RESEARCH ARTICLE

Diagnostic Value of Human Epididymis Protein 4 Compared with Mesothelin for Ovarian Cancer: a Systematic Review and Meta-analysis

Jia-Ying Lin[1&], Jin-Bao Qin[2&], Xiao-Yan Li[3], Ping Dong[1], Bing-De Yin[4*]

Abstract

Background and Purpose: Ovarian cancer is the leading cause of death among gynecologic cancers because of the lack of effective early detection methods. Accuracies of the human epididymis protein 4 (HE4) and mesothelin in detecting ovarian cancer have never been systematically assessed. The current systematic review aimed to tackle this issue. **Methods:** MEDLINE, EMBASE, and Cochrane databases were searched (September 1995–November 2011) for studies on the diagnostic performances of HE4 and mesothelin in differentiating ovarian cancer from other benign gynecologic diseases. QUADAS items were used to evaluate the qualities of the studies. Meta-DiSc software was used to handle data from the included studies and to examine heterogeneity. All included studies for diagnostic performance were combined with sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratios (DORs) with 95% confidence intervals (CIs), summary receiver operating characteristic (SROC) curves, and areas under the SROC curves (AUC). **Results:** A total of 18 studies and 3,865 patients were eligible for the final analysis. The pooled sensitivity estimates for HE4 (74.4%) were significantly higher than those for mesothelin (49.3%). The pooled specificity estimates for mesothelin (94.5%) were higher than those for HE4 (85.8%). The pooled DOR estimates for HE4 (26.22) were higher than those for mesothelin (24.01). The SROC curve for HE4 showed better diagnostic accuracy than that for mesothelin. The PLR and NLR of HE4 were 6.33 (95% CI: 3.58 to 11.18) and 0.27 (95% CI: 0.21 to 0.34), respectively. The PLR and NLR for mesothelin were 11.0 (95% CI: 6.21 to 19.59) and 0.51 (95% CI: 0.42 to 0.62), respectively. The combination of the two tumor markers or their combination with CA-125 increased sensitivity and specificity to different extents. **Conclusion:** The diagnostic accuracy of HE4 in differentiating ovarian cancer from other benign gynecologic diseases is better than that of soluble mesothelin-related protein. Combinations of two or more tumor markers show more sensitivity and specificity.

Keywords: Ovarian carcinoma - HE4 - mesothelin - meta-analysis

Introduction

Epithelial ovarian cancer (EOC), along with primary peritoneal cancer, is the fourth leading cause of cancer death among women, with approximately 15,000 deaths annually in the United States. The poor prognosis of ovarian cancer is primarily due to the fact that the cancer is not detected in a majority of patients until it has spread beyond the ovary (Williams et al., 2007). The detection of a larger fraction of ovarian cancers during the early stages might significantly improve the overall survival rate. Over the last two decades, attempts to develop effective screening methods have focused on ultrasonography and serum markers. Given the low prevalence of ovarian cancer in the general population, an effective strategy must have a high sensitivity (> 75%) during the early stages of

the disease and a very high specificity (> 99.6%) to attain a positive predictive value (PPV) of 10%. A PPV of 10% equates to a situation in which only 1 out of 10 surgical interventions leads to the diagnosis of ovarian cancer (Bast, 2004).

To date, the glycoprotein CA-125 is the most thoroughly investigated biomarker for ovarian cancer screening. CA-125 has been used as a tool to differentiate between benign and malignant ovarian masses. CA-125 is not sufficiently sensitive for screening the general population when used as an individual marker. In most studies, CA-125 is elevated in approximately 50% to 60% of stage I diseases at the time of conventional diagnosis (Jacobs & Bast, 1989). In addition, the specificity of CA-125 is compromised because many benign gynecologic and medical conditions and other malignancies can result

*[1]Department of Otolaryngology Head and Neck Surgery, [4]Department of Urology, Shanghai First People's Hospital, School of Medicine, Shanghai Jiao Tong University, [2]Department of Vascular Surgery, the Ninth People's Hospital, the Affiliated to Shanghai Jiaotong University, School of Medicine, [3]Department of Otolaryngology-Head and Neck Surgery, Shanghai Children's Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China [&]Equal contributors *For correspondence: 13818761532@126.com*

in elevated serum CA-125 levels. Therefore, new serum markers need to be sought to replace or complement CA-125 to detect a greater fraction of ovarian cancers during the early stages of the disease.

Mesothelin or soluble mesothelin-related protein (SMRP) and HE4 are two of the most promising novel ovarian cancer biomarkers that are currently under evaluation. Mesothelin and HE4 are commonly overexpressed in ovarian cancer tissues and are elevated in the sera of patients with ovarian cancer. SMRP and HE4 have many beneficial characteristics that set them apart from other potential markers currently being studied. The preeminent distinctiveness of SMRP and HE4 can be attributed to their temporal stability (Salceda et al., 2005; Moore et al., 2008), which may assist in the early diagnosis of high-risk patients.

Previous studies on the roles of SMRP and HE4 markers in the differential diagnosis of ovarian cancer from benign gynecologic diseases present conflicting results. Therefore, a comprehensive system review would be useful to synthesize the available information. This study aims to assess the diagnostic performances of SMRP and HE4 in differentiating between ovarian cancer and benign gynecologic diseases using standard meta-analysis techniques.

Materials and Methods

Literature search

A systematic review of original articles analyzing the diagnostic performance of HE4 and SMRP was performed by searching MEDLINE, EMBASE, and Cochrane Databases. Original and review articles published from September 1995 to November 2011 were sought; The search terms were "HE4/ WFDC2", " mesothelin/SMRP/ MLSN", "ovarian carcinoma/ovarian cancer/carcinoma of ovary,"sensitivity/specificity/false negative/false positive/diagnosis/detection/accuracy". We evaluated all associated publications to retrieve the most eligible studies. Moreover, their reference lists were searched manually to find other relevant publications. Both original and review articles were sought because the latter were considered as additional sources of unaccounted original works.

Selection of studies

The eligibility criteria for the meta-analysis of the studies included the following:(1) both the sensitivity and specificity of levels of HE4 and SMRP for the diagnosis of ovarian cancer were provided , or HE4 and SMRP values were provided in a scatter plot form, allowing test results to be extracted for each individual. (2) 50 or more patients were included. (3)The study design included women with ovarian cancer and benign gynecologic diseases ,and evaluated the contribution of HE4 and SMRP.

Articles were excluded when data were insufficient to construct a 2×2 table of the test result (serum HE4 and SMRP concentration). The 2×2 tables were constructed independently by two of the authors (JY, L and JB, Q). In the event of disagreement, the judgment of a third author (BD, Y) was decisive.

Data extraction

The final set of English articles was assessed independently by two observers (JY, L and JB, Q). The observers were blinded to publication details, and differences between them were resolved by a consensus. Data retrieved from the reports included the name of the author, publication year, participant characteristics, test method, cut-off value, sensitivity and specificity, and study quality score.

We assessed the quality of the included studies by the criteria selected from the Quality Assessment for Studies of Diagnostic Accuracy checklist for the assessment of diagnostic studies (Whiting et al., 2003): study design (prospective or retrospective); patient selection (consecutive or not); blinding (blind or not to the interpretation of index text); assay method ;study size, etc .The numbers of true-positive (TP), false-negative (FN), false-positive (FP) and true-negative (TN) results in the detection of ovarian cancer were extracted on a per-patient or per-lesion basis.

Statistical analysis

The sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) for each study were calculated, and the results were pooled as per the DerSimonian Liard random effects model (DerSimonian & Laird, 1986). Heterogeneity was assessed using the likelihood ratio I^2 index and X^2 test. I^2 index is a measure of the percentage of total variation across studies due to heterogeneity beyond chance. Values over 50% indicate heterogeneity (Huedo-Medina et al., 2006). In the likelihood ratio X^2 test, $p < 0.05$ was considered to indicate apparent heterogeneity. If heterogeneity existed (Dinnes et al., 2005) a random effects model was used in the primary meta-analysis to obtain a summary estimate for sensitivity with 95% confidence interval (CI). The diagnostic performance of the test under study per unit increase in the covariate was used as the accuracy measure. The current work used Moses' linear model to draw a summary receiver operating characteristic (SROC) curve, which summarized the joint distribution of sensitivity and specificity. The area under the SROC curve (AUC) was calculated. Publication biases were assessed using funnel plots.
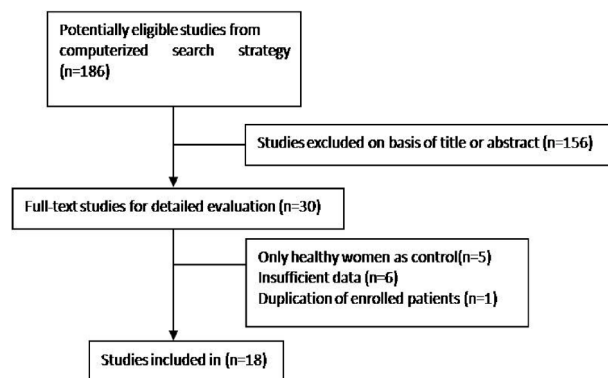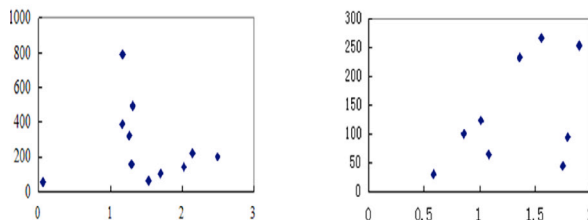
Results

Literature search and study design characteristics

The current research yielded 186 primary studies, of which 156 were excluded after reviewing the title and abstract, and 12 after reviewing the full article (Figure 1). Five articles were excluded because only healthy women were included in the control group. Six articles were excluded because their data were insufficient to calculate the true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) values.

One article was excluded because its author published two reports on the same patients and considered only the study with the best quality. A total of 18 studies (Scholler et al., 1999; Hellstrom et al., 2003; McIntosh et al., 2004; Hassan et al., 2006; Badgwell et al., 2007; Moore et al.,

Diagnostic Value of Human Epididymis Protein 4 Compared to Mesothelin for Ovarian Cancer

Table 1. Main Characteristics of 18 Studies

Auother	Location	temperature(°C)	Number of patients'serum	design	Blind	enrollment
Scholler	USA	4	124	ND	ND	ND
McIntosh	USA	ND	95	ND	ND	ND
Hassan	USA	-80	45	ND	Yes	Consecutive
Badgwell	USA	-80	254	ND	ND	ND
Scholler	USA	-80	124	ND	Yes	ND
Palmer	USA,Sweden	ND	101	Prospective	Yes	Consecutive
Moore	USA	-80	233	Prospective	Yes	Consecutive
Shah	USA	-80	267	ND	ND	ND
Abdel-Azeez	Egypt	-80	65	ND	ND	ND
Hellstrom	USA	ND	66	Prospective	Yes	Consecutive
Huhtinen	Finland	-20	148	Retrospective	ND	Consecutive
Nolen	USA	ND	790	ND	ND	ND
Gorp	Belgium	-80	65	Prospective	ND	Consecutive
Jacob	Switzerland	-80	160	Prospective	ND	ND
Holcomb	USA	ND	229	Prospective	Yes	Consecutive
Chang	China	-70	202	ND	Yes	Consecutive
Montagnana	Italy	-80	104	Retrospective	ND	Consecutive
Park	Korea	-70	323	Retrospective	ND	Consecutive

Figure 1. Results of Search Strategy

Potentially eligible studies from computerized search strategy (n=186)

Studies excluded on basis of title or abstract (n=156)

Full-text studies for detailed evaluation (n=30)

Only healthy women as control(n=5)
Insufficient data (n=6)
Duplication of enrolled patients (n=1)

Studies included in (n=18)

Figure 2. Funnel Plot of Included Studies. The plot of this meta-analysis showed symmetric, demonstrating that there was probably no publication bias

2008; Palmer et al., 2008; Scholler et al., 2008; Huhtinen et al., 2009; Montagnana et al., 2009; Shah et al., 2009; Abdel-Azeez et al., 2010; Nolen et al., 2010; Chang et al., 2011; Holcomb et al., 2011; Jacob et al., 2011; Van Gorp et al., 2011; Park et al., 2012) with 3865 patients who satisfied the inclusion criteria were analyzed. The results are shown in Table 1. A total of 11 articles studied the diagnostic value of serum HE4, and 9 studied the diagnostic value of serum SMRP. Two articles evaluated both HE4 and SMRP. The assay methods, cut-off point, TP, FP, TN, and FN of HE4 or SMRP from each article are listed in Tables 2 and 3.

Study description and quality
As shown in Table 1, the average sample size of the HE4 studies was 217 (range: 65 to 323), whereas that of the SMRP studies was 145 (range: 45 to 267). The studies were published between 1999 and 2011. Ten studies

Table 2. Test for Heterogeneity and Threshold Effect in the Meta-analysis

	Likelihood ratio		I^2 idex(%)
	χ^2	p	
sensitivity			
HE4	43.37	0	76.9
SMRP	51.59	0	84.5
CA125+HE4	10.29	0.173	32
CA125+SMRP	26.41	0	92.4
SMRP+HE4	6.02	0.049	66.8
CA125+SMRP+HE4	7.58	0.023	73.6
specificity			
HE4	196.21	0	94.9
SMRP	14.04	0.081	43
CA125+HE4	177.89	0	96.1
CA125+SMRP	12.13	0.0023	83.5
SMRP+HE4	7.69	0.021	74
CA125+SMRP+HE4	0.95	0.283	20.8

(55.6%) collected samples from consecutive patients. Seven studies (38.9%) reported blinded interpretation of tumor marker assays. Six studies (33.4%) had prospective study designs. Three studies (16.7%) were retrospective. All studies used histopathologic analysis as their reference standard.

Diagnostic accuracies of HE4 and SMRP
Publication bias and heterogeneity To assess possible publication biases, scatter plots were designed using the log DORs of individual studies against their sample size. The plots of these meta-analyses (Figure 2) were symmetric, demonstrating no publication bias. Heterogeneity was observed for most markers, which was confirmed by either the likelihood ratio χ^2 test or the I^2 index (Table 2). The random-effects model was used to estimate the overall effect.

Diagnostic accuracy
The SROC curve graphs for the determination of tumor markers that show TP rates versus FP rates from individual studies are shown in Figure 2. Pooled results of the diagnostic accuracy of each tumor marker and marker

Table 3. Pooled Results of Diagnostic Accuracy of Serum Tumor Marker Combinations

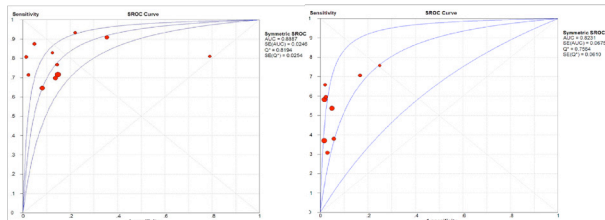| combinations	Nmubers of studies	Sensitivity (95% CI)	Specificity (95% CI)	PLR(95% CI)	NLR(95% CI)	DOR (95% CI)	AUC (SEM)
HE4	11	0.74(0.72-0.77)	0.86(0.84-0.87)	6.33(3.58-11.18)	0.27(0.21-0.34)	26.22(14.83-46.38)	0.8887
SMRP	9	0.49(0.45-0.53)	0.95(0.93-0.96)	11.02(6.21-19.59)	0.51(0.42-0.62)	24.01(14.46-39.85)	0.8231
CA125	11	0.76(0.74-0.86)	0.81(0.73-0.89)	4.17(2.75-6.37)	0.27(0.25-0.30)	25.19(14.49-41.97)	0.8689
CA125+HE4	8	0.80(0.77-0.83)	0.86(0.83-0.88)	11.42(3.80-34.35)	0.23(0.20-0.28)	44.88(25.98-77.54)	0.9096
CA125+SMRP	3	0.59(0.51-0.67)	0.96(0.93-0.98)	12.03(2.93-49.42)	0.42(0.26-0.68)	43.40(18.52-101.66)	0.9251
SMRP+HE4	3	0.78(0.70-0.84)	0.94(0.90-0.97)	10.07(3.56-32.16)	0.25(0.18-0.36)	54.86(26.87-112.01)	0.9297
CA125+SMRP+HE4	3	0.67(0.59-0.74)	0.95(0.91-0.97)	10.82(5.92-19.80)	0.32(0.21-0.51)	35.83(13.21-97.22)	0.9032

Figure 3. Summary ROC Curves of HE4 and SMRP

combinations are shown in Table 1. The AUC of HE4 and SMRP were 0.8887 and 0.8231, respectively. A significant difference was found between the two markers (p < 0.05). The SROC curves reflected very good but not excellent overall diagnostic performance for the HE4 assay (Figure 2) and bad performance for the SMRP assay (Figure 3).

Some publications included in the meta-analysis evaluated the simultaneous determination of two or more serum tumor markers in ovarian cancer diagnosis. The pooled results of the diagnostic accuracy of the combinations of CA-125, HE4, and SMRP are shown in Table 3. The results indicated that some combinations of tumor markers had a greater diagnostic role than one tumor marker alone. CA-125 + HE4 achieved the highest sensitivity in ovarian cancer diagnosis at 0.801 (range: 0.766 to 0.833). The specificity, PLR, NLR, and DOR of CA-125 + HE4 were 0.855 (range: 0.830 to 0.877), 11.42 (range: 3.80 to 34.35), 0.23 (range: 0.20 to 0.28), and 44.88 (range: 25.98 to 77.54), respectively. CA125+SMRP showed the highest specificity at 0.961 (range: 0.928 to 0.982). The sensitivity, PLR, NLR, and DOR of CA125+SMRP were 0.588 (range: 0.507 to 0.665), 12.03 (range: 2.93 to 49.42), 0.42 (range: 0.26 to 0.68), and 43.40 (18.52 to 101.66), respectively.

Discussion

EOC is the leading cause of death among gynecologic malignancies in the U.S. because of the lack of sensitive screening methods. Over the past 20 years, many studies have focused on the use of the tumor marker CA-125 and ultrasound imaging in ovarian cancer screening. Recently, the use of novel and multiple biomarkers has been proven to be an accurate method for preoperatively assessing the risk of ovarian cancer in women with ovarian masses. To the best of our knowledge, the present study is the first meta-analysis to estimate the pooled diagnostic accuracy characteristics of serum HE4 and SMRP, both individually and in various combinations with CA-125, in ovarian cancer.

Mesothelin is a 40 kDa cell surface glycoprotein that is highly expressed in pancreatic cancers, ovarian cancers, mesotheliomas, and other cancers (Chang & Pastan, 1996). The measurement of mesothelin in the blood may be useful for diagnosing and monitoring the response of patients to therapy because small amounts of mesothelin can be detected in the blood of patients with mesothelin-positive cancers (Scholler et al., 1999). Serum SMRP assays exhibited a limited diagnostic sensitivity of 0.493 (range: 0.454 to 0.532) because they failed to identify approximately half of the patients but demonstrated a high specificity of 0.945 (range: 0.926 to 0.958) for the diagnosis of ovarian cancers. Hence, SMRP assays are not suitable for screening general patients and may have limited value when utilized as a first-line diagnostic test but may be useful in confirming cancer diagnosis.

HE4, a secreted glycoprotein overexpressed by serous and endometrioid epithelial ovarian carcinomas (Drapkin et al., 2005), is believed to be a sensitive and specific screening test that can detect ovarian cancer at a curative stage. Currently, CA-125 remains the most widely used biomarker for ovarian cancer. Medeiros et al. (2009) conducted a meta-analysis to estimate the accuracy of CA-125 assays in the diagnosis of ovarian tumors. The pooled sensitivity and specificity for the diagnosis of borderline tumors or ovarian cancer were 0.80 (95% CI: 0.76 to 0.82) and 0.75 (95% CI: 0.73 to 0.77), respectively. The meta-analysis performed in the current paper showed that HE4 assays had pooled sensitivity and specificity of 74% (95% CI: 0.72 to 0.77) and 86% (95% CI: 0.84 to 0.87), respectively. HE4 assays had similar sensitivity and greater specificity than CA-125 assays. AUCs were used in the current meta-analysis to compare test accuracies other than pooled sensitivities and specificities. The estimation of the overall performance and test accuracy is difficult because pooled sensitivity and specificity are often used in homogeneity studies. In the present study, HE4 and SMRP showed significant heterogeneity. Thus, a simple pooling of sensitivity and specificity is inappropriate. In addition, HE4 had a higher sensitivity and a lower specificity compared with SMRP when sensitivity or specificity was pooled. Thus, drawing a clear conclusion is difficult. The SROC curve presents a global summary of test performance and shows the trade-off between sensitivity and specificity. Our data showed that the AUCs of HE4 + SMRP, CA125 + SMRP, HE4, SMRP and CA125 were 0.9297, 0.9251, 0.8887, 0.8231 and 0.8689, respectively. Hence, the accuracy of HE4 + SMRP or CA-125 + SMRP for predicting ovarian cancers was better than that of HE4 or SMRP alone.

DOR is a single overall indicator of diagnostic accuracy. It indicates the frequency (expressed as odds)

of positive test results among patients with the condition of interest compared with patients without the condition (Glas et al., 2003). The value of a DOR ranges from 0 to infinity; higher values indicate better discriminatory test performances (higher accuracy). The current meta-analysis showed that SMRP + HE4 exhibited the highest DOR in detecting ovarian cancer (DOR = 54.86) compared with CA125 + SMRP (DOR = 44.88) or CA125 + SMRP (DOR = 43.40). However, the SROC curve and DOR are not easy to interpret and use in clinical practice, and likelihood ratios are considered to be more clinically meaningful (Deeks, 2001). The current work included both PLR and NLR as the measures of diagnostic accuracy, which indicates high accuracy. Likelihood ratios greater than 10 or less than 0.1 generate large and often conclusive changes from pre-test to post-test probability. Likelihood ratios of 5 to 10 and 0.1 to 0.2 generate moderate but substantial shifts from pre-test to post-test probability (Jaeschke et al., 1994). In the current study, the PLR values of HE4 + CA-125 and CA-125 + SMRP were 11.42 and 12.03, respectively. These values are high enough for clinical purposes. However, the NLR values of HE4 + CA-125 and CA-125 + SMRP were 0.23 and 0.42, respectively. When HE4 + CA-125 or CA-125 + SMRP assay results are negative, the probability that the patient has ovarian cancer is 23% or 42%, which is not low enough to rule out ovarian cancer.

The current meta-analysis has several limitations. First, the exclusion of conference abstracts and unpublished data may have led to publication bias. Second, despite rigorous application of QUADAS criteria, most results showed heterogeneity in the assessment of sensitivity or specificity, and the heterogeneity in the diagnostic performance application of QUADAS criteria could not be clearly explained. This result may be mainly attributed to one or a combination of confounding factors in the literature. These factors include, the unclear age of the enrolled women, different study designs, study population, study quality, test equipment or methods, menstrual state, and different cut-off values of the HE4 or SMRP assays.

The current meta-analysis results showed that the diagnostic accuracy of HE4 to differentiate ovarian cancer from benign gynecologic diseases was better than that of SMRP. The combination of two or more tumor markers seems to be more sensitive and specific. Based on the gathered data, the serum tumor marker CA-125, in combination with HE4 or SMRP, can be used to help predict the presence of malignancies in patients with a pelvic mass.

Acknowledgements

References

Abdel-Azeez HA, Labib HA, Sharaf SM, Refaie AN (2010). He4 and mesothelin: Novel biomarkers of ovarian carcinoma in patients with pelvic masses. *Asian Pac J Cancer Prev*, **11**, 111-6.

Badgwell D, Lu Z, Cole L, et al (2007). Urinary mesothelin provides greater sensitivity for early stage ovarian cancer than serum mesothelin, urinary hcg free beta subunit and urinary hcg beta core fragment. *Gynecol Oncol*, **106**, 490-7.

Bast RC, Jr (2004). Early detection of ovarian cancer: New technologies in pursuit of a disease that is neither common nor rare. *Trans Am Clin Climatol Assoc*, **115**, 233-47; discussion 47-8.

Chang K, Pastan I (1996). Molecular cloning of mesothelin, a differentiation antigen present on mesothelium, mesotheliomas, and ovarian cancers. *P Natl Acad Sci USA*, 93, 136-40.

Chang XH, Ye X, Dong L, et al (2011). Human epididymis protein 4 (he4) as a serum tumor biomarker in patients with ovarian carcinoma. *Int J Gynecol Cancer*, **21**, 852-8.

Deeks JJ (2001). Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. *BMJ*, **323**, 157-62.

DerSimonian R, Laird N (1986). Meta-analysis in clinical trials. *Controlled Clin Trials*, **7**, 177-88.

Dinnes J, Deeks J, Kirby J, Roderick P (2005). A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. *Health Technol Asses*, **9**, 1-6.

Drapkin R, von Horsten HH, Lin Y, et al (2005). Human epididymis protein 4 (he4) is a secreted glycoprotein that is overexpressed by serous and endometrioid ovarian carcinomas. *Cancer Res*, **65**, 2162-9.

Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM (2003). The diagnostic odds ratio: A single indicator of test performance. *J Clin Epidemiol*, **56**, 1129-35.

Hassan R, Remaley AT, Sampson ML, et al (2006). Detection and quantitation of serum mesothelin, a tumor marker for patients with mesothelioma and ovarian cancer. *Clin Cancer Res*, **12**, 447-53.

Hellstrom I, Raycraft J, Hayden-Ledbetter M, et al (2003). The he4 (wfdc2) protein is a biomarker for ovarian carcinoma. *Cancer Res*, **63**, 3695-700.

Holcomb K, Vucetic Z, Miller C, Knapp RC (2011). Human epididymis protein 4 offers superior specificity in the differentiation of benign and malignant adnexal masses in premenopausal women. *Am J Obstet Gynecol*, **205**, 358.e1-6.

Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J (2006). Assessing heterogeneity in meta-analysis: Q statistic or i2 index? *Psychol Methods*, **11**, 193-206.

Huhtinen K, Suvitie P, Hiissa J, et al (2009). Serum he4 concentration differentiates malignant ovarian tumours from ovarian endometriotic cysts. *Brit J Cancer*, **100**, 1315-9.

Jacob F, Meier M, Caduff R, et al (2011). No benefit from combining he4 and ca125 as ovarian tumor markers in a clinical setting. *Gynecol Oncol*, **121**, 487-91.

Jacobs I, Bast RC, Jr (1989). The ca 125 tumour-associated antigen: A review of the literature. *Hum Reprod*, **4**, 1-12.

Jaeschke R, Guyatt G, Sackett DL (1994). Users guides to the medical literature .3. How to use an article about a diagnostic-test .A. Are the results of the study valid. *Jama-J Am Med Assoc*, **271**, 389-91.

Jemal A, Siegel R, Ward E, et al (2006). Cancer statistics, 2006. *CA: Cancer J Clin,* **56**, 106-30.

McIntosh MW, Drescher C, Karlan B, et al (2004). Combining ca 125 and smr serum markers for diagnosis and early detection of ovarian carcinoma. *Gynecol Oncol*, **95**, 9-15.

Medeiros LR, Rosa DD, da Rosa MI, Bozzetti MC (2009). Accuracy of ca 125 in the diagnosis of ovarian tumors: A quantitative systematic review. *Eur J Obstet Gynecol Reprod Biol*, **142**, 99-105.

Montagnana M, Lippi G, Ruzzenente O, et al (2009). The utility

of serum human epididymis protein 4 (he4) in patients with a pelvic mass. *J Clin Lab Anal*, **23**, 331-5.

Moore RG, Brown AK, Miller MC, et al (2008). The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. *Gynecol Oncol*, **108**, 402-8.

Nolen B, Velikokhatnaya L, Marrangoni A, et al (2010). Serum biomarker panels for the discrimination of benign from malignant cases in patients with an adnexal mass. *Gynecol Oncol*, **117**, 440-5.

Palmer C, Duan XB, Hawley S, et al (2008). Systematic evaluation of candidate blood markers for detecting ovarian cancer. *Plos One,* **3**, e2633.

Park Y, Kim Y, Lee EY, Lee JH, Kim HS (2012). Reference ranges for he4 and ca125 in a large asian population by automated assays and diagnostic performances for ovarian cancer. *Int J Cancer,* **130**, 1136-44.

Salceda S, Tang T, Kmet M, et al (2005). The immunomodulatory protein b7-h4 is overexpressed in breast and ovarian cancers and promotes epithelial cell transformation. *Exp Cell Res*, **306**, 128-41.

Scholler N, Fu N, Yang Y, et al (1999). Soluble member(s) of the mesothelin/megakaryocyte potentiating factor family are detectable in sera from patients with ovarian carcinoma. *P Natl Acad Sci USA*, 96, 11531-6.

Scholler N, Lowe KA, Bergan LA, et al (2008). Use of yeast-secreted in vivo biotinylated recombinant antibodies (biobodies) in bead-based elisa. *Clin Cancer Res*, **14**, 2647-55.

Shah CA, Lowe KA, Paley P, et al (2009). Influence of ovarian cancer risk status on the diagnostic performance of the serum biomarkers mesothelin, he4, and ca125. *Cancer Epidem Biomar*, **18**, 1365-72.

Van Gorp T, Cadron I, Despierre E, et al (2011). He4 and ca125 as a diagnostic test in ovarian cancer: Prospective validation of the risk of ovarian malignancy algorithm. *Brit J Cancer*, **104**, 863-70.

Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J(2003). The development of quadas: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*, **3**, 25.

Williams TI, Toups KL, Saggese DA, et al (2007). Epithelial ovarian cancer: Disease etiology, treatment, detection, and investigational gene, metabolite, and protein biomarkers. *J Proteome Res*, **6**, 2936-62.