

Diagnostic accuracy of urine HE4 in patients with ovarian cancer: a meta-analysis

Meng-Meng Jia¹, Jie Deng¹, Xiao-lin Cheng¹, Zhen Yan¹, Qing-Chun Li¹, Ying-Ying Xing¹, Dong-Mei Fan¹, Xiao-Yu Tian¹

¹Department of Obstetrics and Gynecology, The First Affiliated Hospital, and College of Clinical Medicine of Henan University of Science and Technology, Luoyang, 471003, China

Correspondence to: Xiao-Yu Tian, email: tianxiaoyu816@163.com

Keywords: HE4, tumour marker, ovarian cancer, diagnosis, meta-analysis

Received: September 09, 2016

Accepted: November 23, 2016

Published: December 26, 2016

ABSTRACT

Urine HE4 has been reported as the potential novel diagnostic biomarker for ovarian cancer in several studies, but their results were inconsistent. Therefore, we conducted a systematic analysis to evaluate the diagnostic value of urine HE4 in detecting ovarian cancer. A comprehensive electronic and manual search was conducted for relevant literatures through several databases up to May 5, 2016. The quality of the studies included in the systematic review was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool. All analyses were conducted using Meta-DiSc 1.4 and STATA 12.0 software. A total of seven publications were included in this study, and these studies included 413 ovarian cancer patients and 573 controls. The summary estimates were: sensitivity 0.76 (95% confidence interval [CI]: 0.72–0.80), specificity 0.92 (95% CI: 0.89–0.94), positive likelihood ratio 8.39 (95% CI: 4.81–14.63), negative likelihood ratio 0.23 (95% CI: 0.13–0.39), diagnostic odds ratio 37.90 (95% CI: 18.69–76.83), and area under the curve 0.93. According to our results, urine HE4 has greater diagnostic value in detecting ovarian cancer. In addition, considering the high heterogeneity, further research studies with more well-designed and large sample sizes are needed in the future.

INTRODUCTION

Ovarian cancer, the most lethal tumour type among gynaecologic cancers, is the fifth most common cause of cancer death in women [1]. Each year, 220 000 women develop epithelial ovarian cancer worldwide [2]. Since its lethality may be due to the lack of specific symptoms, effective screening strategies and early diagnostic methods, over 75% of the patients are diagnosed at an advanced stage (stage III or IV), and the 5-year survival rate is only 20%-25% [3]. Nevertheless, the 5-year survival rate for stage I ovarian cancer is up to 90%. Therefore, the early diagnosis of malignant ovarian tumours is a key factor for improving the survival rate of patients.

Histopathological examination is currently considered the gold standard in the diagnosis of ovarian cancer, but the invasive nature of obtaining tumour samples creates a challenge that limits its application

in the early diagnosis of ovarian cancer because most patients have no symptoms. Thus, bimanual pelvic examination, transvaginal sonography (TVS) and carbohydrate antigen (CA) 125 levels are widely employed diagnostic tools for the early detection of ovarian cancer. Unfortunately, several high-quality studies have demonstrated that bimanual pelvic examination lacks accuracy as a screening method for distinguishing benign from malignant lesions [4]. TVS is accurate in detecting abnormalities in ovarian volume and morphology but is less reliable in differentiating benign from malignant ovarian tumours, and its diagnostic accuracy is largely affected by the experience of the examiner [5]. CA125 was the only FDA-approved biomarker for ovarian cancer before 2008. Although CA125 is elevated in 80% of women with advanced ovarian cancer, it is increased in only 50% of women with early stage ovarian cancer [6].

Furthermore, CA125 levels are elevated in some benign gynaecological diseases (including endometriosis) and non-gynaecological malignancies [7]. Therefore, it may be necessary to identify other biomarkers for the early detection and disease surveillance of ovarian cancer.

Among a wide spectrum of biomarkers, human epidermis protein 4 (HE4) is the most promising. HE4 protein is encoded by WAP four-disulfide core domain 2 (WFDC2) [8], which was found to be highly expressed in ovarian carcinoma, especially in serous and endometrioid cancers [9]. Unlike CA125, HE4 is not overexpressed in benign ovarian disease, normal ovarian tissue or tumours with low malignant potential [10]. In 2008, HE4 was the first biomarker since CA125 to be approved by the FDA for monitoring patients with ovarian cancer for disease recurrence. In 2010, Hellstrom I et al reported higher urinary concentrations of HE4 in patients with early and late stage ovarian cancer, indicating that urinary and serum HE4 measurements had similar sensitivity and specificity [11]. Although extensive analyses have indicated the potential of urine HE4 as a novel diagnostic marker for ovarian cancer, the previous studies have been limited by relatively small patient populations, and no previously published meta-analyses have addressed this research question. In the present study, we conducted a meta-analysis of data from multiple studies to systematically evaluate the potential of urine HE4 as a non-invasive biomarker for the diagnosis of ovarian cancer.

RESULTS

Screening results

As show in Figure 1, 114 potentially relevant articles were retrieved after the initial database searches. After removing duplicates, we obtained 75 publications. Of these 75 articles, we excluded 62 that were not relevant to our study on the basis of title and abstract. Afterward, the remaining 13 articles were subjected to a full-text review, and 6 articles were excluded (four conference papers and two articles without sufficient information to calculate the sensitivity and specificity). Consequently, we obtained 7 publications [11–17] that met all of the inclusion and exclusion criteria for this meta-analysis.

Characteristics of the included studies and quality assessments

In this meta-analysis, the final set of 7 diagnostic studies included a total of 413 ovarian cancer patients and 573 controls (patients with benign gynaecological disease and healthy women). All the ovarian cancer

patients were diagnosed based on histopathological examination. Regarding the origin of the studies, four studies were performed in Asia (China) [13–16], which compared the value of serum HE4 and urine HE4 in the diagnosis of ovarian cancer, and three were conducted in European countries [11, 12, 17], which analyzed the value of urine HE4 in the diagnosis of ovarian cancer. All the studies were published from 2010 to 2015, and the sample size ranged from 78 to 279. Enzyme-linked immunosorbent assay (ELISA) was used in these studies to determine urine HE4 and serum HE4 levels. The characteristics of the included studies are shown in Table 1.

Details of the methodological assessment are shown in Figure 2. We evaluated the quality of the seven included studies using the QUADAS-2 tool. According to the results of the methodological assessment, all the included studies were of acceptable quality.

Diagnostic accuracy

The between-study variability (i.e., heterogeneity) was high for both sensitivity ($I^2 = 87.7\%$) and specificity ($I^2 = 76.8\%$), suggesting high levels of heterogeneity in the seven studies. Therefore, the random effects model was applied. The threshold effect was the major cause of heterogeneity. In the meta-analysis, the Spearman correlation coefficient was -0.071 ($P = 0.879$), confirming that the threshold effect was not significant and that the heterogeneity was caused by other factors. Overall, the pooled sensitivity and specificity were 0.76 (95% CI, 0.72–0.80) (Figure 3a) and 0.92 (95% CI, 0.89–0.94) (Figure 3b), respectively. In addition, the pooled PLR was 8.39 (95% CI, 4.81–14.63) (Figure 4a), the NLR was 0.23 (95% CI, 0.13–0.39) (Figure 4b), and the DOR was 37.90 (95% CI, 18.69–76.83) (Figure 5). The AUC of the SROC was 0.93 (standard error (SE) = 0.03) with a Q^* of 0.86 (SE = 0.04) (Figure 6). These results demonstrated that urine HE4 is an effective diagnostic marker for ovarian cancer.

Besides, the pooled SEN of serum HE4 in the 4 Asian studies for the diagnosis of ovarian cancer was 0.74 (95% CI, 0.68–0.80) and the pooled SPE was 0.88 (95% CI, 0.84–0.92). The PLR was 5.69 (95% CI, 3.60–9.00), the NLR was 0.30 (95% CI, 0.17–0.55), and the DOR was 19.39 (95% CI, 7.11–52.90), respectively. The AUC of the SROC was 0.95 (Table 2).

Subgroup analysis

The results of the subgroup analysis are shown in Table 2. Of the 7 included trials, there were 4 Asian studies and 3 European studies. Both the pooled sensitivity (0.70 vs. 0.82) and specificity (0.90 vs. 0.93) were slightly lower for the European studies than for the Asian studies, but significant heterogeneity was observed in these subgroups.

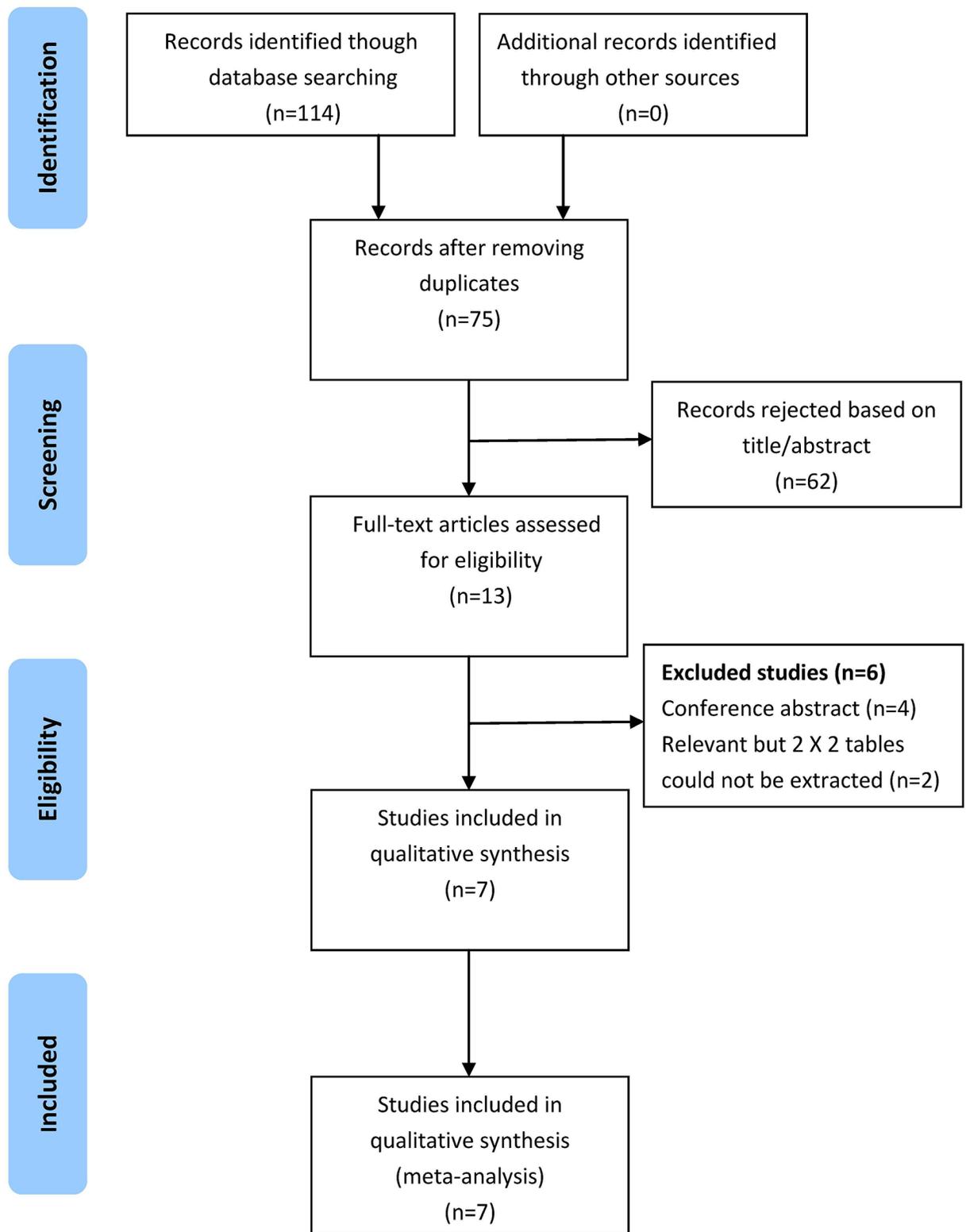


Figure 1: Flowchart depicting the study selection process for this systematic review and meta-analysis.

Table 1: Summary of the diagnostic results of the included studies

study	Year	Country	Simple size		Simple type		Test methods	TP (n)	FP (n)	FN (n)	TN (n)	Cut-off Values (pmol/L)
			Cases	Controls								
			a	b	c	d						
Hellstrom	2010	USA	79	56	Urine	/	ELISA	70	5	9	51	NR
Macuks	2012	Riga	23	55	Urine	/	ELISA	18	14	5	41	1300
Jian HQ	2012	China	30	78	Urine	Serum	ELISA	28	1	2	77	71.519
Yuan ZF	2012	China	50	78	Urine	Serum	ELISA	40	7	10	71	6.51
Zhang YJ	2013	China	98	57	Urine	Serum	ELISA	84	7	14	50	18.74
Wang Y	2014	China	41	62	Urine	Serum	ELISA	28	4	13	58	6.51
Liao	2015	USA	92	187	Urine	/	ELISA	47	10	45	177	Specificity of 95%

Abbreviations: ELASA, Enzyme Linked Immunosorbent Assay; TP, true positive rate; FP, false-positive rate; FN, false-negative rate; TN, true negative rate; NR, not reported;

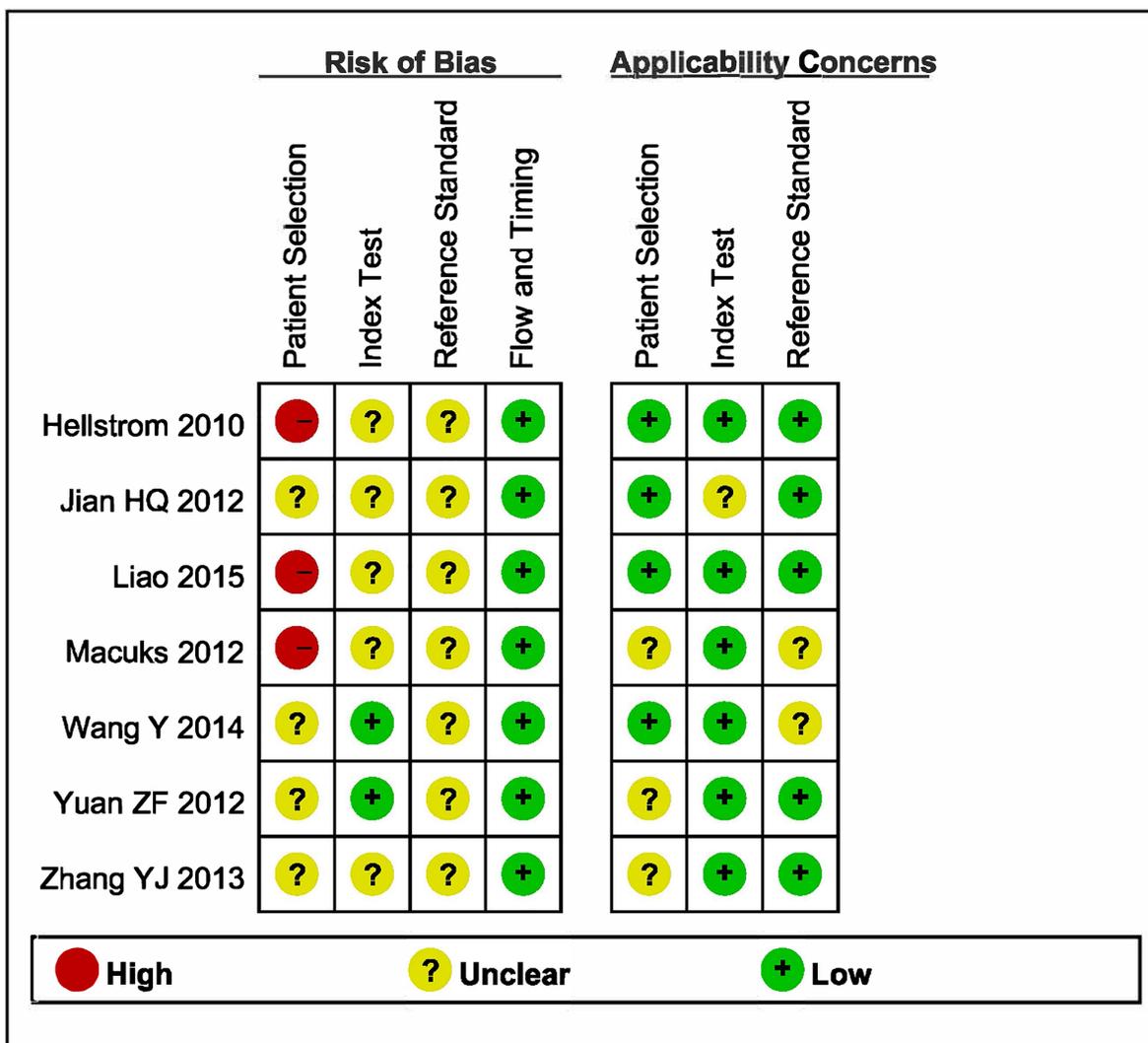


Figure 2: Summary the assessment of methodological quality of included studies by QUADAS-2 tool.

Sensitivity analysis and publication bias

A single study included in this meta-analysis was evaluated each time to determine the influence of the individual data set on the sensitivity and specificity. The results were not impacted by a single study (Table 2). The P-value for the Q test and the I² value also showed that no single study affected the heterogeneity in this meta-analysis. These findings showed that the results of our study were relatively stable.

With regard to publication bias, Deeks' funnel plot asymmetry test was conducted. No significant publication bias was found in the pooled analysis of these studies (P = 0.431) (Figure 7).

DISCUSSION

The aim of this study was to conduct a systematic review to evaluate the accuracy of urine HE4 as a biomarker for ovarian cancer, a disease with high

mortality and a high rate of diagnosis at a late stage. The detection of biomarkers in urine is a less invasive and more convenient way of identifying ovarian cancer. There are some reasons why urinary HE4 analysis may be more sensitive than serum HE4 analysis in certain cases of ovarian cancer. The protein profile in urine may be less complex than that in blood, and certain proteins may be more stable in urine than in blood; moreover, a urinary test would be more convenient than a more invasive blood test [18]. In addition, antibodies to HE4 may be present in women with ovarian carcinoma or with certain types of infertility [19]. These antibodies can be retained by the kidneys; thus, the titre can be analysed at early stages with higher accuracy than serum HE4 measurements [20]. Several urine markers have been identified, including eosinophil-derived neurotoxin, a fragment of osteopontin [21], mesothelin [22, 23], and Bcl-2 [24].

In 2010, Hellstrom et al reported that measuring HE4 in urine can identify patients with ovarian

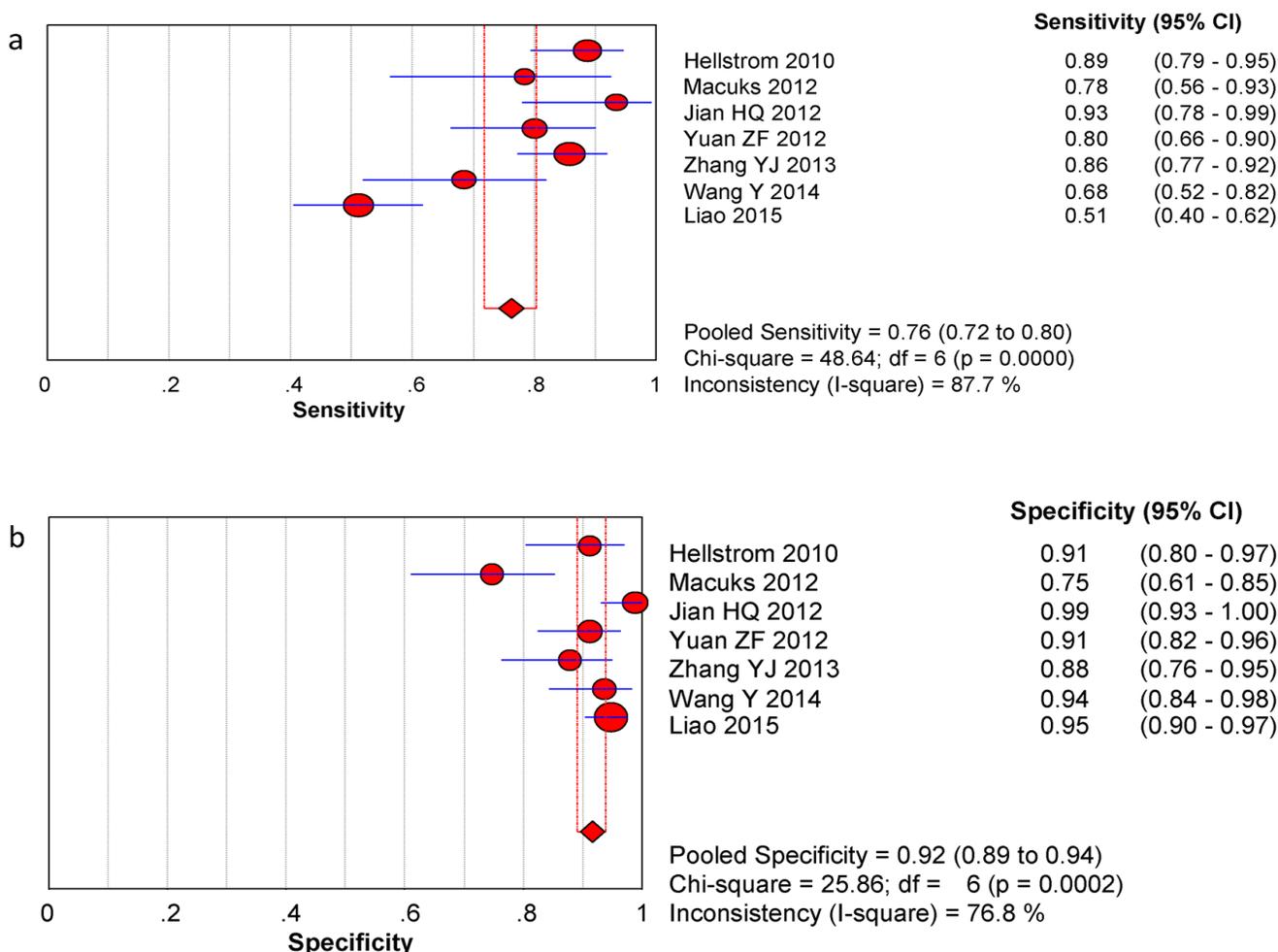


Figure 3: Forest plots of estimated sensitivity (a) and specificity (b) for urine HE4 in the diagnosis of ovarian cancer. a. Forest plots of estimated sensitivity. b. Forest plots of estimated specificity.

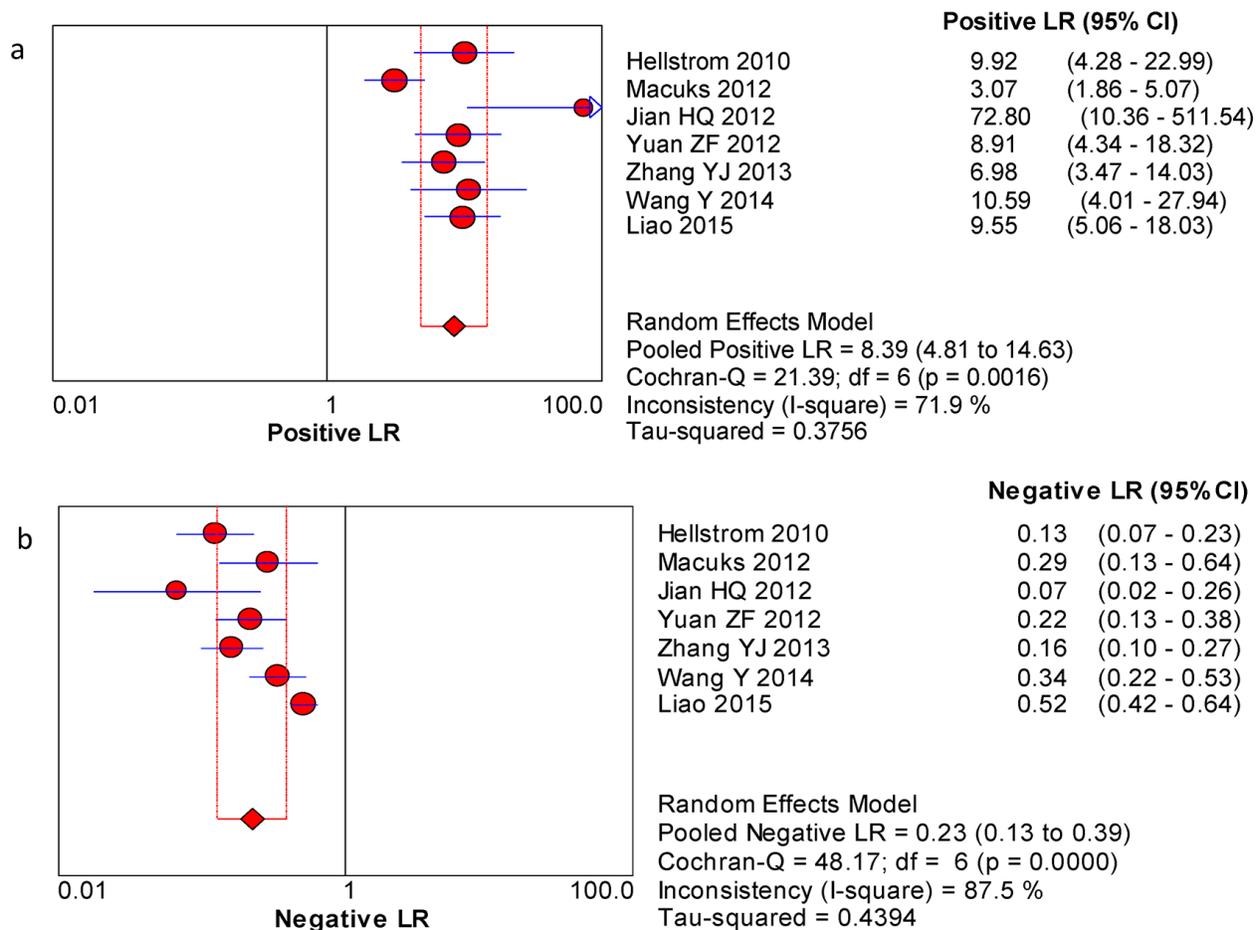


Figure 4: Forest plots of estimated PLR (a) and NLR (b) for urine HE4 in the diagnosis of ovarian cancer. a. Forest plots of estimated positive LR. **b.** Forest plots of estimated negative LR.

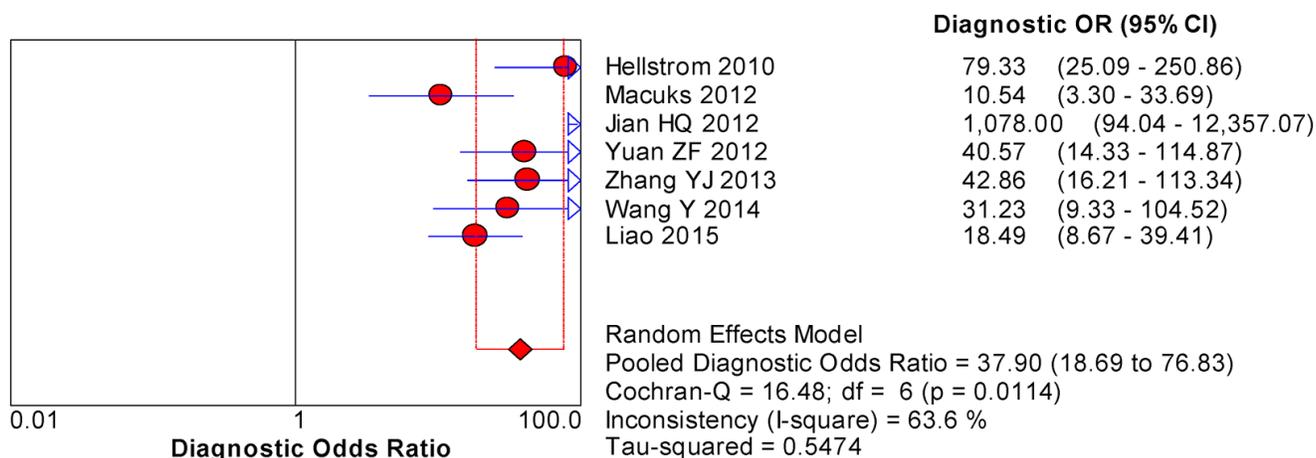


Figure 5: Forest plots of the pooled diagnostic odds ratio (DOR) for urine HE4 in the diagnosis of ovarian cancer.

cancer with a similar sensitivity to assaying sera. However, changes in urine creatinine and urine protein concentration are related; therefore, the ratio of urine HE4/creatinine in early and late ovarian cancer was a better predictor than urine HE4 alone. This result was consistent with the conclusions by Liao et al. However, Macuks et al.[12] found that the diagnostic value of urine HE4 alone was better than that of the urine HE4/creatinine ratio, and Jian HQ [13] reported that there was no significant difference between the two measurements. The predictive value of the urinary HE4/creatinine ratio is better than that of urine HE4, which requires further testing for validation.

To the best of our knowledge, this study is the first meta-analysis to evaluate urine HE4 as a tumour marker for ovarian cancer. The results of the meta-analysis indicated that the pooled sensitivity and specificity of urine HE4 were 0.76 (95% CI, 0.72–0.80) and 0.92 (95% CI, 0.89–0.94) respectively, meaning that 76% of the ovarian cancer patients had high HE4 levels and that

92% of the non-ovarian cancer patients had low HE4 levels. DOR combines the strengths of sensitivity and specificity as a prevalence-independent indicator and is useful from a statistical standpoint [25]. DOR values range from 0 to infinity, with higher values indicating better discriminatory test performance [26]. The DOR value of 37.90 indicates that urine HE4 could be a useful biomarker for ovarian cancer diagnostics. In addition, a good diagnostic test with the best performance will have a PLR greater than 10 and an NLR less than 0.1 [27]. Nevertheless, this review found an overall PLR of 8.39 and an NLR of 0.23, indicating that urine HE4 can neither confirm nor exclude patients with ovarian cancer. Overall, although the sensitivity was compromised, urine HE4 had a good specificity in the diagnosis of ovarian cancer. Urine HE4 was an effective biomarker for ovarian cancer diagnosis.

Of the 7 included trials, only 4 Asian studies had directly compared the diagnostic value of urine HE4 and serum HE4. The results of meta-analyses showed that

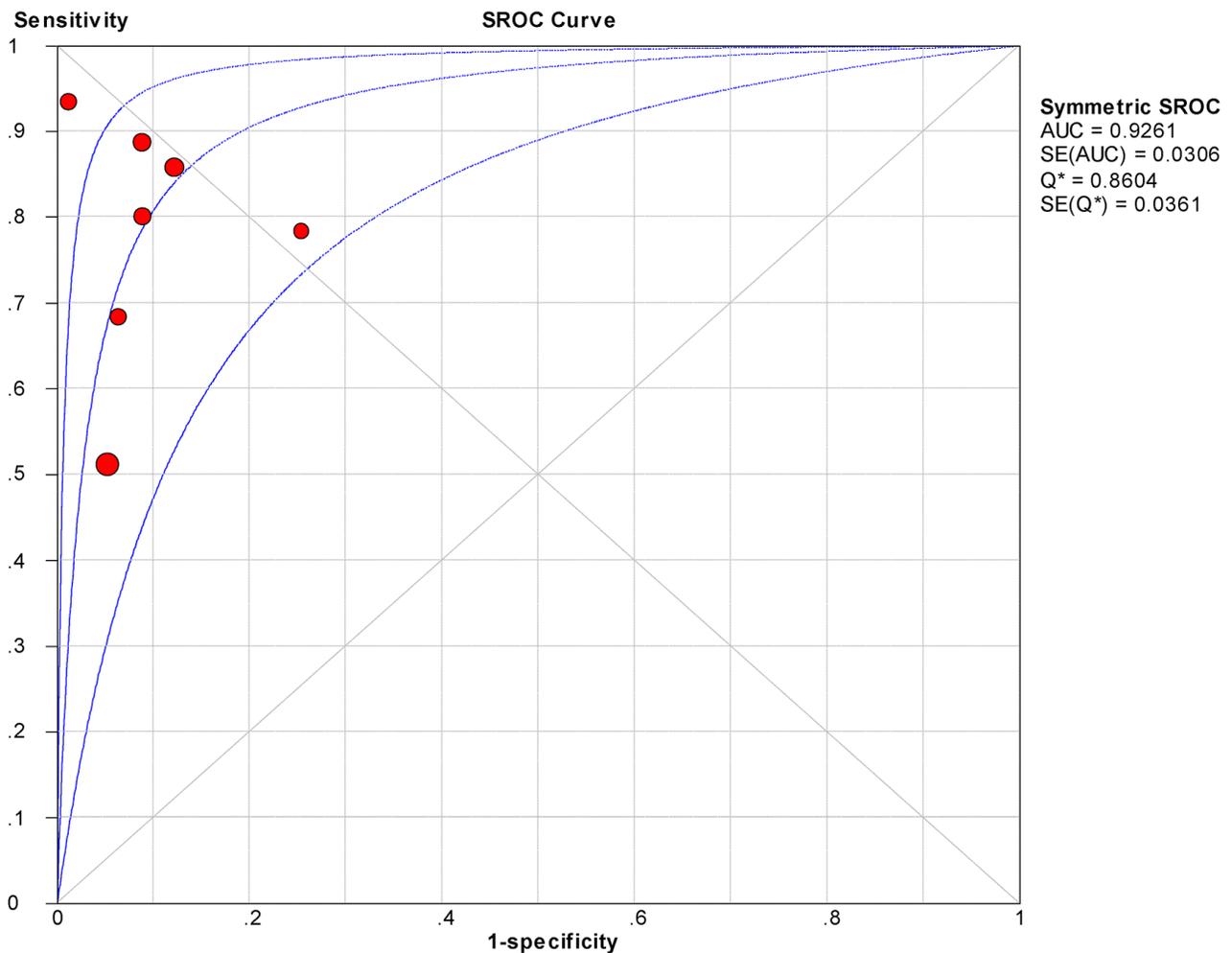


Figure 6: Summary receiver operating characteristic (SROC) curve for urine HE4 in the diagnosis of ovarian cancer.

Table 2: The result of subgroup analyses and sensitivity analysis

Variables	SEN (95% CI)	I ²	SPE (95% CI)	I ²	PLR (95% CI)	NLR (95% CI)	DOR (95% CI)	AUC
Subgroup analysis								
Asia	0.82 (0.72-0.80)	66.1	0.93 (0.89-0.96)	63.7	10.22 (5.54-18.85)	0.20 (0.12-0.34)	56.40 (21.54-147.68)	0.92
European	0.70 (0.63-0.76)	93.5	0.90 (0.86-0.93)	87.6	6.42 (2.56-16.10)	0.27 (0.10-0.74)	24.37 (8.48-69.99)	0.91
Sample types								
Serum	0.74 (0.68-0.80)	85.0	0.88 (0.84-0.92)	0.0	5.69 (3.60-9.00)	0.30 (0.17-0.55)	19.39 (7.11-52.90)	0.95
Sensitivity analysis								
Hellstrom 2010	0.73 (0.68-0.78)	87.3	0.92 (0.89-0.94)	80.6	8.29 (4.38-15.71)	0.25 (0.15-0.44)	33.40 (15.56-71.72)	0.91
Macuks 2012	0.76 (0.72-0.80)	89.7	0.93 (0.91-0.95)	46.1	9.49 (6.79-13.28)	0.21 (0.12-0.40)	47.17 (22.35-95.38)	0.95
Jian HQ 2012	0.75 (0.70-0.79)	88.1	0.91 (0.88-0.93)	71.2	7.22 (4.45-11.71)	0.25 (0.14-0.43)	46.17 (22.35-53.62)	0.92
Yuan ZF 2012	0.76 (0.71-0.80)	89.6	0.92 (0.89-0.94)	80.6	8.52 (4.38-16.60)	0.22 (0.12-0.42)	38.66 (16.59-90.06)	0.93
Zhang YJ 2013	0.73 (0.68-0.78)	88.0	0.92 (0.89-0.94)	79.8	8.99 (4.52-17.88)	0.24 (0.13-0.43)	38.29 (16.33-89.78)	0.92
Wang Y 2014	0.77 (0.73-0.81)	89.4	0.91 (0.89-0.94)	80.4	8.21 (4.39-15.37)	0.20 (0.10-0.41)	40.24 (17.56-92.23)	0.93
Liao 2015	0.83 (0.79-0.87)	55.3	0.90 (0.87-0.93)	77.6	8.42 (4.29-16.53)	0.20 (0.13-0.29)	45.30 (19.89-103.20)	0.91
Total	0.76 (0.72-0.80)	87.7	0.92 (0.89-0.94)	76.8	8.39 (4.81-14.63)	0.23 (0.13-0.39)	37.90 (18.69-76.83)	0.93

Abbreviations: SEN, sensitivity; SPE, specificity; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; CI, confidence interval.

compared with serum HE4, urine HE4 had demonstrated a higher pooled sensitivity (0.82 vs. 0.74), and a higher pooled specificity (0.93 vs. 0.88). When compared with serum HE4, the summary DOR of urine HE4 was better (56.40 vs. 19.39), the pooled PLR of urine HE4 was better (10.22 vs. 5.69), and the pooled NLR was higher (0.30 vs.

0.20). The AUC of SROC for urine HE4 and serum HE4 were 0.92 and 0.95, respectively, which indicated that urine HE4 may be superior to serum HE4 in screening ovarian cancer. More studies are needed for future analyses.

Heterogeneity is a potential problem when interpreting the results of any meta-analysis [28]. In

the present study, we found considerable heterogeneity among the included studies, and the Spearman analysis showed that heterogeneity could not be explained by a threshold effect. Subgroup analyses were performed based on ethnicity to explore the potential source of heterogeneity. Urine HE4 exhibited higher diagnostic accuracy for ovarian cancer detection in non-Caucasian persons than in Caucasian persons, but the heterogeneity was still very high. We speculate that the inclusion of ovarian cancer patients at different disease stages and the use of different cut-off values, types of instruments, and reagent sources contributed to the heterogeneity. Due to the limited number of eligible studies, we could not further elucidate the source of the heterogeneity. Thus, these hypotheses need to be investigated in future studies.

In addition to heterogeneity, there were other limitations. First, we included studies that enrolled healthy women in the control group, which may have resulted in a false increase in the pooled specificity for differentiating benign disease. Second, our study evaluated the performance of HE4 regardless of menopausal status, age, smoking status, decreased

renal function, or estimated glomerular filtration rate [29, 30]. This may have introduced bias in the conclusions, which may make it hard to extrapolate the results. Third, as the research focusing on the diagnosis value of urine HE4 has only been conducted recently, there are only seven studies and the sample sizes of studies included in this meta-analysis are small, which may appear small-study effects. Fourth, there was significant heterogeneity among our included studies, which might indicate the presence of factors that introduce bias. Fifth, only articles published in English or Chinese were included in this meta-analysis, which could have introduced inevitable bias.

In conclusion, the current evidence suggests that urine HE4 is a useful diagnostic biomarker for ovarian cancer, and diagnosis using this non-invasive method could be highly efficient. However, due to the quality of the included studies and the fact that patients suspected of having ovarian cancer were not included, additional studies must be performed to verify the results of this study and the accuracy of this method in terms of achieving high-quality diagnoses.

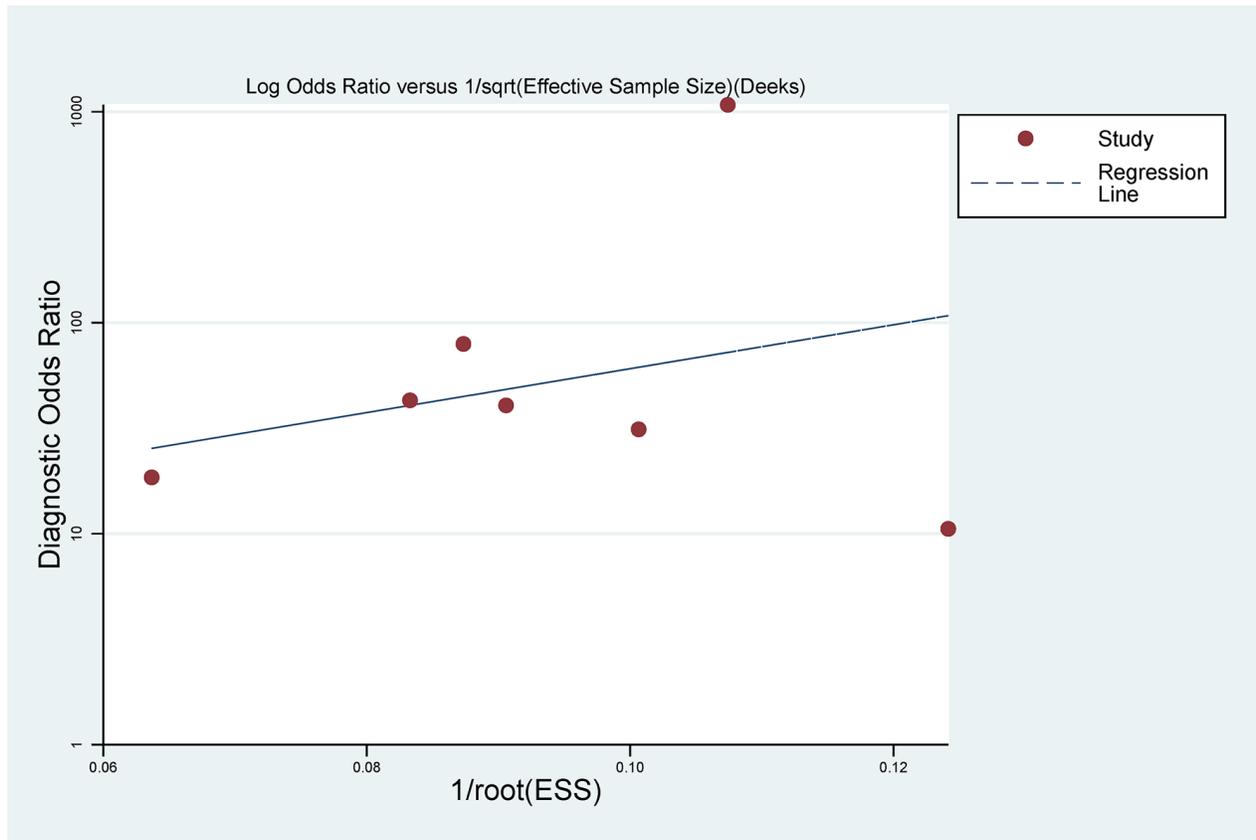


Figure 7: Deek's Funnel Plot Asymmetry Test for the assessment of potential publication bias.

MATERIALS AND METHODS

Search strategy

The meta-analysis was conducted following the criteria of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [31]. The protocol of this review was registered in PROSPERO (<http://www.crd.york.ac.uk/PROSPERO/>), and the registration number is CRD42016038755. The protocol and the PRISMA checklist have been uploaded as supporting information (S1 PRISMA Checklist and S1 Protocol).

Inclusion and exclusion criteria

Inclusion criteria were as follows: (1) a case group of patients with a diagnosis of pathologically confirmed primary ovarian cancer and a control group of women with pathologically confirmed benign gynaecological disease and/or healthy women; this criterion, pathological examination of biopsy specimens, represents the diagnostic gold standard in accordance with the International Federation of Gynecology and Obstetrics guidelines; (2) studies regarding the diagnostic potential of urine HE4 for ovarian cancer; and (3) studies reporting sufficient information regarding the sensitivity and specificity of HE4 to construct two × two contingency tables.

Exclusion criteria were as follows: (1) letters, editorials, case reports, reviews or studies without complete data; (2) sensitivity and specificity were not reported or could not be calculated; and (3) studies with duplicate data reported in other studies.

Data extraction

Two reviewers independently selected all of the literature based on the inclusion and exclusion criteria. Then, these reviewers extracted the following information from the eligible studies: author, year of publication, country of origin, sample size, assay methods, cut-off values, and data regarding true positive (TP), false positive (FP), false negative (FN) and true negative (TN) rates with histology as the gold standard. Disagreements on the eligibility of studies were resolved by full-text review and discussion; if a consensus could not be reached, then a third reviewer was consulted.

Methodological quality assessment of the included studies

Two reviewers independently assessed the quality and potential for bias of all the studies using the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool [32]. This tool is composed of two parts: the risks of bias and concerns regarding applicability. The risks of bias were assessed in four domains: patient selection, index test, reference standard,

and flow and timing. The concerns regarding applicability were assessed in three domains: patient selection, index test, and reference standard. A study received an overall judgement of “low risk of bias” or “low concern regarding applicability” if it was judged as “low” on all domains. In contrast, it would be judged as “risk of bias” or having “concerns regarding applicability” if it scored “high” or “unclear” in one or more domains. This process was performed using Review Manager 5 (<http://ims.cochrane.org/revman/download>).

Statistical analyses

The statistical analysis was performed using Meta-DiSc software (version 1.4) [33]. We extracted the number of participants with a TP, FP, FN or TN from each study to calculate the pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio (DOR) and corresponding 95% confidence intervals (95% CI). Forest plots of accuracy indexes were also constructed. To describe the relationship between test sensitivity and specificity, a summary receiver operating characteristic (SROC) curve was constructed based on the TP and FP rates. An area under the curve (AUC) close to 1 indicated a good diagnostic performance of urine HE4 [34].

A threshold effect is an important cause of heterogeneity in diagnostic testing that can be confirmed by the Spearman correlation coefficient and probability (P) value between the logistic regression of sensitivity and 1 – specificity, and $P < 0.05$ indicated a significant threshold effect [35]. Heterogeneity caused by non-threshold effects was assessed by the Q and I^2 tests. $P < 0.10$ for the Q test or an I^2 value greater than 50% indicates substantial heterogeneity, and the random effects model was applied; otherwise, a fixed effects model was adopted. Subgroup and sensitivity analyses were performed to explore the potential sources of between-study heterogeneity. Deeks’ funnel plots were used to assess potential publication bias with STATA 12.0 software [36]. All statistical tests were two-sided, and $P < 0.05$ was considered to indicate statistical significance.

ACKNOWLEDGMENTS

We thank Xiao-Yu Tian M.D. and Dong-Mei Fan M.D. for their discussion and help.

CONFLICTS OF INTEREST

There is no conflict of interest.

Supporting material

S1 Protocol. The protocol is registered in PROSPERO.(PDF)

REFERENCES

1. Clarke-Pearson, DL. Clinical practice. Screening for ovarian cancer. *N Eng J Med.* 2009; 361: 170-177. doi: 10.1056/NEJMcp0901926.
2. Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. *Lancet.* 2014; 384: 1376-1388. doi: 10.1016/s0140-6736(13)62146-7.
3. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature.* 2011; 474: 609-615. doi: 10.1038/nature10166.
4. Ebell MH, Culp M, Lastinger K, Dasigi T. A systematic review of the bimanual examination as a test for ovarian cancer. *Am J Prev Med.* 2015; 48: 350-356. doi: 10.1016/j.amepre.2014.10.007.
5. Van Nagell JR, Hoff JT. Transvaginal ultrasonography in ovarian cancer screening: current perspectives. *Int J Womens Health.* 2013; 6: 25-33. doi: 10.2147/IJWH.S38347.
6. Jelovac D, Armstrong DK. Recent progress in the diagnosis and treatment of ovarian cancer. *CA Cancer J Clin.* 2011; 61: 183-203. doi: 10.3322/caac.20113.
7. Park Y, Lee JH, Hong DJ, Lee EY, Kim HS. Diagnostic performances of HE4 and CA125 for the detection of ovarian cancer from patients with various gynecologic and non-gynecologic diseases. *Clin Biochem.* 2011; 44: 884-888. doi: 10.1016/j.clinbiochem.2011.04.011.
8. Clauss A, Lilja H, Lundwall A. A locus on human chromosome 20 contains several genes expressing protease inhibitor domains with homology to whey acidic protein. *Biochem J.* 2002; 15: 233-242. doi:10.1042/BJ20020869.
9. Drapkin R, von Horsten HH, Lin Y, Mok SC, Crum CP, Welch WR, Hecht JL. Human epididymis protein 4 (HE4) is a secreted glycoprotein that is overexpressed by serous and endometrioid ovarian carcinomas. *Cancer Res.* 2005; 65: 2162-2169. doi:10.1158/0008-5472.CAN-04-3924.
10. Galgano MT, Hampton GM, Frierson HF Jr. Comprehensive analysis of HE4 expression in normal and malignant human tissues. *Mod Pathol.* 2006;19: 847-853. doi: 10.1038/modpathol.3800612.
11. Hellstrom I, Heagerty PJ, Swisher EM, Liu P, Jaffar J, Agnew K, Hellstrom KE. Detection of the HE4 protein in urine as a biomarker for ovarian neoplasms. *Cancer Lett.* 2010; 296: 43-48. doi: 10.1016/j.canlet.2010.03.013.
12. Macuks R, Baidekalna I, Donina S. Urinary concentrations of human epididymis secretory protein 4 (He4) in the diagnosis of ovarian cancer: a case-control study. *Asian Pac J Cancer Prev.* 2012; 13: 4695-4698. doi: 10.7314/apjcp.2012.13.9.4695.
13. Jian HQ, Xu CL. Significance of urine human epididymis protein 4 in the diagnosis and differential diagnosis of ovarian and endometrial carcinoma. 2012. Shijiazhuang, Hebei Province, China:Hebei Medical University, Cnki.
14. Yuan ZF, Shang RY. Diagnostic value of combined detection of human epididymis protein 4 and serum CA125 in diagnosis of epithelial ovarian cancer. *Progress in Obstetrics and Gynecology.* 2012; 21: 354-356.
15. Zhang YJ, Yuan ZF. Expression and significance of C-OPN, OPN and HE4 in cancer tissues, ascites, serum and urine of epithelial ovarian cancer patients. *Maternal and Child Health Care of China.* 2013; 28: 2290-2293. doi: 10.7620/zgfybj.j.issn.1001-4411.2013.28.38.
16. Wang Y, Yang CL, Cheng BR, Guo XX. Preliminary study on the value of urine HE4 combined with CA125 in the detection of ovarian cancer. *Journal of Practical Obstetrics and Gynecology.* 2014; 30: 220-223.
17. Liao JB, Yip YY, Swisher EM, Agnew K, Hellstrom KE, Hellstrom I. Detection of the HE4 protein in urine as a biomarker for ovarian neoplasms: clinical correlates. *Gyneco Oncol.* 2015; 137: 430-435. doi: 10.1016/j.ygyno.2015.03.044.
18. Chambers AF, Vanderhyden BC. Ovarian cancer biomarkers in urine. *Clin Cancer Res.* 2006; 12: 323-327. doi: 10.1158/1078-0432.CCR-05-2563.
19. Hellstrom I, Swisher E, Hellstrom KE, Yip YY, Agnew K, Luborsky JL. Anti-HE4 antibodies in infertile women and women with ovarian cancer. *Gynecol Oncol.* 2013; 130: 629-633. doi: 10.1016/j.ygyno.2013.05.028.
20. Hellstrom I, Hellstrom KE. fTwo novel biomarkers, mesothelin and HE4, for diagnosis of ovarian carcinoma. *Expert Opin Med Diagn.* 2011; 5: 227-240. doi: 10.1517/17530059.2011.559459.
21. Ye B, Skates S, Mok SC, Horick NK, Rosenberg HF, Vitonis A, Edwards D, Sluss P, Han WK, Berkowitz RS, Cramer DW. Proteomic-based discovery and characterization of glycosylated eosinophil-derived neurotoxin and COOH-terminal osteopontin fragments for ovarian cancer in urine. *Clin Cancer Res.* 2006; 12: 432-441. doi: 10.1158/1078-0432.CCR-05-0461.
22. Badgwell D, Lu Z, Cole L, Fritsche H, Atkinson EN, Somers E, Allard J, Moore RG, Lu KH, Bast RC Jr. 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.* 2007; 106: 490-497. doi: 10.1016/j.ygyno.2007.04.022.
23. Hellstrom I, Hellstrom KE. SMRP and HE4 as biomarkers for ovarian carcinoma when used alone and in combination with CA125 and/or each other. *Adv Exp Med Biol.* 2008; 622: 15-21. doi: doi: 10.1007/978-0-387-68969-2_2. Review.
24. Anderson NS, Bermudez Y, Badgwell D, Chen R, Nicosia SV, Bast RC Jr, Kruk PA. Urinary levels of Bcl-2 are elevated in ovarian cancer patients. *Gynecol Oncol.* 2009; 112: 60-67. doi: 10.1016/j.ygyno.2008.09.037.
25. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test

- performance. *J Clin Epidemiol.* 2003; 56: 1129-1135. doi: 10.1016/s0895-4356(03)00177-x.
26. Pai M, McCulloch M, Enanoria W, Colford JM Jr. Systematic reviews of diagnostic test evaluations: What's behind the scenes?. *ACP J Club.* 2004; 141: A11-13.
 27. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *J Natl Cancer Inst.* 2008; 100: 1432-1438. doi: 10.1093/jnci/djn326.
 28. Coory MD. Comment on: Heterogeneity in meta-analysis should be expected and appropriately quantified. *Int J Epidemiol.* 2010; 39: 932; author reply 933. doi: 10.1093/ije/dyp157.
 29. Nagy B Jr, Krasznai ZT, Balla H, Csobán M, Antal-Szalmás P, Hernádi Z, Kappelmayer J. Elevated human epididymis protein 4 concentrations in chronic kidney disease. *Ann Clin Biochem.* 2012; 49: 377-380. doi: 10.1258/acb.2011.011258.
 30. Kappelmayer J, Antal-Szalmás P, Nagy B Jr. Human epididymis protein 4 (HE4) in laboratory medicine and an algorithm in renal disorders. *Clin Chim Acta* 2015; 438: 35-42. doi: 10.1016/j.cca.2014.07.040.
 31. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* 2009; 339: b2535. doi: 10.1136/bmj.b2535.
 32. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011; 155: 529-536.
 33. Zamora J, Abaira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med. Res. Methodol.* 2006; 6: 31. doi: 10.1186/1471-2288-6-31.
 34. Habel LA, Capra AM, Oestreicher N, Greendale GA, Cauley JA, Bromberger J, Crandall CJ, Gold EB, Modugno F, Salane M, Quesenberry C, Sternfeld B. Mammographic density in a multiethnic cohort. *Menopause.* 2007; 14: 891-899. doi: 10.1097/gme.0b013e318032569c.
 35. Devillé WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, Bezemer PD. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol.* 2002; 2: 9.
 36. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol.* 2005; 58:882-893. doi: 10.1016/j.jclinepi.2005.01.016.