomarker level in stage III-IV women were significantly higher than stage I-II women. Sensitivity and specificity was calculated at 27% and 94% respectively, with a positive predictive value of 86%. This study also assessed IL-6 and CA 19-9 as possible diagnostic biomarkers for endometriosis. It concluded that women with and without endometriosis have similar values and there was no statistical significant difference in endometriotic and normal women.

Kafali et al., conducted a case control study of 28 women who underwent laparoscopy at a tertiary centre for infertility. The study size is very small compared to other studies (28 women in total for both groups), and considerable selection bias exists as all women were infertile women who underwent laparoscopy as part of their investigation. The inclusion criteria consisted of visual and histological confirmation of endometriosis, while exclusion criteria of co-existing malignancy, pregnancy, or other etiologies that may increase CA 125 levels. Staging of case women was performed based on the revised ASRM criteria. Similar CA 125 collection methodology was followed as most studies, but collection day varied among participants, thus allowing day of menstrual cycle have an effect on results. Results show a significant difference of CA 125

levels between controls (12.2 U/ml) and cases (35.8 U/ml), while sensitivity and specificity were 93% and 92% respectively. Sensitivity was markedly increased compared to other studies, a fact that will be addressed during the meta-analysis.

The study by Amaral et al., is a case control study of 52 women referred for laparoscopy because of infertility investigation, pelvic pain investigation or request for tubal ligation. Study size was small as the two groups consisted of 17 and 35 women respectively. Considerable selection bias exists as all women were referrals and no community controls were included. Inclusion criteria for the case group were visual and histological confirmation of endometriosis, while disease staging complied with the ASRM revised set of criteria. Exclusion criteria included pregnancy, PID, hormonal medication for the past 6 months and ovarian neoplasm. The mean age for the case group is somewhat smaller than that of the control group, but as age is not a confounding factor of CA 125 levels, it is not relevant. Day of sampling and surgery was the same (early follicular stage), a fact that limits the effect of menstrual cycle day on CA 125 levels. The assay and collection protocol used is similar to the rest of studies included. Results show a significant difference of serum CA 125 levels between controls and cases, as well as between stages I-II and III-IV. There was no statistical analysis regarding sensitivity, specificity and positive predictive value. This study also reported on peritoneal CA 125 levels and compared them among cases and controls as well as among stages I-II and III-IV. Results show that peritoneal fluid CA 125 levels are significantly higher among endometriotic women and correlate well to disease stage. It is also related to superficial peritoneal disease (pigmented, non-pigmented lesions) as well as deep disease.

Harada et al., conducted a case control study among a homogenous Japanese population of women that were referred for laparoscopy with various

indications. The study population of cases was rather large (101) but the controls were small (22). The inclusion criteria consisted of histological confirmation of endometriosis, while the exclusion were presence of pregnancy or malignancy. The presence of myoma or endometriosis or a past history of hormonal treatment of symptoms were not part of the exclusion criteria. The cases were staged according to the AFS criteria, but due to little difference to the criteria set by the ASRM, the study will be included in the meta-analysis. There was an over-representation though of stages III and IV (28 and 35) respectively and an under-representation of stage II (8 women). Mean age for control and case groups was 33.3 and 35.9, which has no statistical significant difference. The assay used was the one used in all studies, but the day of sampling was not a specific cycle day, but a random one. Cut-off limit of the assay for assessing sensitivity and specificity was 35 U/ml. There is a statistically significant difference of CA 125 levels between normal and endometriotic women (11.3 and 69.8 U/ml), and among stages I-II and III-IV (52.2 and 95.8 U/ml). These results are somewhat higher than other similar studies, but it may be explained by the higher number of severe disease women participated, participation of women with leiomyomas and endometriosis as well as a non-specific cycle day of blood sampling. The sensitivity and specificity rates reported are 49% and 100% respectively, which again differ from other studies, in terms of lower sensitivity and higher specificity figures, something that requires further analysis during the meta-analysis. This study also assessed CA 19-9 sensitivity and specificity for disease diagnosis. By using a cut-off limit of 37 U/ml, sensitivity was reported at 34% which is low, with a 100% specificity rate. It is considered a useful biomarker according to the authors for disease severity determination.

Chen et al., conducted a cohort study of 131 women diagnosed with various stages of endometriosis following laparoscopy for dysmenorrhea, in a

homogenous Taiwanese population of reproductive aged women. The study size was large and included a good representation of stages I-II and III-IV (56 and 75 women respectively). Exclusion criteria did not include hormonal treatment (especially danazol) or any other benign gynecological conditions. Staging was performed by the AFS criteria, which as previously mentioned are similar to the revised ASRM, so results will be included in the meta-analysis. The day of blood sampling varied among participants, a fact that may influence results. By using a cut-off limit of 35 U/ml, 26.8% of stage I-II had higher levels compared to 86.7% of stages III-IV. The reported sensitivity and specificity was 61.1% and 87.5% respectively. Results reported from this study will be handled with caution as hormonal treatment with danazol was not an exclusion criterion.

Figure I. Funnel plot of sensitivity-number of patients (cases)

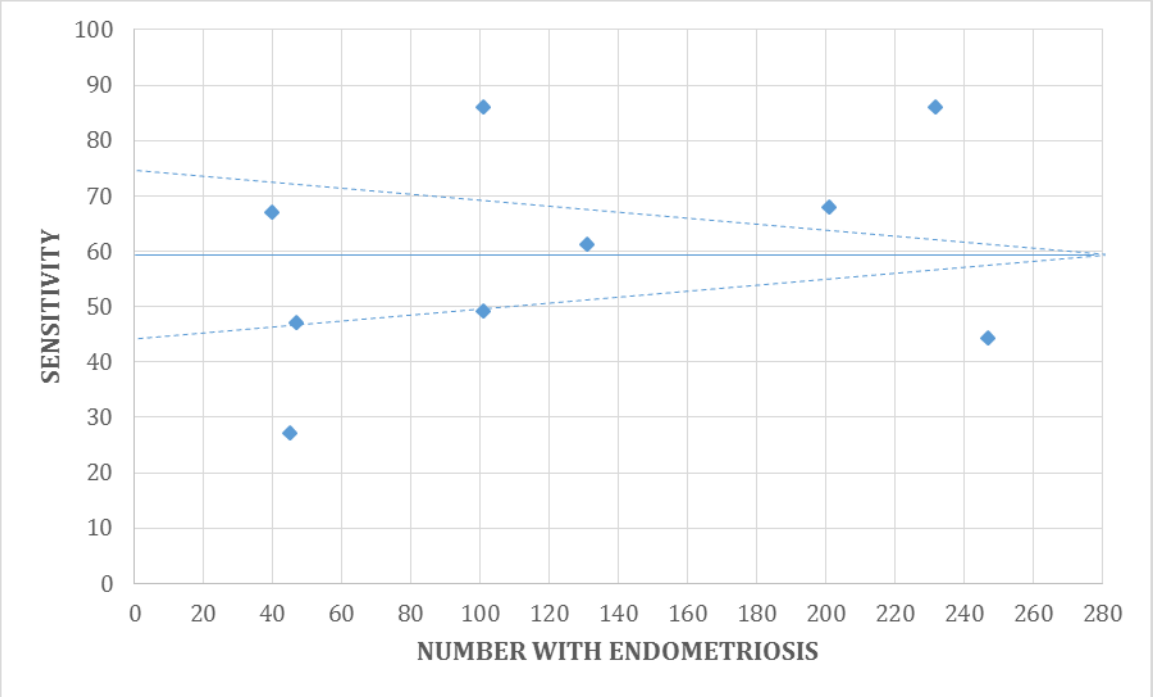
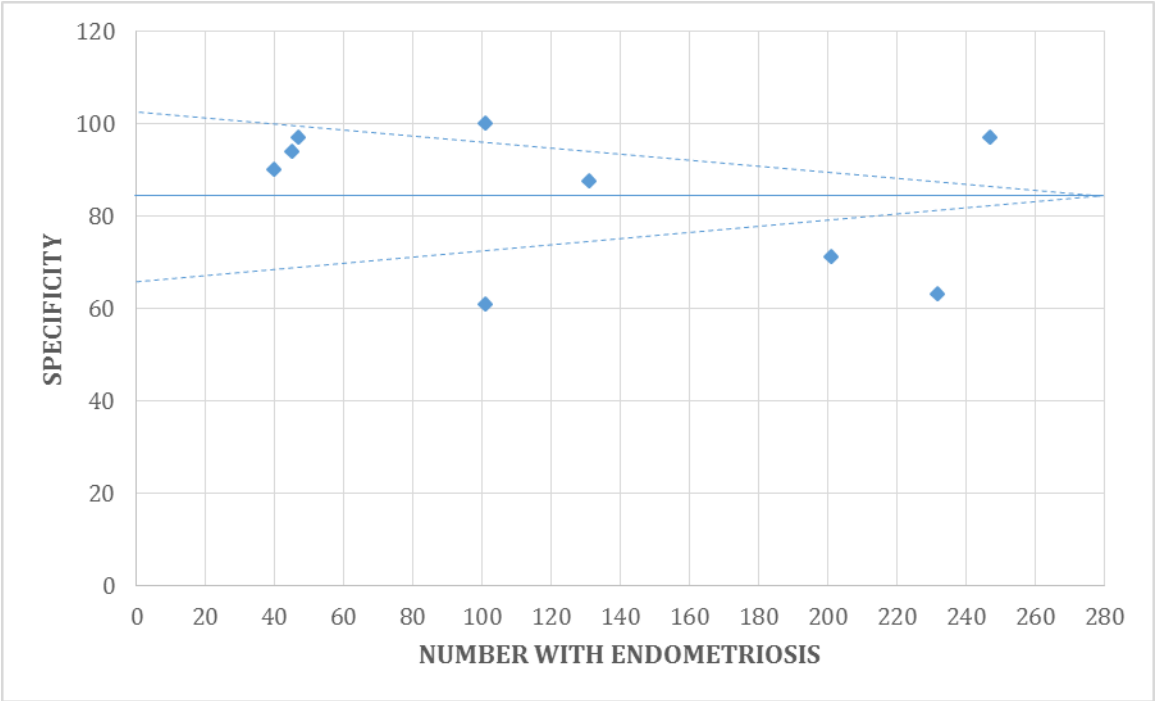


Figure II. Funnel plot of specificity-number of patients (cases)



DISCUSSION

Principal findings

Sensitivity, specificity and positive predictive value for endometriosis diagnosis:

Extracted data from the 13 studies included in the meta-analysis were used in order to assess mean values of sensitivity, specificity and positive predictive value. The mean age for cases and control groups was similar (~30 years), with an age range of 18-45 years. As endometriosis is considered a disease of reproductive age women, no adolescents or menopausal women were included. Data regarding sensitivity were provided by 10 of included studies, ranging from 27% to 93%. Due to the substantial difference between reported values, further analysis is required. The smallest sensitivity rate is reported by Somigliana et al., who conducted a small case-control study of 45 endometriotic women. The mean CA 125 value of cases was also one of the smallest reported, while specificity was high. The critical appraisal of the study did not reveal any problems in selection, comparability or exposure, except of the fact that the blood sampling day was not the same in all women. Three more studies reported sensitivity values below 50%, with no common features in their study design such as size or inclusion criteria used. Two of these studies were among the largest included while the third was one of the smallest. Specificity rates were among the highest reported in all three studies (97%, 97%, and 100% respectively). The mean sensitivity rate is estimated at 62.83%, for all stages of endometriosis. There is not a reported sensitivity rate for stages I-II and III-IV due to lack of specific data from all studies included. The reported rate is higher than what was reported by previous meta-analyses of case-control or cohort studies. In a similar meta-analysis conducted by Mol et al., prior to1998, it concluded that sensitivity rates of CA 125 were significantly smaller than those reported by later studies for all stages of disease. Cohort studies seem to report

higher sensitivity rates than case-control studies, but overall for a specificity rate of 90% the reported sensitivity was only 28% and in order to increase sensitivity to 50%, the specificity dropped to 72% (Mol*et al*., 1998). The study by Vodolazkaia et al., which is a large case control study of 232 endometriotic women, reported a sensitivity value of 86% for an AUC of 0.85 for 95% CI. The mean CA 125 value was 20 U/ml for a cut-off value of 11.5 U/ml which is one of the lowest reported. This may explain the high sensitivity rate reported. The other study that reported higher than expected value was a case control study by Kurdoglu et al. The study size was medium (101 cases) but with an over-representation of stages III-IV (75) which might offer an explanation for the increased sensitivity rate of 86%. An AUC was not reported for this study. The cut-off limit was at 35U/ml which is comparable with most studies. The paradox of this study is the lower specificity rate reported (63%) as it is the same with the Vodolazkaia study (63%). The highest sensitivity rate reported (93%) was by Kafali et al., for an AUC of 0.830 (95% CI). The study size is the smallest included (28 women of both groups),but due to lack of data there was not a total number of cases reported, nor stages I-II and III-IV. This does not allow further analysis of the reported sensitivity rate. The mean CA 125 value of both groups though are comparable to other studies (12.2 U/ml and 35.8 U/ml respectively). Serum CA 125 is not a highly sensitive test in order to be used as a screening tool for detecting endometriosis, but is of use as part of the diagnostic workout in women which have related symptomatology or imaging signs of the disease (endometriomas).

Specificity rates are more consistent that sensitivity rates reported. The mean specificity rate is 85.26%, which is similar to the mean rate reported by the meta-analysis conducted by Mol et al. Out of the 10 studies, 8 have reported rates around 90% (87.5 to 100%) with only two studies reporting rates much lower

(~60%). The two studies that reported such low numbers are the same that reported higher than expected sensitivity rates, as reported above. The study by Vodolazkaia et al., has published a very low CA 125 cut-off value, so the low specificity rate is probably attributed to that. In regards to the study by Kurdoglu et al., the low specificity rate is not attributed to any apparent reason as study critical appraisal shows no significant problems except of a high mean CA 125 value (52.56 U/ml) which is the only probable etiology for the low rate reported. Overall, this meta-analysis confirms results of previous analyses that serum CA 125 has a high specificity rate in diagnosing endometriosis.

The positive predictive value is reported by only five studies included with a mean value of 90.14%, which is anticipated as it correlates well with specificity values. All included studies that reported positive predictive values have also reported AUC that range from 0.759 to 0.935.

Correlation of CA 125 to endometriosis staging:

All included studies have reported on mean control and case values of serum CA 125. The case groups have reported mean values in total and in stages I-II and III-IV. Mean control values range between 10.5 U/ml (Amaral*et al*.) and 21.20 U/ml (de Luna Ramos *et al*.). The mean control group value is calculated at 14.05 U/ml, a value which is well below the 30-35 U/ml cut-off limit, used by the majority of studies. By keeping the cut-off limit at such levels, it is possible to identify the large majority of endometriotic women and avoid false positive results. The mean case group values range between 20 U/ml (Vodolazkaia*et al*.,) and 72.8 U/ml (Kitawaki*et al*.), with a mean case group value of 43.19 U/ml. As shown in table V there is a wide range of case group mean values, a fact that requires further analysis. In most studies the mean value is above the 30-35 U/ml but there are studies that fall short of the cut-off value. Vodolazkaia et al., who reported the lowest case group value, have included 232 cases, of which

stage I-II 148 and stage III-IV 84. The inclusion of many stage I-II women may have contributed to the low figures and further analysis is needed, when the mean stage values are addressed. The same pattern applies for the results by Mihalyi*et al*., who also reported a mean case group value of 22 U/ml. The stage I-II women are 132 out of 201, but again mean values of stage I-II and III-IV need to be analyzed further in order to confirm or not this observation. Kitawaki*et al*., who reported the highest case group value, have included cases with endometrioma presence alone as stage III-IV, which is not what the revised ASRM criteria suggest. Further analysis of the included endometrioma-present and endometrioma-absent patients is needed, as this could be a confounding factor. The study by Harada *et al*., also reported high values (69.8 U/ml), a fact that is probably attributed to the over-representation of stage III-IV women (63) to stage II (8). Overall, it is needed to carefully select the case population, as over-representation of stages III-IV may produce false high results and follow the revised ASRM criteria for classification in order to avoid classification of minimal and mild cases to moderate and severe cases. A mean total endometriosis value of more than 35 U/ml will enable us in differentiating cases from controls.

A subgroup analysis of case CA 125 mean values is presented in table VI. Cases are classified according to ASRM revised criteria into stage I-II and stage III-IV. Results from included studies vary significantly in terms of range and mean levels but all agree that there is statistically significant difference between mean stage I-II and III-IV levels. Mean stage I-II CA 125 values range from 12 to 52.2 U/ml, with a mean value of 25.05 U/ml which is below the 30-35 cut-off limit. It is therefore apparent that CA 125 does not perform as well in stage I-II as in stage III-IV cases. Mean stage III-IV level is 63.98 U/ml, well above the cut-off limit and with a significant difference from stage I-II mean level. From all the included studies only a single one has published a mean stage III-IV below 30

U/ml. Somigliana*et al*., which reported this low value conducted a case control study including 45 cases, 14 of which were stage I-II and 31 stage III-IV. No clear etiology is found for this low value, as the revised ASRM criteria were used to stage the disease and the same assay was used as in the rest of included studies. Overall CA 125 is a sensitive and specific biomarker for disease staging.

Correlation of CA 125 to presence of endometrioma:

Three of the included studies reported mean serum CA 125 values in the presence of endometriomas confirmed by laparoscopy. The reported values are compared to mean serum CA 125 of stages I-II and mean CA 125 of total cases, in order to investigate a possible correlation of values to presence of the lesion.

Anastasi et al., reporter a mean value of 46.1 U/ml for the endometrioma subgroup, but did not report on mean stage I-II values or total cases values, in order to compare them to the endometrioma subgroup. The 46.1 U/ml value though, has a statistically significant difference from the mean control value of 16.8 U/ml reported in the same study.

Kitawaki et al., reported a mean subgroup value of 50.8 U/ml. The endometrioma subgroup included 88 cases and a ROC analysis reported an AUC of 0,935 for a $p<0.05$, while the non-endometrioma subgroup (161 cases) reported a mean value of 84.8 U/ml. The ROC analysis reported an AUC of 0.788 for a $p<0.05$. These results indicate that the non-endometrioma subgroup has a statistical significant difference compared to the endometrioma subgroup. For a cut-off value of 30 U/ml specificity is at 97%, while sensitivity for the non-endometrioma subgroup is at 44.3% and for the endometrioma subgroup at 78.9%. Positive predictive value is 92.9% and 97.7% respectively. This study concluded that in the presence of endometrioma, serum CA 125 has a higher sensitivity rate for any cut-off value (20, 26 or 30 U/ml). Finally, when compared to the mean control CA 125 value of 14.8 U/ml, a statistical significant difference

from the endometrioma subgroup of 50.8 U/ml is reported.

Somigliana et al., reported a mean serum CA 125 of 25.9 U/ml in the endometrioma subgroup, the lowest of the three included studies. The subgroup included 31 women diagnosed with endometrioma by laparoscopy and were compared with the 49 women of the subgroup without endometrioma. The non-endometrioma subgroup had a mean serum CA 125 of 12.9 U/ml, with a statistical significant difference compared to the endometrioma subgroup.

Overall, the meta-analysis of the three included studies regarding the correlation of CA 125 value to the presence of endometrioma, indicates that there is a positive correlation between the two.

<u>Sensitivity, specificity and positive predictive value of peritoneal fluid CA 125 for endometriosis diagnosis:</u>

Amaral et al., conducted a case control study of 35 endometrioticand 17 normal women ,who underwent laparoscopy at which peritoneal fluid collection was performed. Serum and peritoneal fluid CA 125 were determined and assessed for correlation to disease stage and to each other. In the case group, mean serum CA 125 was 39.1 U/ml and mean peritoneal fluid 1469.4 U/ml, while for the control group the values were 10.5 U/ml and 888.7 U/ml respectively. There is positive correlation of serum and peritoneal fluid values as well as to disease stage of both values.

<u>Correlation of serum CA 125 to other biomarkers in endometriosis diagnosis:</u>

Cancer antigen CA 19-9 is one of the most frequently assessed biomarkers besides CA 125 in endometriosis diagnosis. Harada et al., reported a statistical significant difference of mean CA 19-9 among case and control women (55.1 and 9.3 U/ml) respectively, for a cut-off limit of 37 U/ml. Sensitivity was at 34%, lower than the 49% reported for CA 125 and specificity at 100%. This study concluded that CA 19-9 is a useful marker for severity prediction of the disease.

The case control study by Kurdoglu et al., also assessed serum CA 19-9 and reported significantly higher values among endometriotic women compared to normal, and a positive correlation to disease stage. For a cut-off limit of 37 U/ml, they reported sensitivity of 89% and specificity of 52%, very similar to those of CA 125 (86% and 61% respectively). The third study that assessed CA 19-9 was by Somigliana et al., who measured serum CA 19-9 of case and control women, but reported that there was not a statistical significant difference of the mean values between the two groups (9.8 U/ml and 7.4 U/ml respectively). Sensitivity and specificity was reported at 16% and 91%. This study concludes that use of CA 19-9 for endometriosis detection is not useful, even as a complementary test to CA 125.

Interleukin-6 (IL-6) is another marker that has been studied as a possible endometriosis serum biomarker. Somigliana et al. reported mean serum results in the same groups that were evaluated for CA 125 and CA 19-9. Mean serum values of endometriotic and normal women are similar (0.6 and 1.0 pg/ml) as were among stages I-II and III-IV (0.7 and 0.6 pg/ml). Sensitivity rate reported was at 11% while specificity at 91%. The study therefore concludes that use of serum IL-6 is not useful in endometriosis diagnosis.

Anastasi et al., measured serum Human Epididymis protein 4 (HE4) in the same case control study as for CA 125 in order to assess sensitivity and specificity for endometriosis diagnosis. HE4 does not show a significant statistical difference between normal and endometriotic women as for a cut-off of 150 pmol/L, mean serum results were 48.6 and 53.8 pmol/L respectively. There is no correlation to disease stage as well, while as the authors concluded it may have a complementary role to CA 125 use, as it is significantly increased in ovarian carcinoma cases compared to normal and endometriotic cases. In ovarian carcinoma cases the mean value is 508.3 pmol/L.

The case control study by de Luna Ramos et al., assessed soluble CD-23 besides CA 125 and compared results. By using a cut-off of 10-91 U/ml, mean control (54.47 U/ml) and case (42.85 U/ml) levels had no statistical significant difference. Although there was a slight increase of sCD-23 among stage III-IV endometriotic women, the statistical difference was not important (p value 0.274). The only statistical significant difference observed was among endometriotic cases with cyclic intestinal symptoms (p value 0. 657).

Interpretation of findings

Sensitivity, specificity and positive predictive value for endometriosis diagnosis:

CA 125 appears to be a good biochemical marker for disease diagnosis. This is evident by the statistical significant difference of the mean serum levels of normal and endometriotic women reported by all the studies included. Most of the included studies also reported higher mean serum levels of the case groups than the cut-off limit of 30-35 U/ml. Sensitivity rates vary substantially among studies included but as shown in table III, most of them report rates of more than 50%. Specificity rates are substantially higher than sensitivity rates, as they are approximately 90%, as are positive predictive values. Due to the lack of a homogenous case group (due to included women of all 4 stages of disease), these results need to be addressed with care, as there is a possible difference of sensitivity and specificity of CA 125 among advanced stages of disease.

Correlation of CA 125 to endometriosis staging:

As all studies that have assessed mean CA 125 levels among stage I-II and III-IV women concluded, there is statistically significant difference among them. As shown in table VI, CA 125 correlates well with advanced disease stage, a fact that enables disease staging by using serum CA 125 as a complementary test to laparoscopic visual criteria classification system.

Correlation of CA 125 to presence of endometrioma:

Endometrioma is a disease feature that is associated with advanced stage, so based on the previous findings, it would be expected to be increased than in non-endometriotic cases. The three studies that have measured CA 125 levels in women with endometriomas alone, reported values that are higher than those with no endometrioma and in two of them, higher than the cut-off limit of 35 U/ml. Anastasi et al., reported a mean value of 46.1 U/ml while Kitawaki et al., 50.8 U/ml. The study by Somigliana et al., has reported the lowest mean value which was calculated at 25.9 U/ml.

Sensitivity, specificity and positive predictive value of peritoneal fluid CA 125 for endometriosis diagnosis:

Peritoneal fluid CA 125 levels have high sensitivity, specificity and positive predictive value as well as a strong correlation to disease stage as one of the included studies suggest. This is out of the scope of this study though, as a serum biomarker for disease diagnosis is being investigated.

Correlation of serum CA 125 to other biomarkers in endometriosis diagnosis:

CA 19-9 is the only of the serum biomarkers that have also being assessed in terms of sensitivity and specificity in endometriosis diagnosis that shows comparable rates to that of CA 125. Although it seems that sensitivity is somewhat lower than that of CA 125, specificity is at the same level. As for any other biomarker included in the assessment, none seems to have a positive correlation to endometriosis.

Strengths and weaknesses

Strengths of the studies included are:

(i) similar inclusion and exclusion criteria for both control and case women. Inclusion criteria were reproductive age (18-45 years) and confirmation of disease by laparoscopy for case women. Exclusion criteria included pregnancy, malignancy, benign gynecological conditions such as leiomyomas, pelvic

inflammatory disease, adenomyosis and no hormonal treatment for the past 6 months.

(ii) Confirmation of disease by visual and histological criteria during laparoscopic procedures.

(iii) Staging system for grading disease was the same in all studies (revised ASRM criteria).

(iv) same assay for CA 125 assessment.

(v) same cut-off limit of CA 125 (30-35 U/ml).

Weaknesses of all included studies, are:

(i) significant selection bias as all control and cases were referral to a tertiary center requiring laparoscopy for various indications such as infertility or pelvic pain.

(ii) a non-homogenous case population as the percentage of each stage was not the same in each study, thus affecting CA 125 level.

(iii) lack of a specific cycle time frame in which laparoscopy and blood sampling to be conducted. In cases were blood sampling was conducted during menstruation, CA 125 levels were somewhat increased.

Clinical and research implications / Future directions

Clinical and research implications:

The main clinical and research implications that have been pointed out after reviewing all included studies are:

(i) application of a set of strict exclusion criteria for case and control women such as presence of any benign gynecological condition that affects serum CA 125 levels (leiomyoma, pelvic inflammatory disease, adenomyosis), hormonal medication used in the past 6 months prior to the inclusion in the study, pregnancy and any form of malignancy.

(ii) histological confirmation of endometriosis by laparoscopic guided biopsies of

all women included in the case group, avoiding use of visual criteria alone.

(iii) application of the revised ASRM criteria for disease staging and not any other staging criteria previously used.

(iv) the inclusion of a representative number of all four stages of disease in order to avoid overrepresentation of either minimal-mild or moderate-severe disease, something that would affect mean CA 125 levels.

(v) blood sampling should be performed after menstruation in order to avoid higher serum levels that have been associated with sampling during menstruation.

(vi) consensus on a universal cut-off limit of serum CA 125. Most studies agree on a 30-35 U/ml limit, above which is considered abnormal.

Future directions:

Future directions apply for all research questions analyzed above. Intermsof thefirstquestiononspecificity, sensitivityandpositivepredictivevalueon diseasediagnosis, morecasecontrolstudiesarerequiredonassessingsensitivity of the assay. Mean sensitivity values show great variability among studies and as the meta analysis shows, study design may affect these outcome. More representative population samples are needed in terms of stage I-II and III-IV since if a greater representation of stage III-IV greatly affects the assay sensitivity rate. Mean specificity and positive predictive values are more consistent and correlate well to diseases severity and to each other. In relation to the second research question, more case control studies are also need with a more balanced representation between minimal-mild and moderate-severe disease is needed. Studies that will assess CA 125 correlation to endometrioma alone cases are also needed, as these patients pose some difficulties in terms of classification. If these cases are classified as stage III-IV, correlation to stage I-II is also needed. The fourth research question about peritoneal CA 125 values and

its' sensitivity, specificity and positive predictive values has not been studied extensively. In cases confirmed via laparoscopy, simultaneous peritoneal washings and CA 125 estimation may prove useful in order to assess sensitivity and specificity as well as positive correlation to disease stage. Few studies that did that, published promising results. Other biomarkers that have been assessed include CA 19-9, IL-6, HE4 and CD 23, but statistical significant difference has only been published about CA 19-9. Due to the small number of studies that assessed this, there is a need for more evidence and other possible biomarkers that have not been studied yet.

CONCLUSIONS

Endometriosisisadiseasethataffects primarily womenofreproductiveage withseverecomplications especially interms of infertility and chronic pelvic pain. Prevalence rates are largely unknown since there is a lack of specific diagnostic modality that would enable disease diagnosis. Imaging modalities such as transvaginal ultrasound or MRI only diagnose ovarian endometriosis but not peritoneal implants or adhesions. Laparoscopy is considered to be the gold standard technique for disease diagnosis but as studies suggest, there is overestimation of disease prevalence when applying this as a diagnostic tool. Furthermore, since it is an invasive and expensive modality, it cannot be used as a first line diagnostic or screening tool. The need for a highly sensitive and specific examination remains, which should also be cheap, easy to use and apply, have a standardised assay and well tolerated by patients.

CA 125 is a biomarkers that has been evaluated as a possible diagnostic and screening tool for endometriosis. There are several factors that affect CA 125 levels such as menstruation, pelvic inflammatory disease, leiomyoma, adenomyosis, hormone replacement, pregnancy, liver cirrhosis and malignancy. These factors prevent this biomarker from being considered as disease specific. There is a standardised assay used universally, as well as a specific cut-off limit, above which there is need for further investigation. Evidence suggest that serum CA 125 has positive correlation to superficial as well as deep infiltrating endometriosis. It correlates well to disease stage according to the rASRM classification system, and in the presence of endometriomas.

The systematic literary review and meta-analysis had specific research questions that were investigated in order to assess CA 125 as a diagnostic and screening tool for endometriosis diagnosis. Sensitivity of the assay in disease diagnosis is variable among the included studies ranging between 27-93%. Mean

sensitivity is at 62.83%, for a cut off limit of 35 U/ml, which is not high enough in order to allow use of it for screening purposes. It correlates well with cut-off limits suggested by various studies, but not as well to study size. It therefore has a role in diagnosis but as evidence suggest, it requires further case control studies with modifications of inclusion criteria in order to assess sensitivity rates. Data on specificity rates are more consistent among studies included with most of them reporting rates more than 80% and a mean rate of 85.26%. Specificity also shows positive correlation to cut-off limits reported by different studies. As a result of that, CA 125 is specific enough to be used as a diagnostic test for endometriosis. Positive predictive value shows positive correlation to specificity with a mean value of 90.14% for an AUC 0.759-0.935. Correlation of serum CA 125 levels to disease staging is the second research question investigated and as data extracted suggest, there is positive correlation to stage. Mean levels of stage I-II is at 25.05 U/ml and stage III-IV at 63.98 U/ml. These data suggest that the assay is a sensitive and specific modality for diagnosis. Other extracted data from the meta-analysis show that mean control levels are at 14.05 U/ml and mean case levels at 43.19 U/ml which suggests a statistical significant difference among the two groups. In the subgroup of endometriosis patients that only have an ovarian endometrioma as a feature of disease and cannot be classified according to the rASRM system, results also show positive correlation to serum CA 125 levels, and a negating correlation to endometrioma negative patients.

Peritoneal fluid CA 125 levels although not studied as extensively as serum levels, also demonstrate a strong positive correlation to disease presence and severity. Mean control group levels are at 888.7 U/ml while mean case group levels at 1469.4 U/ml. Disadvantages of this examination though are the need for laparoscopy, an expensive and invasive procedure, which prohibits its' use as a

screening tool for the general population. Application during laparoscopy may be of use in order to confirm diagnosis.

The search for other possible biomarkers to be used in disease diagnosis has expanded beyond CA 125 to include CA 19-9, IL 6, HE4 and CD23. The only other biomarker to demonstrate a statistical significant difference between controls and cases is CA 19-9. When compared to CA 125 sensitivity and specificity rates, there is lower sensitivity but higher specificity rates than CA 125, thus not allowing it to be considered as a better biomarker for CA 125 diagnosis.

Overall, CA 125 is a relatively sensitive and highly specific biomarker for detecting endometriosis. Problems that exist and do not allow its' use as a single diagnostic modality for diagnosis or screening, include factors affecting it such as menstruation, leiomyoma, malignancy etc, as well as a statistical significant difference among stage I-II and III-IV patients. Further studies are needed in order to establish the degree of correlation between CA 125 and disease stage, due to study design issues that still exist. Such issues are mainly small study size, an overrepresentation of stage III-IV cases, and the lack of implementation of a full listof exclusion criteria such as menstruation, leiomyomas, pelvic inflammatory disease and adenomyosis. Further studies are also needed in terms of comparing CA 19-9 and CA 125 sensitivity and specificity rates. The need for a highly sensitive and specific serum biomarkers that is cheap, easy to use, standardised and well tolerated by patients still remains a necessity for a disease that affects a significant portion of the population with severe consequences.

REFERENCES

Abbott JA, Hawe J, Clayton RD, Garry R. The effects and effectiveness of laparoscopic excision of endometriosis: a prospective study with 2-5 year follow up. *Hum Reprod* 2003;**18**(9):1922-1927.

Abrao MS, Podgaec S, Filho BM, Ramos LO, Pinotti JA, de Oliveira RM. The use of biochemical markers in the diagnosis of pelvic endometriosis. *Hum Reprod* 1997;**12**(11):2523-2527.

Abrao MS, da C Goncalves MO, Dias JA, Podgaec S, Chamic LP, Blabalg R. Comparison between clinical examination, transvaginal sonography and magnetic resonance imaging for the diagnosis of deep endometriosis. *Hum Reprod*2007;**22**(12):3092-3097.

Acien P, Velascol I, Guttierez M, Martinez-Beltran M. Aromatase expression in endometriotic tissues and its relationship to clinical and analytical findings. *FertilSteril* 2007;**88**(1):32-38.

Amaral VF, Ferriani RA, Silva de Sa MF, Nogueira AA, Silva JCR, Silva ACJR, de Moura MD. Positive correlation between serum and peritoneal fluid CA-125 levels in women with pelvic endometriosis. *Sao Paulo Med J* 2006;**124**(4):223-7.

American College of Obstetricians and Gynecologists. Washington DC. ACOG 1993. *Endometriosis*. ACOG technical bulletin no 184.

Amaral VF, Ferriani RA, Sa MF, et al. Positive correlation between serum and peritoneal fluid CA-125 levels in women with pelvic endometriosis. *Sao Paulo*

Med J 2006;**124**(4):223-7.

American Society for Reproductive Medicine (ASRM). Revised ASRM classification of endometriosis. *FertilSteril* 1996;**67**:817-821.

Anastasi E, Granato T, Falzarano R, Storelli P, Ticino A, Frati L, Benedetti Panici P, GraziaPorpora M. The use of HE4, CA 125 and CA 72-4 biomarkers for differential diagnosis between ovarian endometrioma and epithelial ovarian cancer. *J Ovarian Res* 2013;**6**:44.

Ang WC, Alvey CM, Marran S, Kennedy SH, Golding S. A systematic review of the accuracy of magnetic resonance imaging (MRI) in the diagnosis of endometriosis. (Submitted for publication).

Armenti VT, Ahlswede KM, Ahlswede BA, Jarrell BE, Moritz MJ, Burke JF. National Transplantation Pregnancy Registry in cyclosporine-treated female kidney transplants. *Transplantation* 1994;**57**:502-506.

Barbieri RL, Niloff JM, Bast RC Jr, Scaetzl E, Kistner RW, Knapp RC. Elevated serum concentrations of CA-125 in patients with advanced endometriosis. *Fertil Steril* 1986;**45**(5):630-634.

Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *ClinPharmacolTher* 2001;**69**(3):89-95.

Brosens IA. Endometriosis- a disease because it is characterized by bleeding. *Am*

J ObstetGynecol 1997;**176**:263-267.

Chapron C, Fauconnier A, Dubuisson JB, Barakat H, Vieira M, Breart G. Deep infiltrating endometriosis: relation between severity of dysmenorrhea and extent of disease. *Hum Reprod* 2003b;**18**:760-766.

Chatzikokkinou P, Thorfinn J, Angelidis KI, Papa G, Trevisan G. Spontaneous endometriosis in an umbilical skin lesion. *ActaDermatoven APA* 2008;**18**:126-30.

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. *ActaObstGyneSc* 1998;**77**(6):665-665.

Cohen J, Cohen P, West SG, Aiken LS. *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Psychology Press 2002.

Cortes C, Mohri M, Riley M, Rostamizadeh A. Sample selection bias correction theory. *Algorithmic Learn Theory* 2008;**5254**:38-53.

Cramer DW, Wilson E, Stillman RJ, Berger MJ, Belisle S, Schiff I, Albrecht B, Gibson M, Stadel BV, Schoembaum SC. The relation of endometriosis to menstrual characteristics, smoking and exercise. *JAMA* 1986;**355**:1904-1908.

Dawood MY, Khan-Dawood FS, Ramos J. Plasma and peritoneal fluid levels of CA 125 in women with endometriosis. *Am J ObstetGynecol* 1988;**159**(6):1526-1531.

Debrock S, Drijkoningen R, Goossens W, Meuleman C, Hill JA, D'Hooghe TM.

Quantity and quality of retrograde menstruation: red blood cells, inflammation and peritoneal cells. Abstract presented at the annual meeting of the American Society for Reproductive Medicine, Toronto, Canada. *FertilSteril* 2000;**74**:3S.

de Luna Ramos IM, Podgaec S, Abrao SM, de Oliveira R, ChadaBaracat E. Evaluation of CA-125 and soluble CD 23 in patients with pelvic endometriosis: a case control study. *Rev Assoc Med Bras* 2012;**58**(1):26-32.

D'Hooghe TM, Bambra CS, Raeymaekers BM, Koninckx PR. Increased prevalence and recurrence of retrograde menstruation in baboons with spontaneous endometriosis. *Hum Reprod* 1996;**11**:2022-2025.

D'Hooghe TM, Bambra CS, Hill JA. Avery sticky riddle. *FertilSteril* 1997;**67**:182-183.

D'Hooghe TM, Bambra CS, Ling Xiao L, Hill JA. The effect of menstruation and intrapelvic injection of endometrium on peritoneal fluid parameters in the baboon. *Am J ObstetGynecol* 2001;**184**:917-925.

dos Reis RM, de Sa MF, de Moura MD, Nogueira AA, Ribeiro JU, Ramos ES, Ferriani RA. Familial risk among patients with endometriosis. *J Assist Reprod Genet* 1999;**16**:500-503.

Fawcelt T. An introduction to ROC Analysis. *Pattern Recognition Letters* 2006;**27**(8):861-874.

Fletcher RH, Fletcher SW. *Clinical epidemiology: the essentials* (4thed). Baltimore

Md: Lippincott Williams & Wilkins 2005;45.

Franssen AM, van der Heijden PF, Thomas CM, Doesburg WH, Willemsen WN, Rolland R. On the origin and significance of serum CA-125 concentrations in 97 patients with endometriosis before, during and after buserelin acetate, nafarelin or danazol. *FertiSteril* 1992;**57**(5):974-979.

Fraser IS, McCarron G, Markham R. Serum CA-125 levels in women with endometriosis. *Aust N Z J ObstetGynaecol* 1989;**29**(4):416-420.

Fujii S. Secondary mullerian system and endometriosis. *Am J ObstetGynecol* 1991;**165**:219-225.

Giudice LC. Clinical practice: endometriosis. *N Engl J Med* 2010;**362**:2389-2398.

Giudice LC, Kao L. Endometriosis. *Lancet* 2004;**364**:1789-1799.

GruppoItaliano per lo Studio dell' Endometriosi. Relationship between stage, site and morphological caracteristics of pelvic endometriosis and pain. *Hum Reprod* 2001;**16**(12):2668-2671.

Halme J, Becker S, Hammond MG, Raj SG, Talbert LM. Retrograde menstruation in healthy women and in patients with endometriosis. *ObstetGynecol* 1984;**64**:151-154.

Harada T, Kubota T, Aso T. Usefulness of CA 19-9 versus CA 125 for the diagnosis of endometriosis. *FertilSteril* 2002;**78**(4):733-739.

Healy DL, Rogers PAW, Hii L, Wingfield M. Angiogenesis: a new theory for endometriosis. *Hum Reprod* 1998;**4**(5):736-740.

Hermsen BBJ, von Mensdorff-Pouilly S, Berkhof J, van Diest PJ, Gille JJP, Menko FH, Blankenstein MA, Kenemans P, Verheijen RHM. Serum CA 125 in relation to adnexal dysplasia and cancer in women at hereditary high risk of ovarian cancer. *J ClinOncol* 2007;**25**:1383-1389.

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

Janssen EB, Rijkers ACM, Hoppenbrouwers K, Meuleman C, D'Hooghe TM. Prevalence of endometriosis diagnosed by laparoscopy in adolescents with dysmenorrhea or chronic pelvic pain: a systematic review. *Hum Reprod* 2013;**19**(5):570-582.

Jirasek JE, Henzl MR, Uher J. Periovarian, peritoneal adhesions in women with endometriosis. Structural patterns. *J Reprod Med* 1998;**43**:276-280.

Kafali H, Artuc H, Demir N. Use of CA 125 fluctuation during the menstrual cycle as a tool in the clinical diagnosis of endometriosis; a preliminary report. *Eur J ObstetGynecolReprodBiol* 2004;**116**:85-88.

Kennedy S, Bergqvist A, Chapron C, D'Hooghe T, Dunselman G, Greb R, Hummelshoj L, Prentice A, Saridogan E. On behalf of the ESHRE special interest group for endometriosis endometrium guideline development group. ESHRE

guideline for the diagnosis and treatment of endometriosis. *Hum Reprod* 2005;**20**:2698-2704.

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

Koninckx PR, Heyns W, Verhoeven G, et al. Biochemical characterization of peritoneal fluid in women during the menstrual cycle. *J ClinEndocrinolMetab* 1980c;**51**:1239-1244.

Koninckx PR, Martin D. Deep endometriosis: a consequence of infiltration or retraction or possiblyadenomyosisexterna? *FertilSteril* 1992;**58**:924-928.

Koninckx PR. Is mild endometriosis a condition occurring intermittently in all women? *Hum Reprod* 1994;**9**(12):2202-2205.

Koper NP, Thomas CMG, MassugerLFAG, van der Mooren MJ, Kiemeney LALM, Verbeek ALM. Serum CA 125 concentrations in women of different ages, hormonal statuses orclinical conditions. *Int J Gynecol Cancer* 1997;**7**:405-411.

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. *FertilSteril* 2009;**92**:1761-3.

Liu DT, Hitchcock A. Endometriosis: its association with retrograde menstruation, dysmenorrhea and tubal pathology. *Br J ObstetGynecol*

1986;**93**(8):859-62.

McLaren J, Prentice A, Charnock Jones DS, et al. Vascular endothelial growth factor (VEGF) concentrations are elevated in peritoneal fluid of women with endometriosis. *Hum Reprod* 1996;**11**:220-223.

Martinez S, Garrido N, Coperias JL, Pardo F, Desco J, Garcia-Velasco JA, Simon C, Pellicer A. Serum interleukin-6 levels are elevated in women with minimal-mild endometriosis. *Hum Reprod* 2007;**22**(3):836-842.

Massad LS. Assessing new technologies for cervical cancer screening: beyond sensitivity. *J Low Genit Tract Dis* 2008;**12**:317-326.

Matsuzaki S, Canis M, Pouly JL, Wattiez A, Okamura K, Mage G. Cyclooxygenase-2 expression in deep endometriosis and matched ectopic endometrium. *Fertil Steril* 2004;**82**(5):1309-1315.

May KE, Conduit-Hulbert SA, Villar J, Kirtley S, Kennedy SH, Becker CM. Peripheral biomarkers of endometriosis: a systematic review. *Hum Reprod Update* 2010;**16**(6):651-674.

Meigs JV. An interest in Endometriosis and its consequences. *Am J ObstetGynaec* 1960;**79**:625-35.

Melin A, Sparen P, Persson I, Bergqvist A. Endometriosis and the risk of cancer with special emphasis on ovarian cancer. *Hum Reprod* 2006;**21**(5):1237-1242.

Mihalyi A, Gevaert O, Kyama CM, et al. Non-invasive diagnosis of endometriosis based on a combined analysis of six plasma biomarkers. *Hum Reprod* 2010;**25**(3):654-664.

Mol BW, Bayram N, Lijmer JG, Wiegerinck MA, Bongers MY, van der VF, Bossuyt PM. The performance of CA-125 measurement in the detection of endometriosis: a meta analysis. *FertilSteril* 1998;**70**:1101-1108.

Moore J, Copley S, Morris J, Lindsell D, Golding S, Kennedy S. A systematic review of the accuracy of ultrasound in the diagnosis of endometriosis. *Ultrasound ObstetGynecol* 2002;**20**:630-634.

Moore RG, Miller MC, Steinhoff MM, Skates SJ, Lu KH, Lambert-Messerlian G, Bast RC Jr. Serum HE4 levels are less frequently elevated than CA 125 in women with benign gynecologic disorders. *Am J ObstetGynecol* 2012;**206**(4):351.

Niloff JM, Knapp RC, Schaetzl E, Reynolds C, Bast RC Jr. CA 125 antigen levels in obstetric and gynecologic patients. *ObstetGynecol* 1984;**64**(5):703-707.

O'Brien TJ, Raymond LM, Bannon GA, et al. New monoclonal antibodies identify the glucoprotein carrying the CA 125 epitope. *Am J ObstetGynecol* 1991;**165**:1857-1864.

O'Brien TJ, Beard JB, Underwood LJ, Dennis RA, Santin AD, York L. The CA 125 gene: an extracellular superstructure dominated by repeat sequences. *Tumour Biol* 2001;**22**:348-366.

Oosterlynck DJ, Cornillie FJ, Waer M, et al. Women with endometriosis show a defect in natural killer activity resulting in a decreased cytotoxicity toautologous endometrium. *FertilSteril* 1991;**56**:45-51.

Pauler DK, Menon U, McIntosh M, Symecko HL, Skates SJ, Jacobs IJ. Factors influencing serum CA 125II levels in healthy postmenopausal women. *Cancer Epidemiol Biomarkers Prev* 2001;**10**:489-493.

Prayudhana S, Yudho J, Hadisaputra W. Moderate-severe endometriosis shows higher CA 125 serum level compared to minimal-mild endometriosis. *Indones J ObstetGynecol* 2012;**36**(3):121-124.

Peters TJ, Wildschut HIJ, Weiner CP. Epidemiological considerations in screening. In:Wildschut HIJ, Weiner CP, Peters TJ (eds). *When to screen in Obstetrics and Gynecology*. London. WB Saunders Company Ltd. 1996. 1-12.

Pittaway DE, Fayez JA. Serum CA-125 antigen levels increase during menses. *Am J ObstetGynecol* 1987;**156**:75-76.

Ruiz-Flores FJ, Garcia-Valesco JA. Is there a benefit for surgery in endometrioma-associated infertility ?*CurrOpinObstetGynecol* 2012;**24**(3):136-140.

Sampson JA. Peritoneal endometriosis due to menstrual dissemination of endometriotic tissue into the pelvic cavity. *Am J ObstetGynecol* 1927;**14**:422-469.

Sayasneh A, Tsivos D, Crawford R. Endometriosis and ovarian cancer: A

systematic review. *International Scholarly Research Notices* 2011.

Scott RB, TeLinde RW, Wharton LR Jr. Further studies on experimental endometriosis. *Am J ObstetGynecol* 1953;**66**:1082.

Somigliana E, Vigano P, Gaffuri B, et al. Human endometrial stromal cells as a source of soluble intercellular adhesion molecule (ICAM)-1 molecules. *Hum Reprod* 1996;**11**:1190-1194.

Somigliana E, Vigano P, Tirelli AS, Felicetta I, Torresani E, Vignali M, Di Blasio AM. 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. *Hum Reprod* 2004;**19**(8):1871-1876.

Somigliana E, Vercellini P, Vigano P, Benaglia L, Crosignani PG, Fedele L. Non-invasive diagnosis of endometriosis: the goal or own goal? *Hum Reprod* 2010;**25**(8):1863-1868.

Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, Hing C, Kwok CS, Pang C, Harvey I. Dissemination and publication of research findings: An updated review of related biases. *Health Tech Ass* 2010;**14**(8):193.

Suginami H. A reappraisal of the coelomic metaplasia theory by reviewing endometriosis occurring in unusual sites and instances. *Am J ObstetGynecol* 1991;**165**(1):214-218.

Szumilas M. Explaining Odds Ratio. *J Can Acad Child Adolesc Psychiatry* 2010;**19**(3):227-229.

Tsoumpou I, Kyrgiou M, Gelbaya TA, Nardo LG. The effect of surgical treatment for endometrioma on in vitro fertilization outcomes: a systematic review and meta-analysis. *FertilSteril* 2009;**92**(1):75-87.

van der Linden PJ, Dunselman GA, de Goeij AF, et al. Epithelial cells in peritoneal fluid- of endometrial origin? *Am J ObstetGynecol* 1995;**173**:566-570.

Vercellini P. Endometriosis: what a pain it is. *SeminReprodEndocrinol* 1997;**15**:251-261.

Vercellini P, Fedele L, Aimi G, Pietropaolo G, Consonni D, Crosignani PG. Association between endometriosis stage, lesion type, patient characteristics and severity of pelvic pain symptoms: a multivariate analysis of over 1000 patients. *Hum Reprod* 2007;**22**(1):266-271.

Vodolazkaia A, El-Aalamat Y, Popovic D, Mihalyi A, et al. Evaluation of a panel of 28 biomarkers for the non-invasive diagnosis of endometriosis. *Hum Reprod* 2012;**27**(9):2698-2711.

Wald NJ. The definition of screening. *J Med Screen* 2001;**8**:1

Wellbery C. Diagnosis and treatment of endometriosis. *Ann Fam Physician* 1999;**60**:1753-1768.

Wells GA, Shea B, O'Connell D, Peterson JEA, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. (available on line at:http://www.medicine.mcgill.ca)

Zomer MT, Ribeiro R, Trippia CH, Cavalcanti TC, Hayashi RM, Kondo W. Correlation between serum CA 125 levels and surgical findings in women with symptoms evocative of endometriosis. *Rev Bras GinecolObstet* 2013;**35**(6):262-267.