A Clinical Diagnostic Value Analysis of Serum CA125, CA199, and HE4 in Women with Early Ovarian Cancer: Systematic Review and Meta-Analysis

2/14 醫學檢驗科 黃致雅

卵巢癌及BIOMARKER介紹

- 在女性生殖系統中，卵巢癌成為三大常見惡性腫瘤之一，
- 約85%-90%的卵巢惡性腫瘤為上皮性卵巢癌。 近年來，卵巢癌的發病率呈逐漸增加的趨勢。由於卵巢癌缺乏的早期症狀和缺乏早期篩查和診斷，患者的生存率不到 30%。但是，如果早期發現分期並給予規範的手術和輔助治療，卵巢癌的5年生存率可高達90%。

- 在卵巢癌患者血清CA125水平可能升高，但該標記在早期階段的卵巢癌敏感性較低。 在其他生理或病理條件下CA125也會增加，如月經、懷孕、子宮內膜異位症和內-腹膜炎性疾病。

- 現已開發標記以改進對卵巢癌的特異性，例如人類附睾蛋白 4 (Human Epididymis Protein 4，以下簡稱HE4) 。
- 該生物標誌物在卵巢癌中過度表達 。儘管這些標記的特異性相當可靠，卻不是很敏感。

比較

	CA125	HE4
	大分子癌症抗原，類似黏液蛋白的醣蛋白	乳清酸性4-二硫化中心（WFDC）蛋白家族
造成**偽陽性**原因	月經，懷孕，出血性囊腫，腹部炎症	懷孕，吸菸，青春期少女(18y以下未納入ROMA)
正常值	<35U/mL	Premenopausal ≦ 40 pmol/L Postmenopausal ≦ 140 pmol/L
檢驗方法	化學冷光微粒免疫分析法CMIA	化學冷光微粒免疫分析法CMIA
		避孕藥，生理期不影響data

步驟1
- 形成臨床問題 →PICO

步驟2
- 搜尋最佳證據 →FAITH

步驟3
- 嚴格評讀證據

步驟4
- 將臨床專業與病人價值觀相結合

步驟5
- 評估執行效果及效用

步驟 1：系統性文獻回顧探討的問題為何？

Question：一個問句簡潔有力表達臨床特定問題			
P	Patient/Problem	病人、臨床問題、患者族群	能否利用HE4、CA125、CA199在早期發現並治療
I	Intervention	關注的(實驗組)： 治療方式、劑量、頻率.... 診斷工具 暴露因子	CA125聯合HE4的檢測
C	Comparison	對照的： 治療：ex. Placebo 診斷工具： ex. Gold standard 暴露因子	單獨檢測CA125 單獨檢測HE4
O	Outcomes	臨床關注、對病患有意義可測量的結果	HE4、CA125、CA199敏感性和特異性 聯合檢測比起單獨檢測的診斷更適合

步驟1：系統性文獻回顧的品質如何？(FAITH)

- F - 研究是否找到 (Find) 所有的相關證據？

良好的文獻搜尋至少應包括二個主要的資料庫 (如：Medline, Cochrane 考科藍實證醫學資料庫, EMBASE 等)，並且加上文獻引用檢索(參考文獻中相關研究、Web of Science, Scopus 或 Google Scholar)、試驗登錄資料等。文獻搜尋應不只限於英文，並且應同時使用MeSH字串及一般檢索詞彙(text words)。

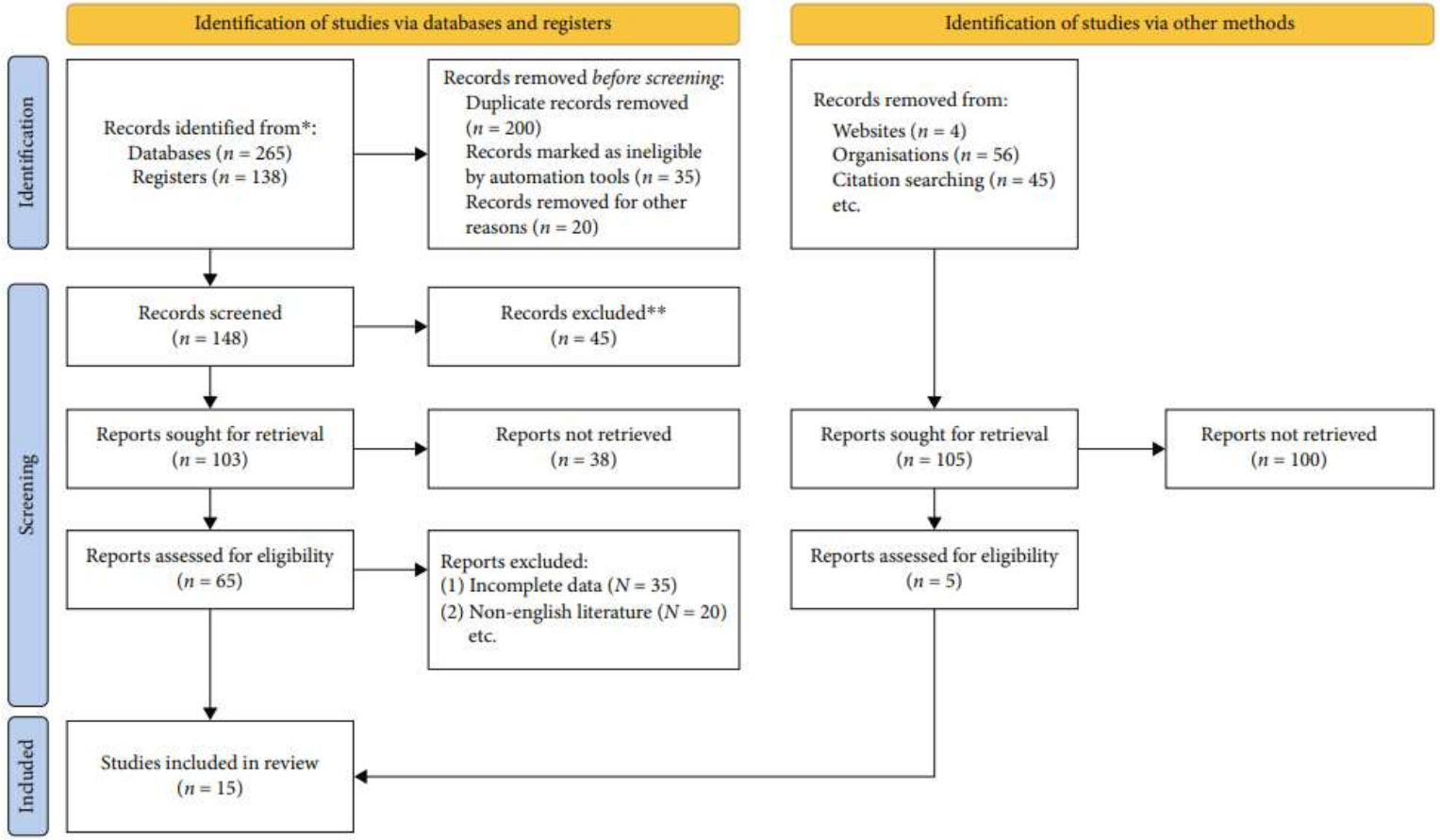
結果(Results) 章節中可以找到本篇系統性文獻回顧評估的摘要及全文文獻數目、文獻納入與排除的數量及原因。資料可能會以圖表或PRISMA的流程圖呈現。

2.3. Retrieval Strategy. PubMed, Elsevier Science, Springer, CNKI, Wanfang, VIP, and other databases were searched by computer. Literature languages are limited to Chinese and English. HE4 and CA125 and ovarian cancer; hE4; ovarian cancer; serum biomarkers; diagnosis, etc. as search terms (Figure 1).

3. Result

3.1. Literature Retrieval Results and Basic Features of Included Studies. After screening, a total of 15 studies were finally included [11–24]. All the included studies were diagnostic tests, including 2262 patients with ovarian cancer, all confirmed by national pathological standards, and the control group included 2300 patients with benign ovarian disease and healthy people, and a total of 4562 cases were included. The basic characteristics included in the study are shown in Table 1.

PRISMA 流程圖



Criteria:

1. Included literature published from January 2005 to December 2021.
2. healthy subjects or patients with the benign disease were taken as a control group.
3. Pathological tissue diagnosis as the gold standard.
4. Complete original data can be obtained.
5. The sample size was ≥20.

評讀結果： ☑ 是 ☐否 ☐不清楚

步驟 2：系統性文獻回顧的品質如何？(FAITH)

- A - 文獻是否經過嚴格評讀 (Appraisal)？

應根據不同臨床問題的文章類型，選擇適合的評讀工具，並說明每篇研究的品質(如針對治療型的臨床問題，選用隨機分配、盲法、及完整追蹤的研究類型)。

方法章節，可以找到所使用的文獻品質評讀標準的描述

結果章節則會列出每篇研究品質的評讀結果。

2.5. Data Extraction. The studies included in this paper were all diagnostic test accuracy studies, and their quality was evaluated from the following aspects: (1) whether the case spectrum included various medical records and cases of easily confused diseases; (2) whether the criteria for the selection of research objects are clear; (3) whether the clinical data available when interpreting test results are consistent with the clinical data available in practice; (4) whether intermediate test results are reported; (5) whether to explain the cases that withdrew from the study.

評讀結果： ☑是 □否 □不清楚

步驟3：系統性文獻回顧的品質如何？(FAITH)

- I - 是否只納入 (Included) 具良好效度的文章？

僅進行文獻判讀是不足夠，系統性文獻回顧只納入至少要有一項研究結果是極小偏誤的試驗。
在文章的方法章節，可以找到文章評估的方式，以及是由誰完成評估的

2.4. Extraction of Literature. This study by two people as evaluators, in strict accordance with the inclusion criteria and exclusion criteria to an independent screening of literature, respectively, after extracting data to cross-check, ensures the quality of literature to extract and review price is the consistency of the results when disagreements are resolved through discussion, as there are still differences through consulting a guidance group of other experts to solve. Extracted data include author, age, country, test method, positive determination value, gold standard, and fruit index.

評讀結果： ☑是 □否 □不清楚

步驟4：系統性文獻回顧的品質如何？(FAITH)

T - 作者是否以表格和圖表「總結」(Total up) 試驗結果？

應該用至少1個摘要表格呈現所納入的試驗結果。若結果相近，可針對結果進行統合分析(meta-analysis)，並以「森林圖」(forest plot)呈現研究結果，最好再加上異質性分析。

在文章的結果章節，可以找到摘要的圖表，以及作者對系統性文獻回顧結果的解釋。

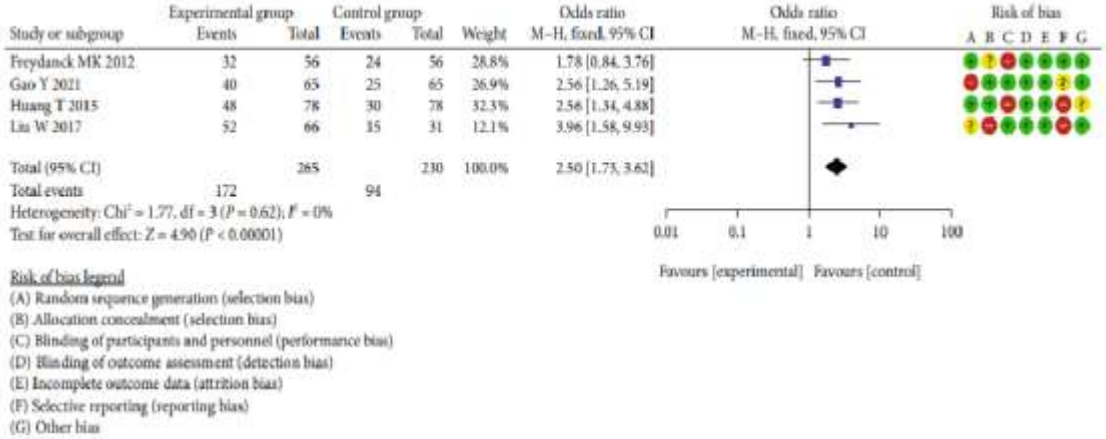
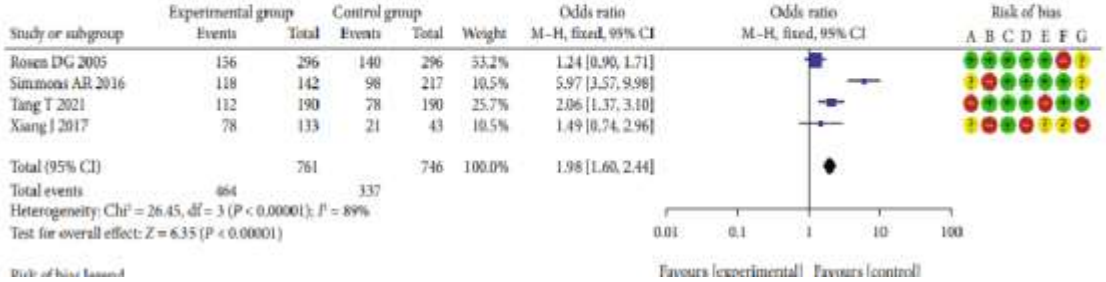


FIGURE 5: Meta-analysis of diagnostic sensitivity analysis of CA125 between two groups.



評讀結果： ☑是 ☐否 ☐不清楚

步驟5：系統性文獻回顧的品質如何？(FAITH)

Systematic review 異質性-文獻結果是否合併，適當與否？

在理想情況下，各個試驗的結果應相近或具同質性，若具有異質性，作者應評估差異是否顯著 (卡方檢定)。根據每篇個別研究中不同的 PICO 及研究方法，探討造成異質性的原因。

常用來闡述研究間異質性的統計方法如下：

I^2值：通常介於0-100%，數字越大異質性越大，但事實上 I^2值低，不代表研究間異質性低，也有可能是檢定力不足。

Low <50%、Moderate 50-75%、High >75%

H - 試驗的結果是否相近 - 異質性 (Heterogeneity)？

- 主要原因在於，並不是所有的系統性回顧 (systematic review)都適合把結果合併成一個值，變成統合分析 (meta-analysis)，例如研究彼此間研究設計差異太大、異質性 (heterogeneity)太大時。
- 通常可以從文章尋找方法 (Methods)中統計分析 (statistical analysis)的描述部分判斷合理性。
- 異質性太大原因:研究間異質性可能由非閾值效應引起，例如受試者差異、測試標準和條件、參考標準、研究間設計和實施方法的差異。

2.7. Statistical Analysis. The heterogeneity among the included studies was analyzed. The threshold effect is one of the main reasons for the heterogeneity of accurate research. The meta-disc soft piece can be used to evaluate the threshold effect by three methods: view test, spearman correlation coefficient calculation, and the image generated by the exact estimator of each study in the SROC curve plan. Interstudy heterogeneity can be caused by nonthreshold effects, such as differences in subjects, differences in test criteria and conditions, reference criteria, interstudy design, and implementation methods. MetaDisc1.4 software provides two methods to evaluate heterogeneity caused by nonthreshold effects: observed forest maps and statistical tests, including chi-square test and Cochran-Q.

評讀結果： ☑是 ☐否 ☐不清楚

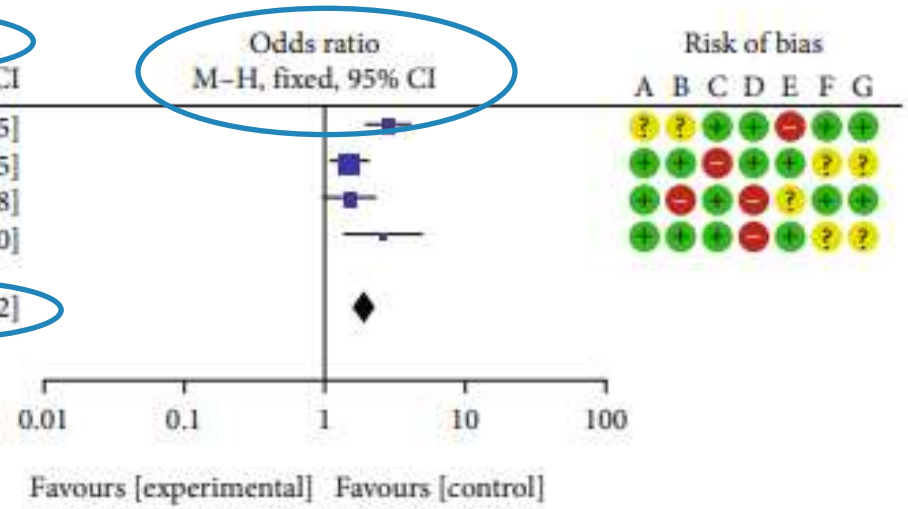
結果:
CA125特異性分析

95% CI (1.58, 2.32)

Study or subgroup	Experimental group Events	Total	Control group Events	Total	Weight	Odds ratio M-H, fixed, 95% CI
Chen F 2018	178	262	112	262	24.1%	2.84 [1.99, 4.05]
Chen X 2015	178	323	145	323	43.7%	1.51 [1.11, 2.05]
Chen X 2017	98	178	80	178	24.2%	1.50 [0.99, 2.28]
Cho HY 2019	50	81	31	81	8.0%	2.60 [1.38, 4.90]
Total (95% CI)		844		844	100.0%	1.91 [1.58, 2.32]
Total events	504		368			

Heterogeneity: $Chi^2 = 9.17$, $df = 3$ ($P = 0.03$); $I^2 = 67\%$

Test for overall effect: $Z = 6.58$ ($P < 0.00001$)



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Figure 4:：兩組間CA125診斷特異性分析。

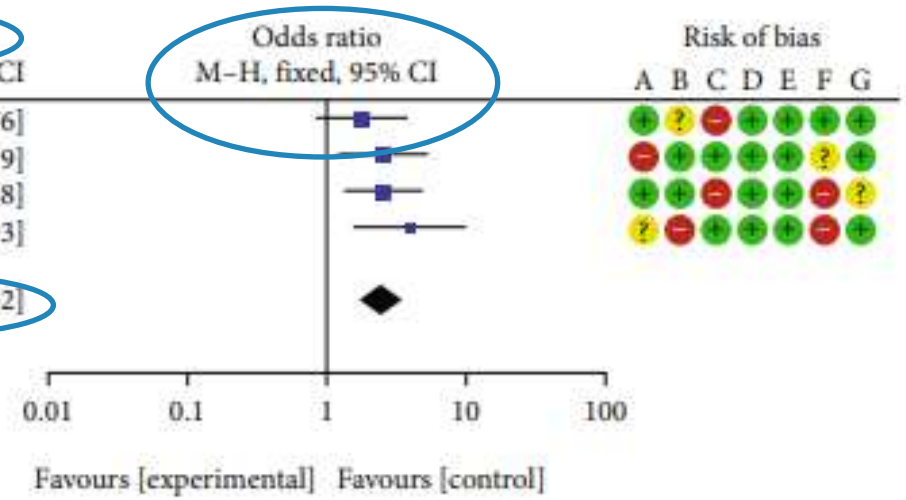
CA125敏感性分析

95% CI(1.73, 3.62)

Study or subgroup	Experimental group Events	Total	Control group Events	Total	Weight	Odds ratio M–H, fixed, 95% CI
Freydanck MK 2012	32	56	24	56	28.8%	1.78 [0.84, 3.76]
Gao Y 2021	40	65	25	65	26.9%	2.56 [1.26, 5.19]
Huang T 2015	48	78	30	78	32.3%	2.56 [1.34, 4.88]
Liu W 2017	52	66	15	31	12.1%	3.96 [1.58, 9.93]
Total (95% CI)		265		230	100.0%	2.50 [1.73, 3.62]
Total events	172		94			

Heterogeneity: $Chi^2 = 1.77$, $df = 3$ ($P = 0.62$); $I^2 = 0\%$

Test for overall effect: $Z = 4.90$ ($P < 0.00001$)



Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

FIGURE 5: Meta-analysis of diagnostic sensitivity analysis of CA125 between two groups.

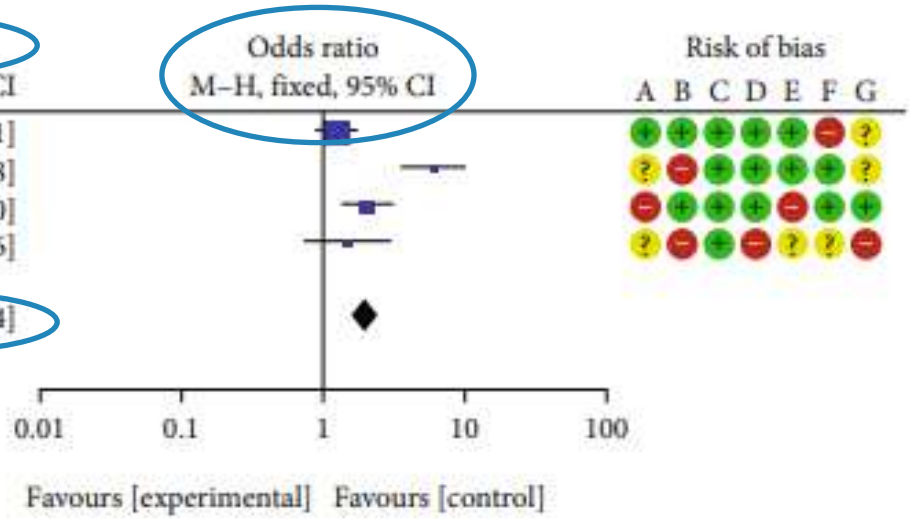
CA199特異性分析

95%CI (1.60, 2.44)

Study or subgroup	Experimental group Events	Total	Control group Events	Total	Weight	Odds ratio M–H, fixed, 95% CI
Rosen DG 2005	156	296	140	296	53.2%	1.24 [0.90, 1.71]
Simmons AR 2016	118	142	98	217	10.5%	5.97 [3.57, 9.98]
Tang T 2021	112	190	78	190	25.7%	2.06 [1.37, 3.10]
Xiang J 2017	78	133	21	43	10.5%	1.49 [0.74, 2.96]
Total (95% CI)		761		746	100.0%	1.98 [1.60, 2.44]
Total events	464		337			

Heterogeneity: $Chi^2 = 26.45$, $df = 3$ ($P < 0.00001$); $I^2 = 89\%$

Test for overall effect: $Z = 6.35$ ($P < 0.00001$)



Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

FIGURE 6: Meta-analysis of diagnostic specificity analysis of CA199 between two groups.

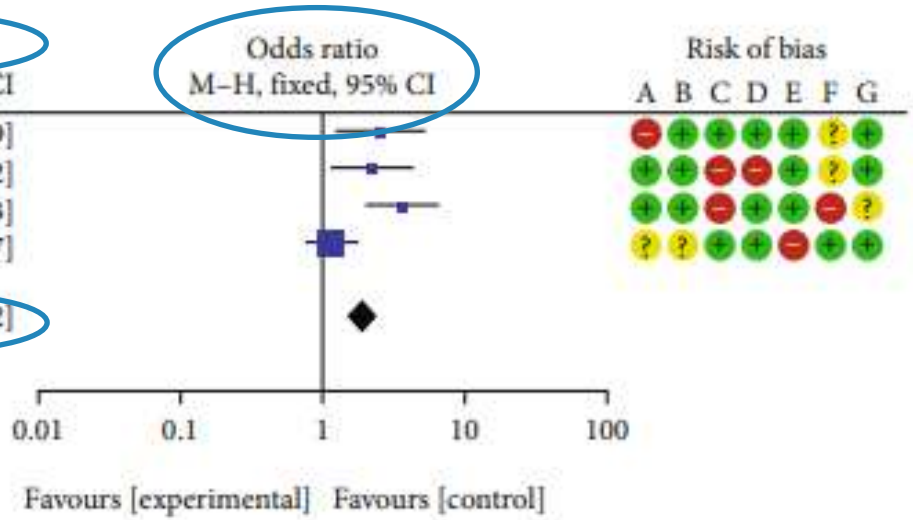
CA199敏感性分析

95%CI (1.46, 2.52)

Study or subgroup	Experimental group Events	Total	Control group Events	Total	Weight	Odds ratio M–H, fixed, 95% CI
Gao Y 2021	40	65	25	65	12.8%	2.56 [1.26, 5.19]
Xu J 2021	45	75	30	75	16.0%	2.25 [1.17, 4.32]
Yu XP 2020	67	102	35	102	16.0%	3.66 [2.06, 6.53]
Zhu C 2021	93	179	86	179	55.1%	1.17 [0.77, 1.77]
Total (95% CI)		421		421	100.0%	1.92 [1.46, 2.52]
Total events	245		176			

Heterogeneity: $Chi^2 = 11.16$, $df = 3$ ($P = 0.01$); $I^2 = 73\%$

Test for overall effect: $Z = 4.70$ ($P < 0.00001$)



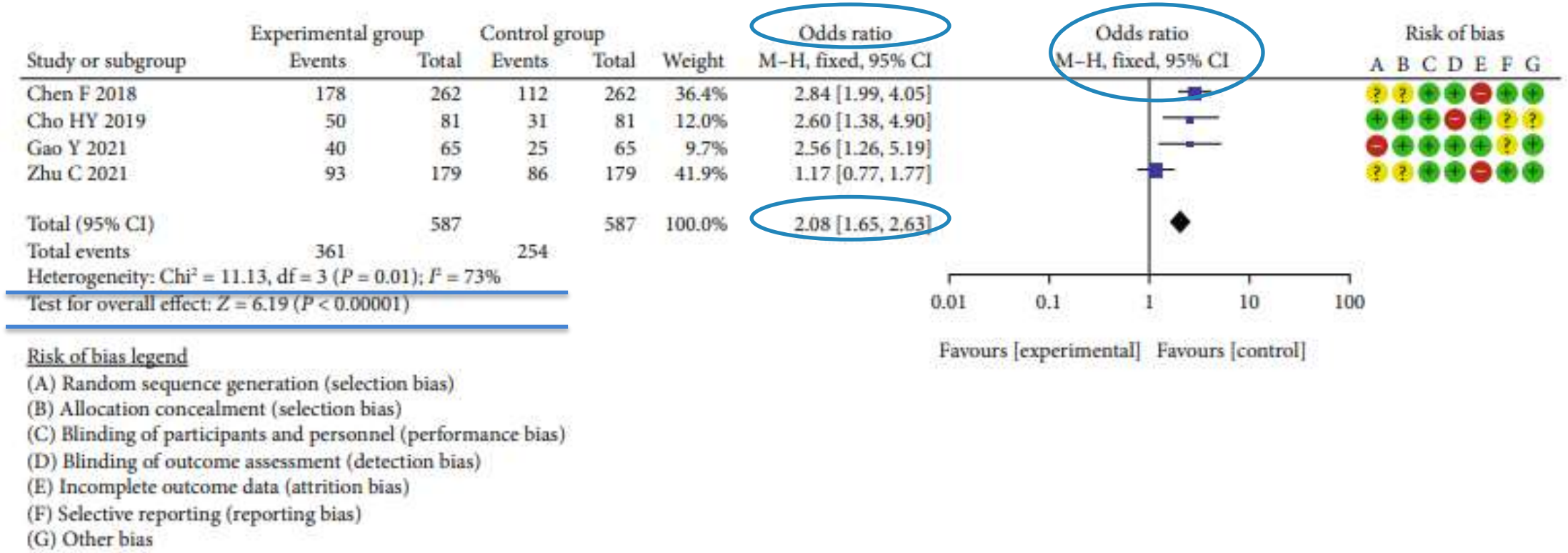
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

FIGURE 7: Meta-analysis of diagnostic sensitivity analysis of CA199 between two groups.

HE4特異性分析

95% CI :2.08 (1.65, 2.63)



Study or subgroup	Experimental group Events	Total	Control group Events	Total	Weight	Odds ratio M–H, fixed, 95% CI
Chen F 2018	178	262	112	262	36.4%	2.84 [1.99, 4.05]
Cho HY 2019	50	81	31	81	12.0%	2.60 [1.38, 4.90]
Gao Y 2021	40	65	25	65	9.7%	2.56 [1.26, 5.19]
Zhu C 2021	93	179	86	179	41.9%	1.17 [0.77, 1.77]
Total (95% CI)		587		587	100.0%	2.08 [1.65, 2.63]
Total events	361		254			

Heterogeneity: $Chi^2 = 11.13$, $df = 3$ ($P = 0.01$); $I^2 = 73\%$

Test for overall effect: $Z = 6.19$ ($P < 0.00001$)

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

FIGURE 8: Meta-analysis of diagnostic specificity analysis of HE4 between two groups.

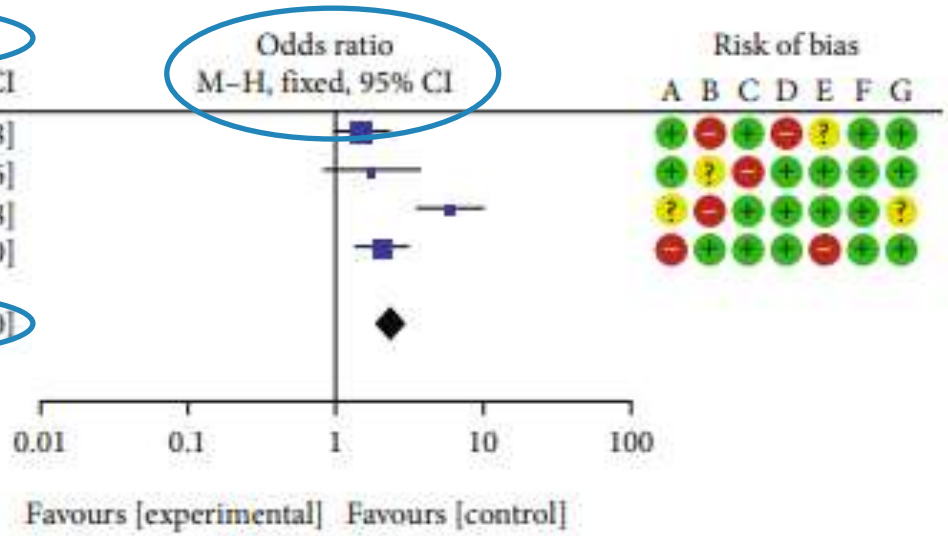
HE4敏感性分析

95% CI(1.87, 3.00)

Study or subgroup	Experimental group Events	Total	Control group Events	Total	Weight	Odds ratio M–H, fixed, 95% CI
Chen X 2017	98	178	80	178	39.4%	1.50 [0.99, 2.28]
Freydanck MK 2012	32	56	24	56	11.3%	1.78 [0.84, 3.76]
Simmons AR 2016	118	142	98	217	14.3%	5.97 [3.57, 9.98]
Tang T 2021	112	190	78	190	35.0%	2.06 [1.37, 3.10]
Total (95% CI)		566		641	100.0%	2.37 [1.87, 3.00]
Total events	360		280			

Heterogeneity: $Chi^2 = 18.03$, $df = 3$ ($P = 0.0004$); $I^2 = 83\%$

Test for overall effect: $Z = 7.19$ ($P < 0.00001$)



Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

FIGURE 9: Meta-analysis of diagnostic sensitivity analysis of HE4 between two groups.

結果

- 針對15篇研究，包括2262例經國家病理標準確診的卵巢癌患者，對照組包括2300例卵巢良性患者和健康人，共計4562例。觀察結果為:

1. **CA125、CA199、HE4具有較高的診斷效能，可提高卵巢癌診斷的敏感性和準確性，對診斷具有一定的臨床價值及鑑別診斷。**

2. **研究最終表達CA125、CA199、HE4特異性和敏感性在卵巢癌有明顯提高，但沒有提出聯合檢測的數據來提供結果說明(缺點)**

補充說明

Biomarkers and algorithms for diagnosis of ovarian cancer: CA125, HE4, RMI and ROMA, a review

Vincent Dochez[1*], Hélène Caillon[2], Edouard Vaucel[1], Jérôme Dimet[3], Norbert Winer[1] and Guillaume Ducarme[4]

- 血清 CA125 檢測在早期敏感性較低，在某些情況下（如月經或子宮內膜異位症）可能會增加。 HE4 的水平在卵巢腫瘤中過表達。 其特異性為 94%，其水平不受子宮內膜異位囊腫的影響。 CA125 和 HE4 的組合措施已被證明是高效的，曲線下面積 (AUC) 高達 0.96。 此外，這種 CA125 的聯合測量可以糾正 HE4 的變化，這些變化是由於吸煙或避孕結合雌激素和孕激素引起的。 雖然 RMI 的特異性有時達到 92%，但 0.86 的相當低的 AUC 並不能使其成為最佳診斷工具。 ROMA 的特異性低於 HE4（84% 比 94%）。 迄今為止，診斷卵巢癌最有效的生物診斷工具是 CA125 和 HE4 的結合。

討論

- 本科室尚未有HE4的檢驗項目，根據研究顯示HE4對卵巢癌的早期治療有很大的幫助，或許未來有機會引進此項目來幫助臨床醫師診斷卵巢癌。

- 也可運用ROMA算式提供卵巢癌診斷的參考。

卵巢惡性腫瘤風險評估因子(ROMA)

- 2009年，Moore提出了一種新的算法：（risk of ovarian malignancy algorithm,ROMA)，他根據停經期HE4和CA125水平，定義為缺乏月經或停經6個月。

1. 停經期前的病患
   ROMA值>=7.4% 有罹患卵巢上皮細胞癌的高風險
   ROMA值< 7.4% 有罹患卵巢上皮細胞癌的低風險

2. 停經期後的病患
   ROMA值>=25.3% 有罹患卵巢上皮細胞癌的高風險
   ROMA值< 25.3% 有罹患卵巢上皮細胞癌的低風險

感謝聆聽

Risk of malignancy index(RMI)

- 使用CA125，超音波評分及停經的條件，做出以下公式來區分良性或惡性腫瘤
- RMI = U x M x CA125 with U =ultrasound score
- U=0 if ultrasound score = 0
- U=1 if ultrasound score = 1
- U=3 if ultrasound score 2 to 5

- M=1 for pre-menopausal women
- M=3 for post-menopausal women

- score>200， strong association with a high risk of malignancy

RMI-2

- 根據超音波形態學標準評估腫塊：雙側性、實性區域、多房性、腹水和轉移。如果滿足0或1個標準，則超聲評分被指定為*U*=1，如果滿足2個或更多標準，則超聲評分*U=4*。計算總分。

- This value 0 maysometimes seem aberrant in the interpretation of a score.Nevertheless, the RMI score is based on a threshold (200).The calculation of specificity, sensitivity and positive andnegative predictive values is not based on the value of thescore but on its value below or above that threshold.