

# Meta-analysis (統合分析)

內科部研究中心 許秋婷 研究員

四期七樓 (分機: 4738)

# 前言

- 研究風氣盛行，以論文作為就業、考核、升遷(升等)的指標，論文發表數量節節升高。
- 閱讀論文的過程中，我們發現有很多研究主題都是類似的。
- 對於同樣的主題，論文之間有不同的結果(困擾)。
- 為了將這些具有相同主題的論文進行整理，估算出一個合併效果（pooling effect size）的統計方法稱為統合分析(Meta-analysis)。
- 整合分析(Meta-analysis)為「證據導向」的研究型態，是量化研究的另一顯學。
- 統合分析是要做已發表文獻的summary data的統合分析。  
非Pooled-reanalysis，要和各論文的研究者聯繫，請他們提供原始資料。





# 新手上路

- 末代健保資料庫研究
- 沒有大量累計的病人，長期追蹤資料
- 沒有研究資源支持臨床基礎實驗

*在不同階段，擁有不同資源時*

*有其適合做的研究類型及努力方向*



## 新手上路

- 入門門檻較低(\$\$)，無資源研究生涯的起步
- 常常閱讀最新期刊，保持對臨床特定議題的敏感度與關注
- 高臨床應用度
- 文獻搜尋、評讀技巧是所需必備的能力

*在不同階段，擁有不同資源時*

*有其適合做的研究類型及努力方向*

# A hierarchy of evidence



# 臨床研究的實驗設計

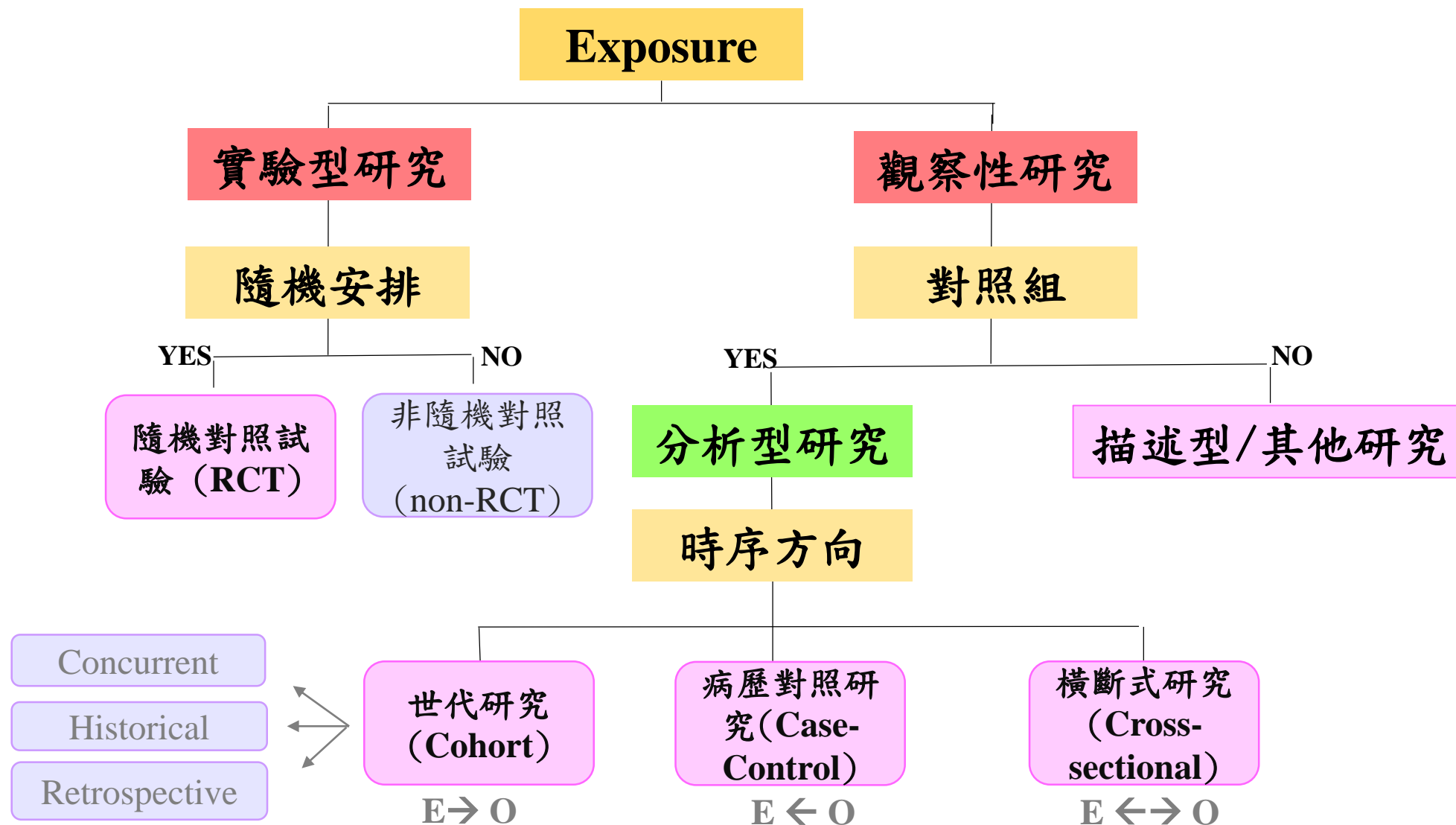
Exposure (Independent Variable)

干預，治療，藥物，檢驗數值...



Outcome (Dependent Variable)

疾病、復發(or復原)、惡化(or緩解)、死亡...



## 形成一個可回答的臨床問題

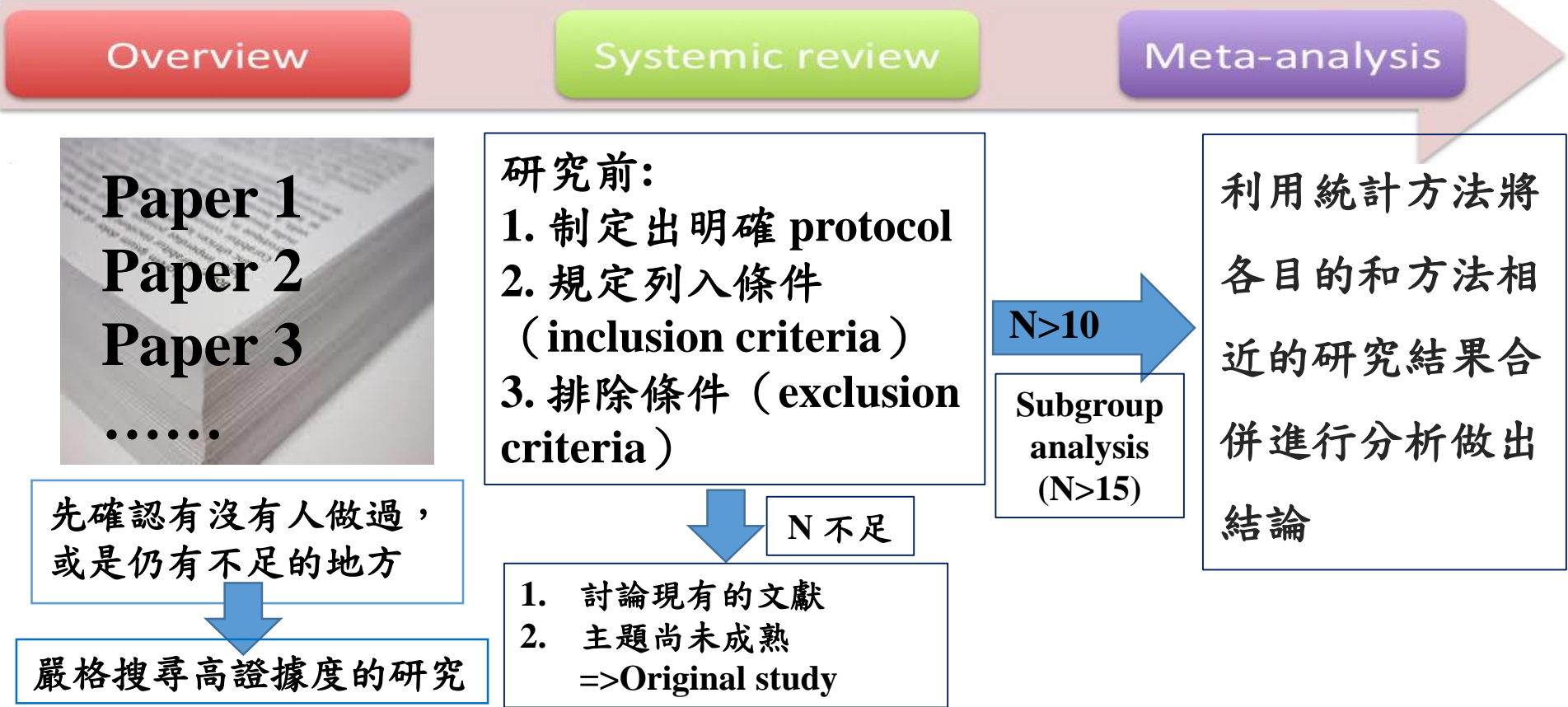
**Introduction: why/ what's the importance to conduct this review**

- **Pulling together the evidence**
  - **Results from individual studies are not consistent (controversial).**
  - **Many of the results have been conflicting and have brought confusion to the direction of further study.**
- **Using appropriate method and model, in order to explore and allow for all important sources of variation.**
- **Reducing the random errors in assessment of intervention.**

# Systematic Reviews and Meta-Analyses



某一臨床問題



對臨床問題做出結論

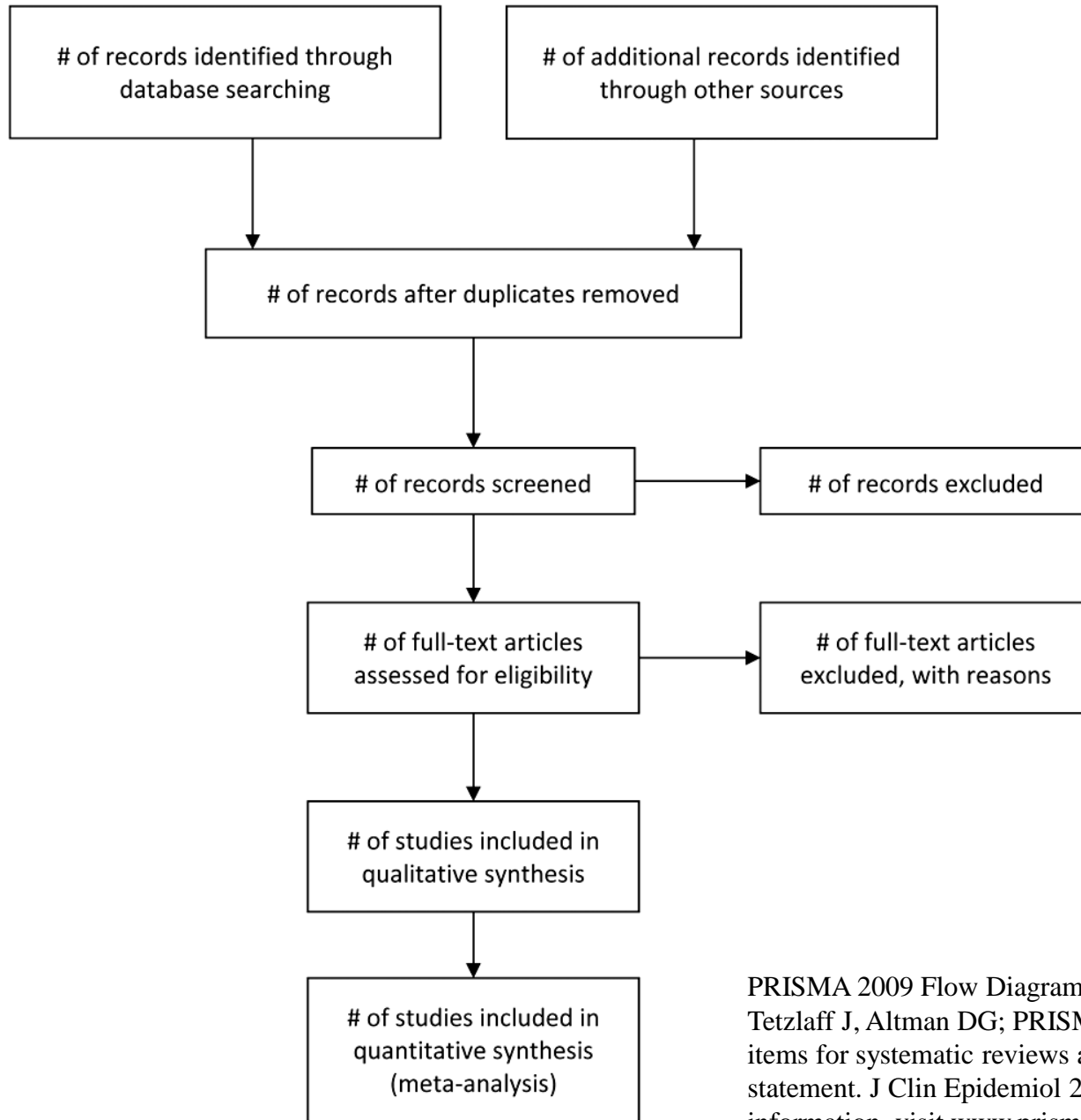


Identification

Screening

Eligibility

Included



研究議題事先定義好:

先定義好主題，關鍵字，才能夠找出真正有相關性的論文。

搜尋論文時要減少漏抓的疏失:

所有的相關資料庫 (database) 都要搜尋，才不會遺漏重要論文。

列入條件和排除條件事先定義好:

作任何研究一定要定義出受試者條件或研究的論文的列入和排除條件，這個流程可以減少異質性 (heterogeneity)。

PRISMA 2009 Flow Diagram (From Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009;62:1006-12, For more information, visit [www.prisma-statement.org](http://www.prisma-statement.org)).

## 步驟二：尋找最佳的實證文獻證據

### Identification:

- Electronic bibliographic databases:  
PubMed, Google scholar, UpToDate, DynaMedScopus, Cochrane Collaboration Central, Register of Controlled Clinical Trials, Cochrane Systematic Reviews, and ClinicalTrials.gov
- Literature written in English(?) or available in full texts (abstracts).
- Keywords and published period  
(?Disease, ?drug, ?Outcome, comparison group, RCT, cohort ...)

可參考 PICO 標準

# 清楚地描述挑選文獻的理由

## PICO 臨床問題模型

- P: patient, population  
這個病人或族群的年齡層、體質、疾病史、過去史
- I: intervention  
治療、診斷、預後、成本效益分析等，研究的主題
- C: comparison  
intervention A 與 B 比較、或與不做比較
- O: outcome  
預後、或與目前問題有關的比較基準

# Association between obesity and kidney disease: A systematic review and meta-analysis

Y Wang<sup>1</sup>, X Chen<sup>1</sup>, Y Song<sup>2</sup>, B Caballero<sup>1</sup> and LJ Cheskin<sup>1</sup>

<sup>1</sup>Center for Human Nutrition, Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA and <sup>2</sup>Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

- **過去不足的地方**: Previous reviews of the association between obesity and KD have either focused on a single disease or have not included recent large prospective studies.
- **現在的目的**: This study aimed to comprehensively assess epidemiologic evidence on the relation between obesity and KD outcomes.
- **文章架構**: We first performed a systematic review of studies that have examined the association in adults, and then conducted a meta-analysis based on cohort studies conducted in the *general population*.

# Association between obesity and kidney disease: A systematic review and meta-analysis

Y Wang<sup>1</sup>, X Chen<sup>1</sup>, Y Song<sup>2</sup>, B Caballero<sup>1</sup> and LJ Cheskin<sup>1</sup>

<sup>1</sup>Center for Human Nutrition, Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA and <sup>2</sup>Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

- **Keywords :** Obesity, overweight, body mass index, kidney disease, chronic kidney disease, CKD, ESRD, kidney stones, kidney cancer, and RCC.
- Our PICO: P=general population. I= Obesity, C=Comparison, O=all kidney disease
- This literature search was carried out independently by 2 academic investigators. Group discussion resolved any disagreement with article selection.

# 清楚地描述挑選文獻的理由

Epidemiology. 14(4):479-487, JUL 2003

DOI: 10.1097/01.EDE.0000071413.55296.c4, PMID: 12843775

Issn Print: 1044-3983

Publication Date: 2003/07/01

 Share

 Print

## Lifestyle Factors, Obesity and the Risk of Chronic Kidney Disease

Bénédicte Stengel; Michelle E. Tarver-Carr; Neil R. Powe; Mark S. Eberhardt; Frederick L. Brancati

	<b>Systematic review</b>	<b>Our PICO</b>
<b>P</b>		<input type="checkbox"/> YES <input type="checkbox"/> NO
<b>I</b>		<input type="checkbox"/> YES <input type="checkbox"/> NO
<b>C</b>		<input type="checkbox"/> YES <input type="checkbox"/> NO
<b>O</b>		<input type="checkbox"/> YES <input type="checkbox"/> NO


# 關於止血帶用於關節鏡 ACL 重建手術之統合分析

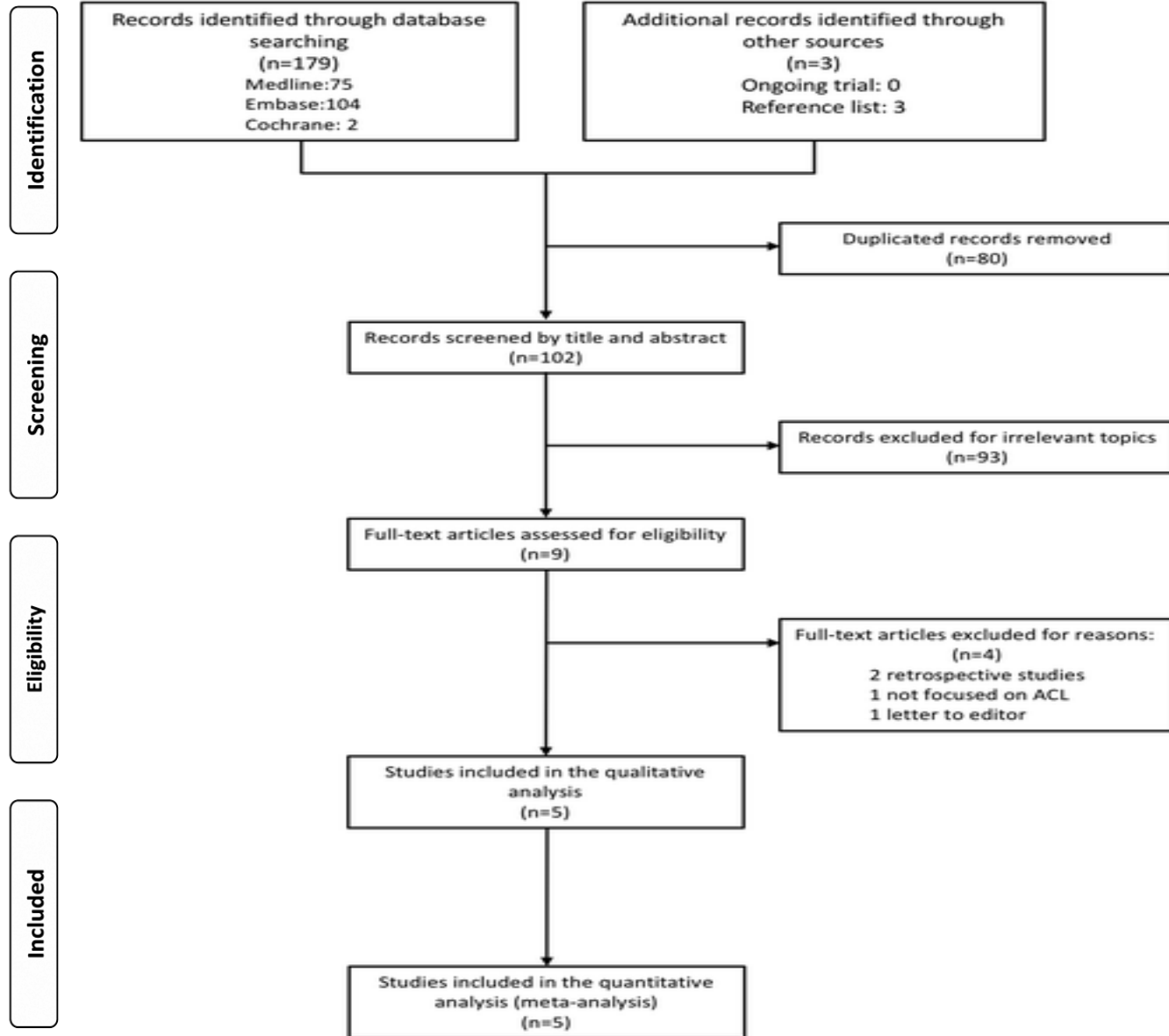
## RESEARCH ARTICLE

## Open Access



### Tourniquet use in arthroscopic anterior cruciate ligament reconstruction: a systematic review and meta-analysis of randomised controlled trials

Liang-Tseng Kuo<sup>1,2,3</sup>, Pei-An Yu<sup>1</sup>, Chi-Lung Chen<sup>1</sup>, Wei-Hsiu Hsu<sup>1,4\*†</sup> and Ching-Chi Chi<sup>2,4,5\*†</sup> 



## PICO 標準

## 步驟三：嚴格評讀文獻

## Qualitative analysis

## Assessment of risk of bias

「研究設計」 「研究執行」 「研究結果」 三大要素



# Qualitative analysis 文獻評讀工具

## Comparison of CATs

▷ 隨機分派研究 (RCT) 常用評讀工具比較:

	CASP	Oxford CEBM	Cochrane RoB
隨機分排	受試者是否被隨機分派?	受試者是否被隨機分派?	如何產生隨機分派順序?
分派隱匿	-	-	為何確保分派隱匿?
維持盲性	受試者、健康照護者、研究人員是否維持盲性?	量測結果是否客觀? 受試者與醫師是否維持盲性?	受試者、照護者、評估者是否維持盲性?
分派結果	研究開始時, 組間是否相仿?	研究開始時, 組間是否相仿?	-
相同對待	除實驗介入, 是否接受相同對待?	除實驗介入, 是否接受相同對待?	-
意向分析	納入研究受試者是否適當的代表結論 (研究結果)?	是否分析納入的所有受試者? 按分派組別分析?	數據是否完整?
相同對待	除實驗介入, 是否接受相同對待?	除實驗介入, 是否接受相同對待?	-
開誠布公	-	-	是否選擇性報告 (數據、資料)?

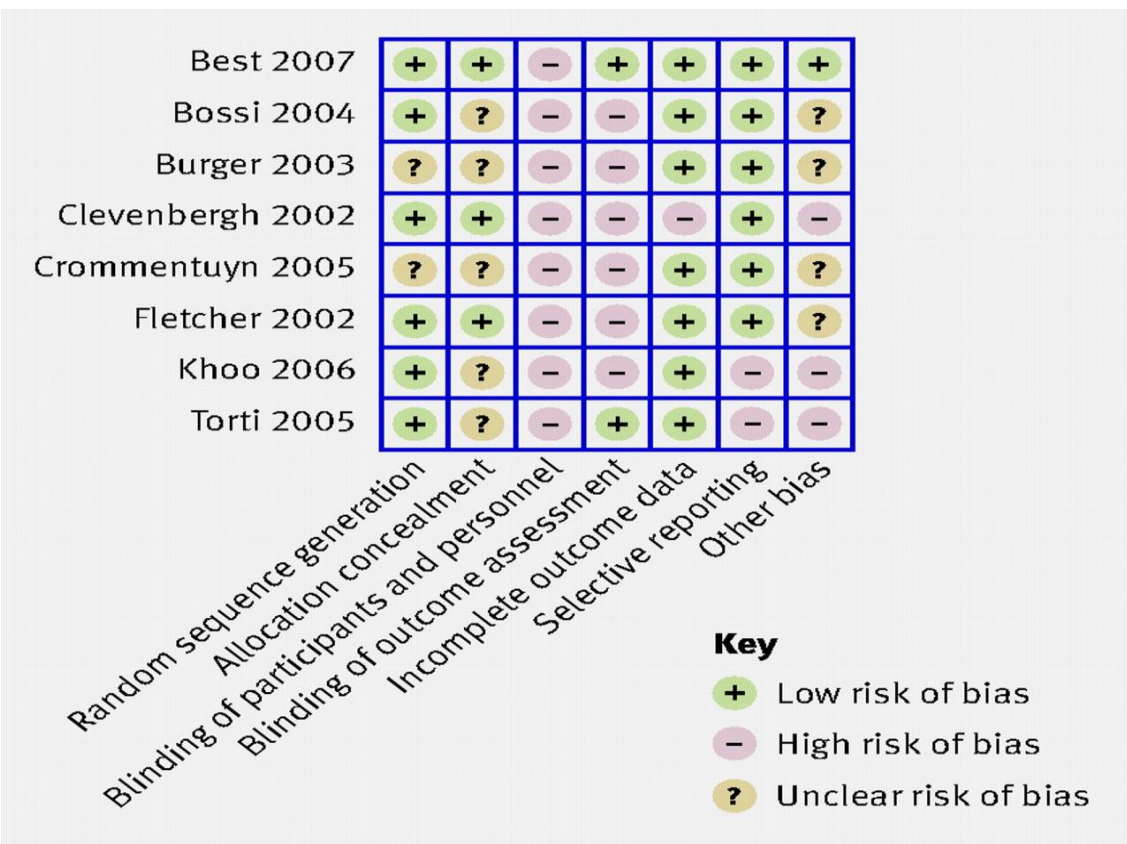




# Cochrane 風險評估工具

The Cochrane Collaboration's tool for assessing risk of bias in randomised trials  
BMJ 2011; 343 doi: <https://doi.org/10.1136/bmj.d5928> (Published 18 October 2011)  
Cite this as: BMJ 2011;343:d5928

偏誤風險	定義	項目
選擇性偏誤	組間比較有「系統性差異」	如何產生隨機分派順序? 如何維持分派隱匿?
表現性偏誤	實驗性介入外，接受的照顧、待遇有「系統性差異」	受試者、照護者是否維持盲性?
偵測性偏誤	結果量測有「系統性差異」	結果評估者否維持盲性?
削弱性偏誤	組間退出試驗有「系統性差異」	數據是否完整?
報告性偏誤	報告與未報告項目有「系統性差異」	是否選擇性報告?



## CASP CHECKLISTS

<http://www.casp-uk.net/casp-tools-checklists>

不同研究設計有不同的評讀工具

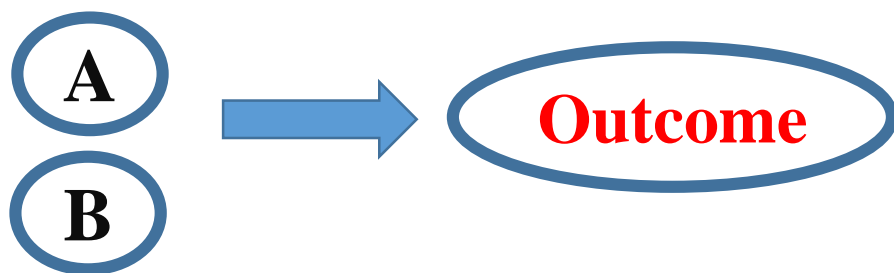
<u>CASP Systematic Review Checklist</u>	<u>CASP Qualitative Checklist</u>
<u>CASP Randomised Controlled Trial Checklist</u>	<u>CASP Case Control Checklist</u>
<u>CASP Diagnostic Checklist</u>	<u>CASP Cohort Study Checklist</u>
<u>CASP Economic Evaluation Checklist</u>	<u>CASP Clinical Prediction Rule Checklist</u>

# 個案納入統計分析的準則

病患有接受完整追蹤嗎？而且分析時有按照病患原來安排的治療方式來分析嗎？

## Intention-to-treat Model (ITT 意向分析)

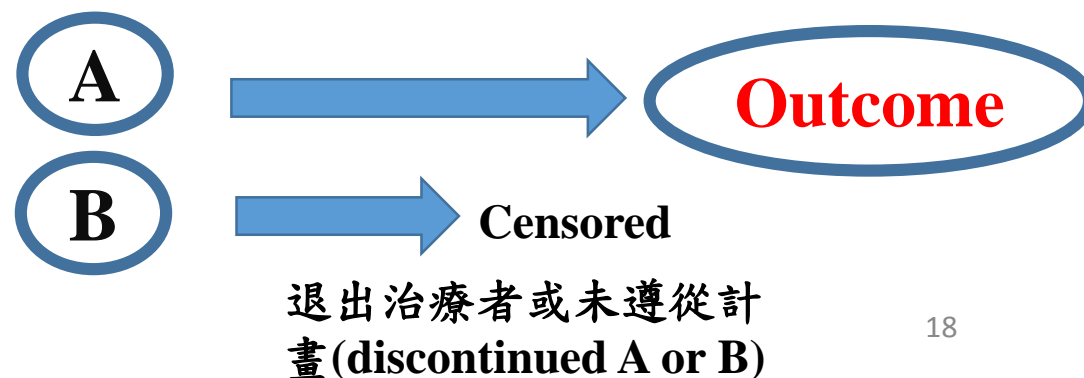
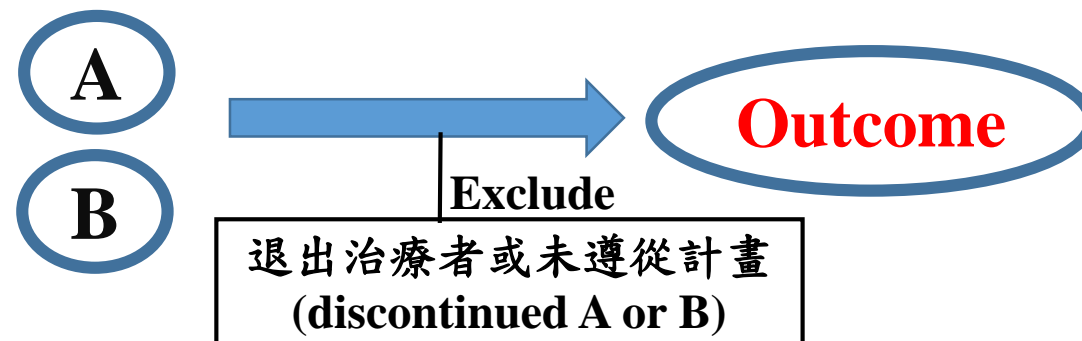
受試者一旦被隨機分組後，不論病人是不是遵從計畫書（可能中途停藥？更換治療？沒有完成療程），受試者組別會完全按照最初分組結果進行分析。



## As-treat Model (AT model 根據治療分析)

## Per-Protocol Model (PP)

排除中途退出，未遵從計畫書者。受試者需符合計畫書進行分析。



# Definition of outcome of the meta-analysis :

- **Binary Outcome:**
  - **Non-comparative:**
    - Odds=(Number of patient having event)/(Number of patient not having event)
    - Incidence rates= (Number of patient having event)/(Total follow-up time)
  - **Comparative**
    - Odds ratio=(Odds for group A)/(Odds for group B)
    - Relative Risk=P(event occurs in group A )/P(event occurs in group B)
    - Hazard ratio
- **Continuous outcomes:**
  - Mean difference (MD): 相同量化工具
  - Standardized mean difference (SMD): 不同量化工具
- **Others:** Number needed to treat (NNT), Number needed to harm(NNH)

# Data extraction

對收集到的論文要做論文描述性的表格，好讓讀者可以清楚這些論文的特點。

Standardized data extraction form:

1. First author, year of publication,
2. Study design (patient selection and concealment),
3. Sample size
4. Participant characteristics (e.g. age, gender, 人種，區域別 ..etc.),
5. Use of intervention/ treatment/ exposure (方式、種類、時機、劑量 and 時間.....),
6. Outcome data (subjective outcome, number of event)

沒有特定格式，詳列越多越好，後續執行subgroup analysis或meta-regression的依據

醫師的工作暫時到這裡，接下來就是統計分析。  
你以為這樣就結束了嗎？還是要知道怎樣解讀統計。

# Heterogeneity of study: Test for presence of heterogeneity

目的是要看個別論文的效果和綜合性效果是否有極大的差異性。

虛無假設：所有的研究都在評估同一個效果

## Cochrane's Q

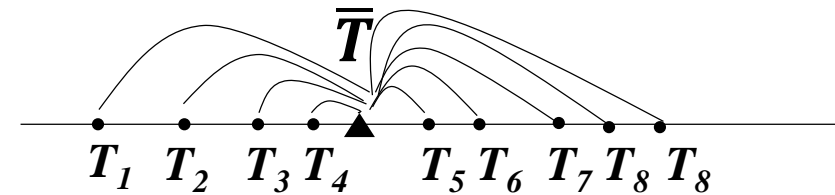
$$\sum_{i=1}^k w_i (T_i - \bar{T})^2 \sim \text{Chi-square } (k-1)$$

where  $k$  is number of studies being combined,  
 $T_i$  is the study effect in the  $i$ -study.

$$\bar{T} = \frac{\sum_{i=1}^k w_i T_i}{\sum_{i=1}^k w_i}$$

Weight(權重):  $w_i = 1/\text{Variance of } T_i$

當用Q-test 來檢定異質性時，因為檢力不足，所以這邊的統計臨界值定成  $p\text{-value} < 0.10$ ，而不是平常熟悉的0.05。



Test whether the individual effect are farther away from the common effect, beyond what is expected by chance.

# I<sup>2</sup> test :

$$I^2 = \left[ \frac{Q - (K - 1)}{Q} \right] \times 100$$

研究間的變異性(between-study variance)有多少  
比例來自於異質性

統計小學堂:

$Q \sim \text{Chi-square}(k-1)$

$= \text{gamma}((k-1)/2, 2)$

$E(Q) = k-1$

$Q - (k-1)$  is unbiased estimator of 0

I<sup>2</sup> 值介於0%~100% ,

I<sup>2</sup> 為 <25% (低度異質性) ,

I<sup>2</sup> 為 25%- 50% (中度異質性) ,

I<sup>2</sup> 為 >75% (高度異質性) ,

好不容易看完所有文獻，文獻的品質符合要求。

**理想狀況**: 各文獻之研究組成、資料結果沒有太大差異的話(異質性低)，才會進行meta-analysis。

**主張 1**: I<sup>2</sup> test >50% 異質性大，不宜統合

**主張 2**: 仍然可以做meta-analysis，但是研究者們必須進行更多分析，找出解釋 I<sup>2</sup> 數值這麼大的原因 (Subgroup analysis)

Meta-analysis: Can We Mix Apples and Oranges?

Paul Moayyedi, Ph.D., F.R.C.P.



# Possible cause of heterogeneity:

- **Clinical Heterogeneity**

- ✓ 不同的病人群研究 (Age dist., diagnosis, severity, in- and exclude criteria ...)
- ✓ 治療方法 Interventions (type, dose, duration,...)
- ✓ 附加治療 Co-interventions
- ✓ 結果評估方式 Outcomes measured (type, scale, cut-off point, follow-up time)

- **Methodological Heterogeneity**

- ✓ 研究設計不同 ( Different study design features )
- ✓ 統計分析的準則 (Intention-to-treat, Per-Protocol, As-treat analysis)
- ✓ 研究品質 ( Study quality )
- ✓ 隨機誤差 ( Random error )



# How to deal with heterogeneity:

## 臨床異質性：

- 異質性多來自於臨床，最初沒有注意。可由收錄文獻的條件(PICO)去找出可能的原因。

## 統計學異質性

- 在滿足臨床同質性的前提下，仍存在有統計學異質性( $I^2$  test)。
- 檢查那些論文造成了異質性的產生，必要時可刪除某些論文(但是要有充分理由)。

不管最終統計能不能確定異質性，還是要回歸文獻本身，研究者應從臨床角度解釋異質性的發生。

## 應對措施

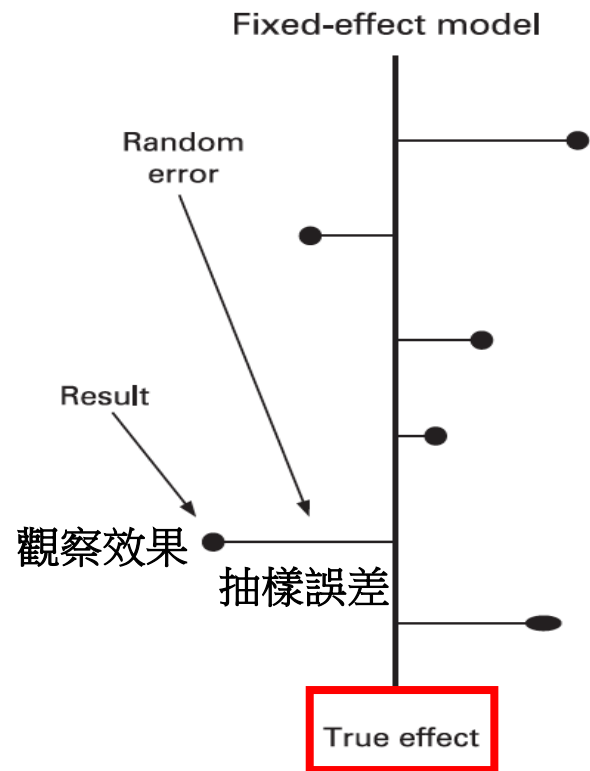
- ➔ Exclude criteria
- ➔ Analysis group of studies separately

## 應對措施

- ➔ 無法解釋，放棄meta-analysis，只做systematic reviews
- ➔ Exclude Studies (充分理由)
- ➔ Random-effect model (默許研究間存在變異或尚有未被發掘因素)
- ➔ Subgroup analysis and meta-regression (根據懷疑的因素做迴歸分析，看加入這個因素後，能不能降低異質性)
- ➔ Sensitivity analysis

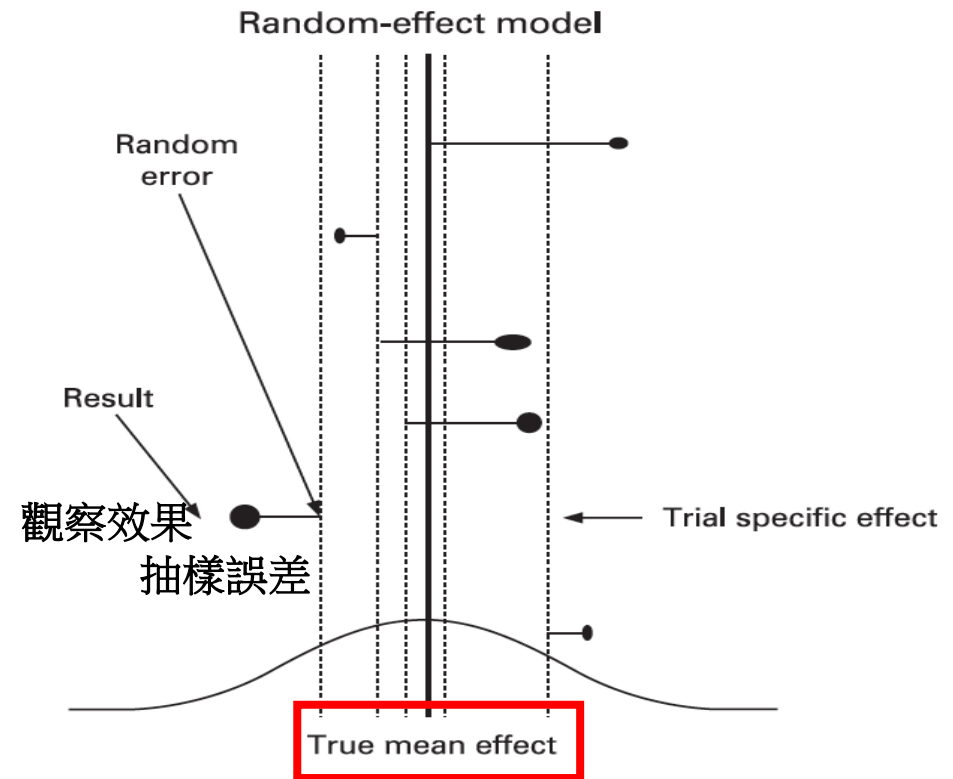


# 統合分析的統計方法 fixed-/random- effects model



假設:每篇研究都有一個共通的真實效果

觀察效果=真實效果(true effect) +  
研究內變異量(within-study variability)  
(sampling error)



假設:每篇研究的真實效果都不同

觀察效果=實際效果平均數(true mean effect)+  
研究內變異量(within-study variability)+  
研究間變異量(between-study variability)

隨機式模  
型的變異  
量<sup>25</sup>

# Combining Results from Multiple Studies

$$\bar{T} = \frac{\sum_{i=1}^k w_i T_i}{\sum_{i=1}^k w_i}$$

$k$  is number of studies being combined,  $T_i$  is the study effect in the  $i$ -study.

**Weight(權重):**  $w_i = 1/\text{Variance of } T_i$  *for fixed-effect model*  
 $= 1/(\text{Variance of } T_i + \tau^2)$  *for random-effect model*  
 $\tau^2$  : *between-study variability*

**Fixed effect model:**

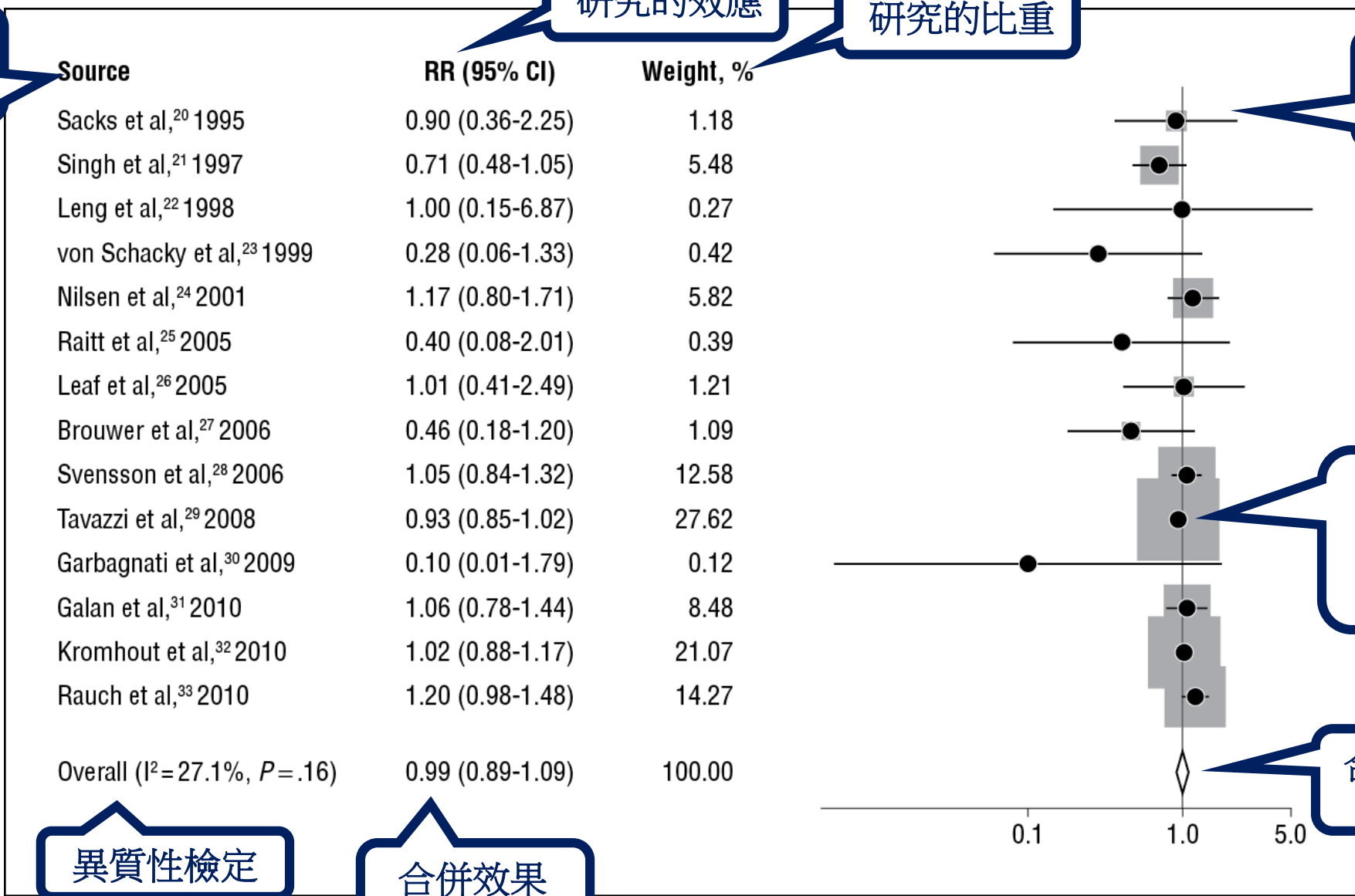
- 某些文章的樣本數大，result 的 variability 小，佔的比重大。
- 凸顯某些文章的重要性

**Random effect model:**

- 權重較 fixed-effect model 小 → 較為保守。
- 沒有凸顯某些文章的重要性

# Forest plot 森林圖：概括結果呈現

文獻第一作者，出刊年



研究的效應

研究的比重

各研究的效果和95% CI

灰階越大，該研究在計算合併效果所佔比重就大

合併效果 (圖示以菱形表示)

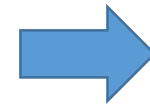
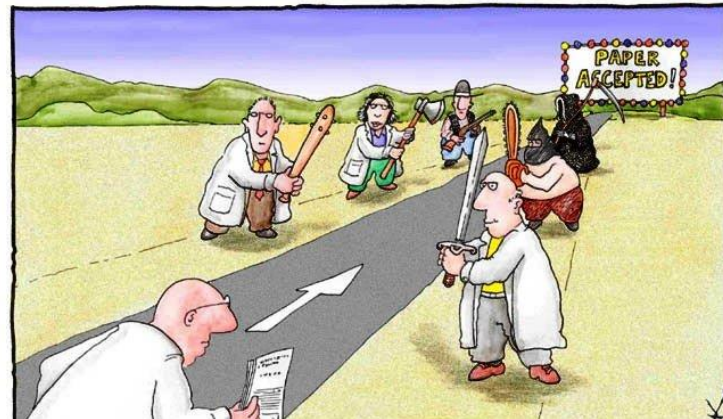
異質性檢定

合併效果  
(fixed effect)

合併效果(pooling effect size)去解釋研究主題的統整結果

# 統合分析(Meta-analysis)：出版偏誤(publication bias)問題

- 隱憂：收集的文獻中可能有出版偏誤(publication bias)而導致結果並不可信。
- 出版偏誤(publication bias):
  - 研究人員：傾向發表正向效果的文章，而負向效果的文章則不加以發表。
  - 期刊的編輯：傾向接受有正向效果的文章，而負向效果或無明顯效果的文章不易為人所知。
  - 收集文獻時，負向的研究結果不易被發現。

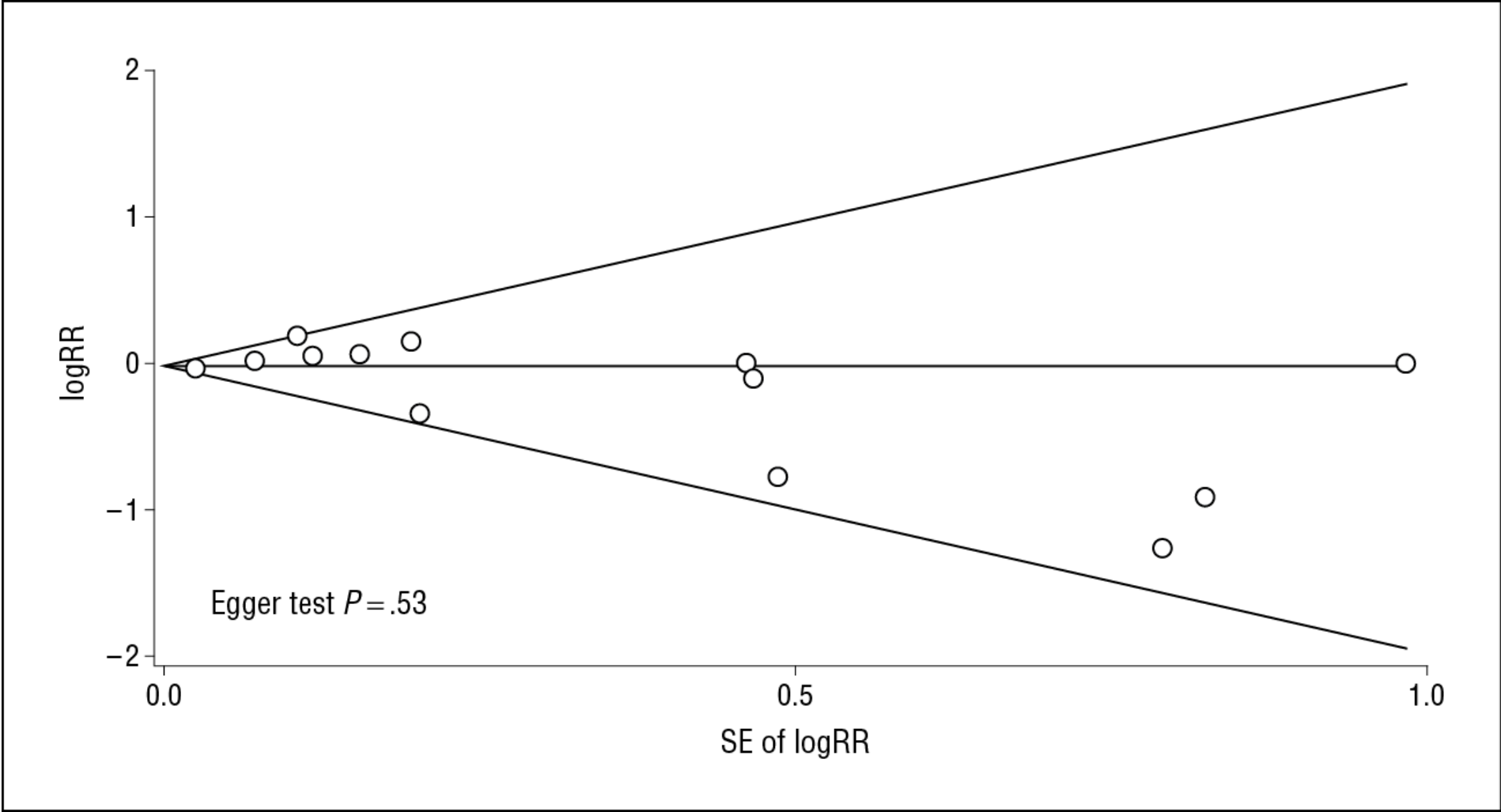


# 出版偏誤的判斷

- 視覺化：漏斗圖(funnel plot)呈現對稱與否
- 客觀的： Egger's test
- 修正合併效果：

以funnel plot 為基礎的理論架構，利用無母數遞迴的估計方法( iterative non-parametric rank-based estimation procedure or called , “Trim and Fill” Method ) ，對未發掘的文獻進行插補(imputation)

# Begg funnel plot and Egger test for identifying publication bias

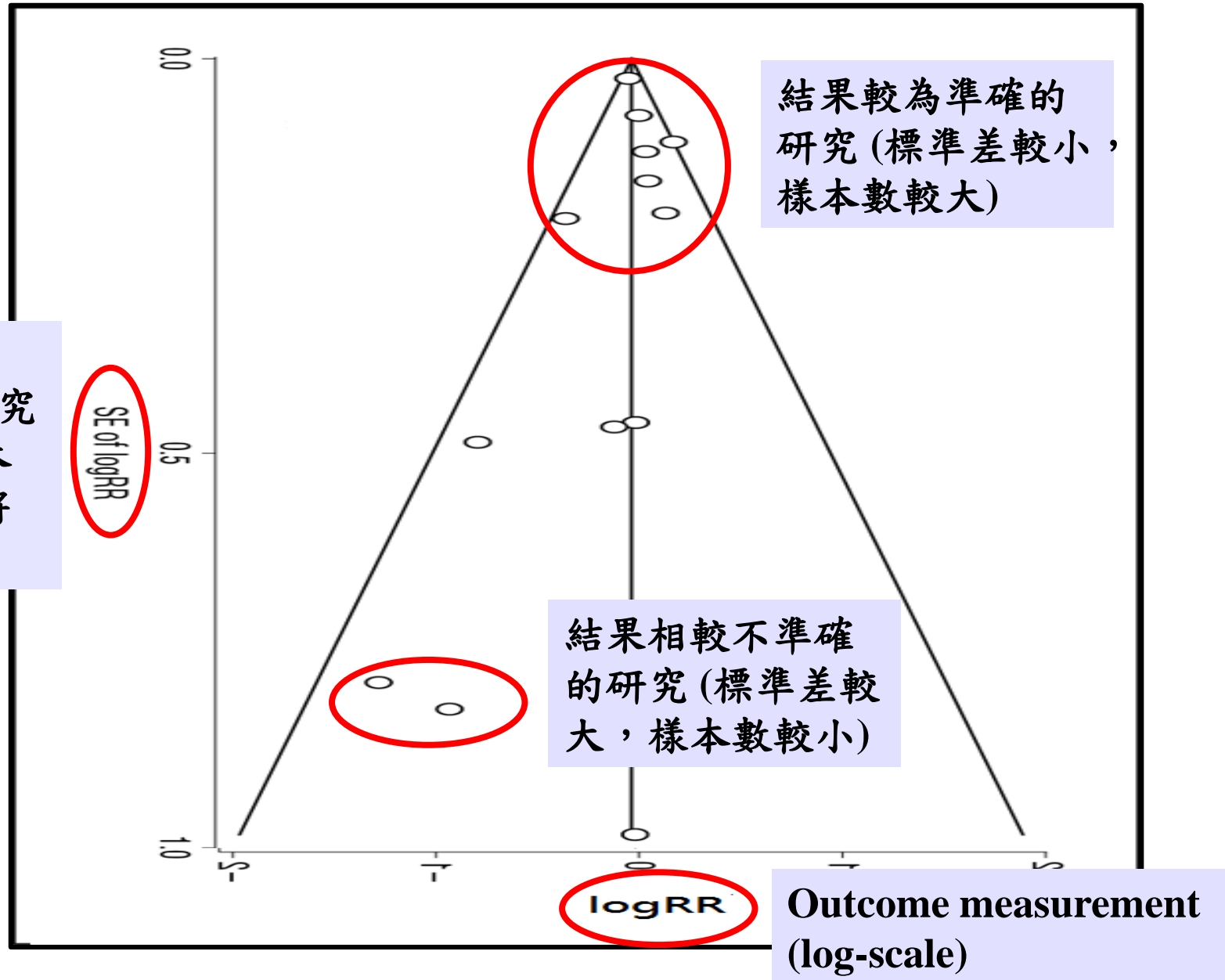


Kwak, Sang Mi, et al. "Efficacy of omega-3 fatty acid supplements (eicosapentaenoic acid and docosahexaenoic acid) in the secondary prevention of cardiovascular disease: a meta-analysis of randomized, double-blind, placebo-controlled trials." *Archives of internal medicine* 172.9 (2012): 686-694.

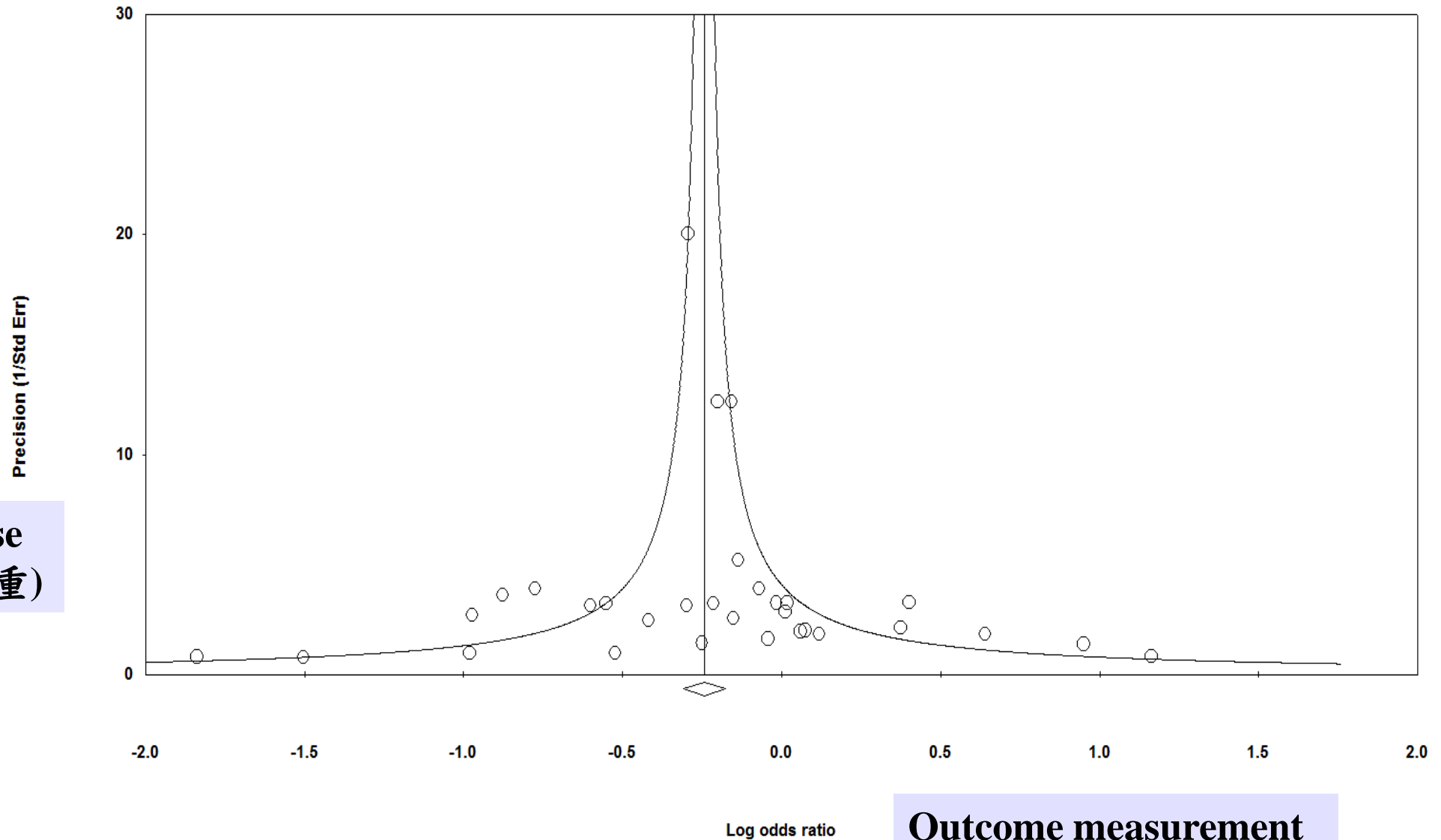
# Begg funnel plot and Egger test for identifying publication bias

所有研究結果的散佈圖(scatter plot)

標準誤(standard error), 代表單一研究結果的精確度, 樣本數目越大, 標準誤將會越小。



# Funnel Plot Precision by Log odds ratio



精確度 precise  
≈ Weight(權重)

Outcome measurement  
(log-scale)



# Egger's test: 檢定發表性偏誤是否存在

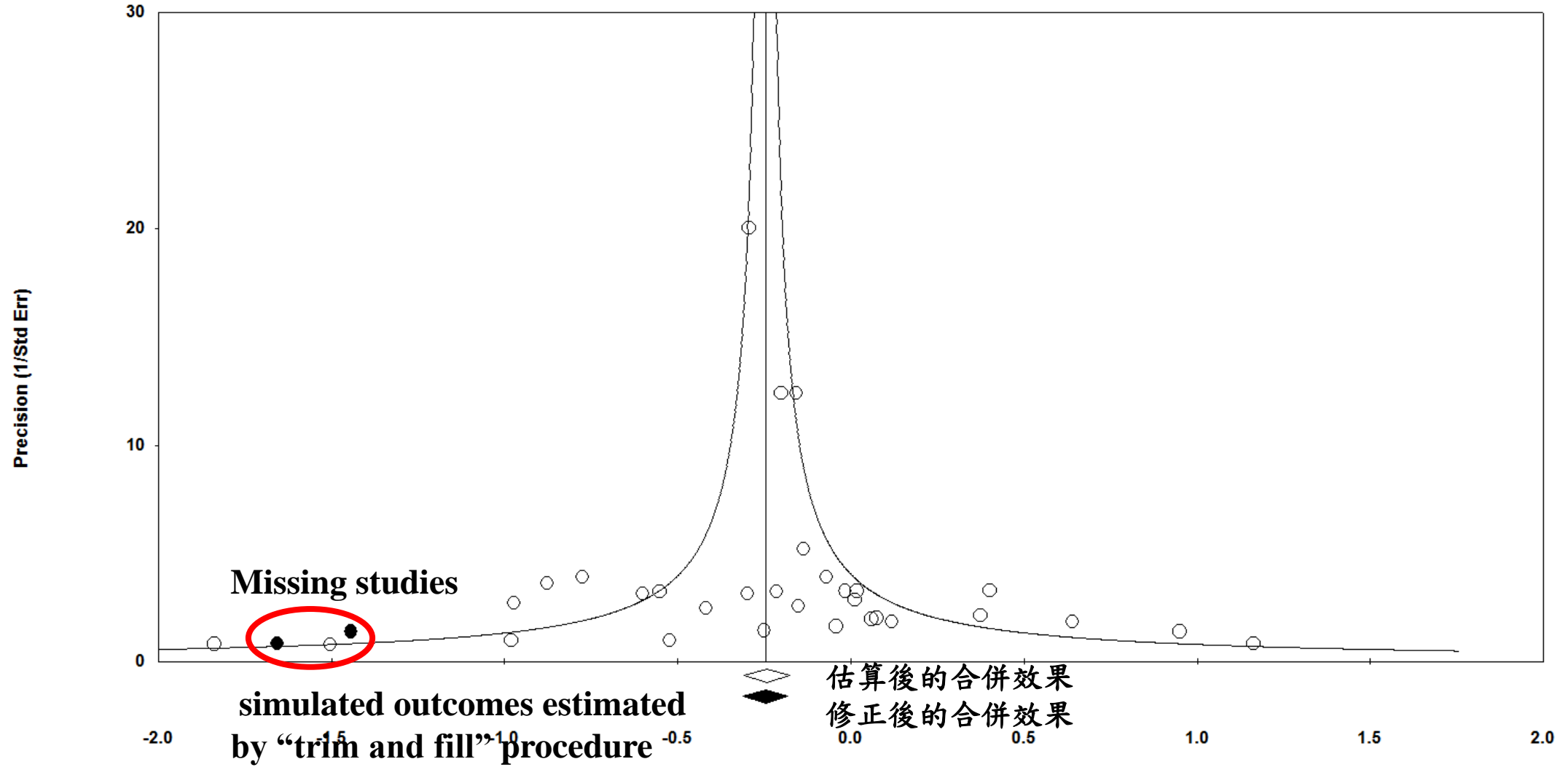
Egger's regression intercept	
Intercept	0.12533
Standard error	0.28024
95% lower limit (2-tailed)	-0.44782
95% upper limit (2-tailed)	0.69849
t-value	0.44724
df	29.00000
P-value (1-tailed)	0.32901
P-value (2-tailed)	0.65803

**H0: intercept = 0 (symmetric, no publication bias)**

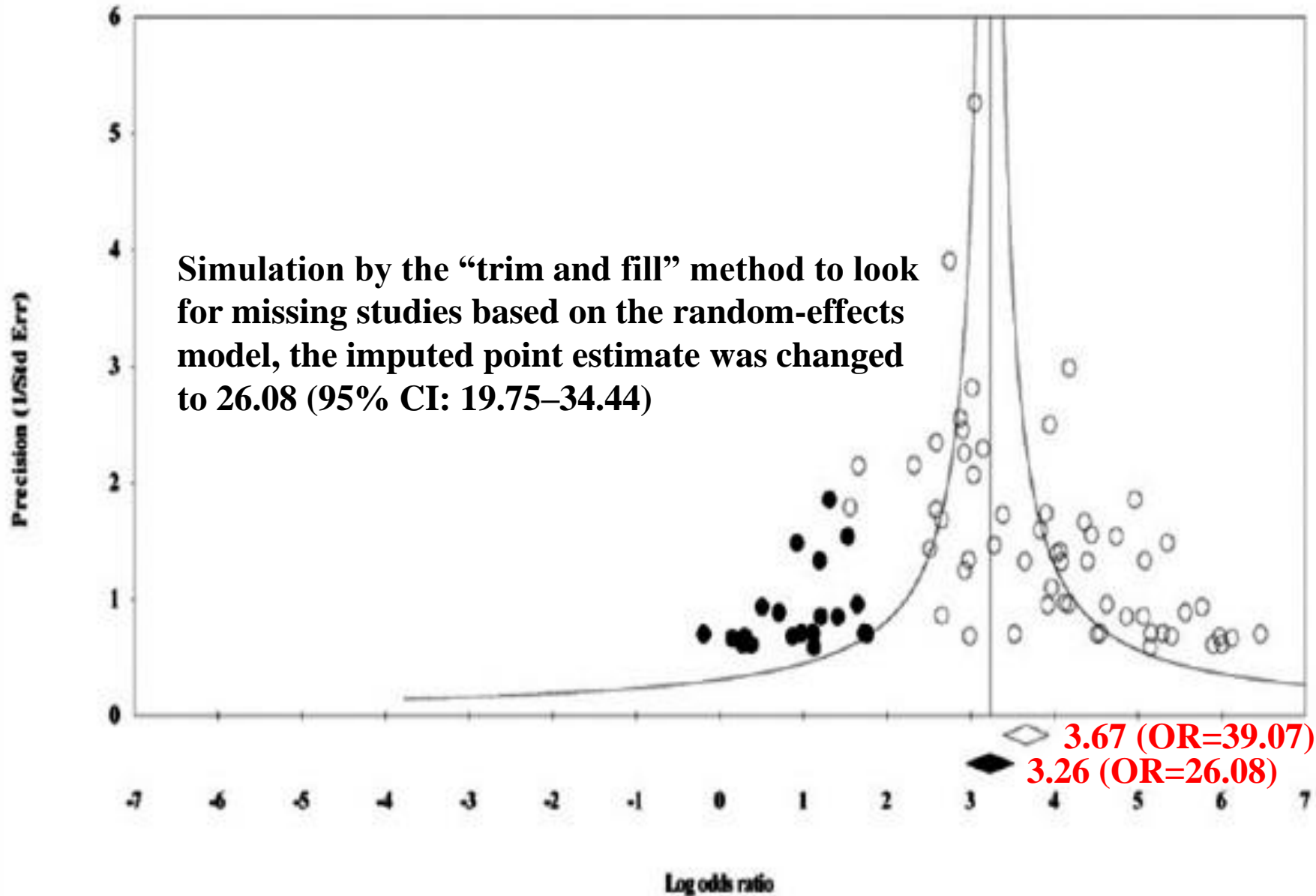
**vs**

**H1: intercept  $\neq$  0 (asymmetric, evidence of publication bias)**

# “Trim and Fill” Method



## Funnel Plot of Precision by Log odds ratio



Funnel plots for DOR showing the distribution of published study outcomes (open circles) and simulated outcomes (black circles) estimated by “trim and fill” procedure.

Geng, Xiao-Xia, et al. "Transient elastography in clinical detection of liver cirrhosis: A systematic review and meta-analysis." *Saudi journal of gastroenterology: official journal of the Saudi Gastroenterology Association* 22.4 (2016): 294.

# Sensitivity analysis

- CMA 中，將逐步將某篇論文刪除後，看看剩餘論文的合併效果是否會因此更改變，藉以測試合併效果的穩定度

# Meta-analysis的陷阱

- 單一（or少數的）研究的結果來下結論是一種比較危險的行為，萬一這些結果有隨機錯誤時（error by chance），我們就有可能對某個醫學議題造成誤判。
- 統合分析結果的可信度是取決於所選取論文的研究品質。
- 如果所選取讀論文品質是差的，那麼統合分析的品質當然也就差（garbage in, garbage out）。
- 我們在看任何一篇論文都應該要有能力來分析出該篇論文的品質好壞，而非只是無條件的全然接受論文的研究結果

*Thanks for your attention!!*