

# 文獻評讀

吳貞蓉

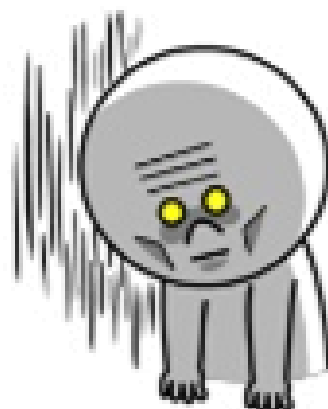


# 實證醫學步驟

---

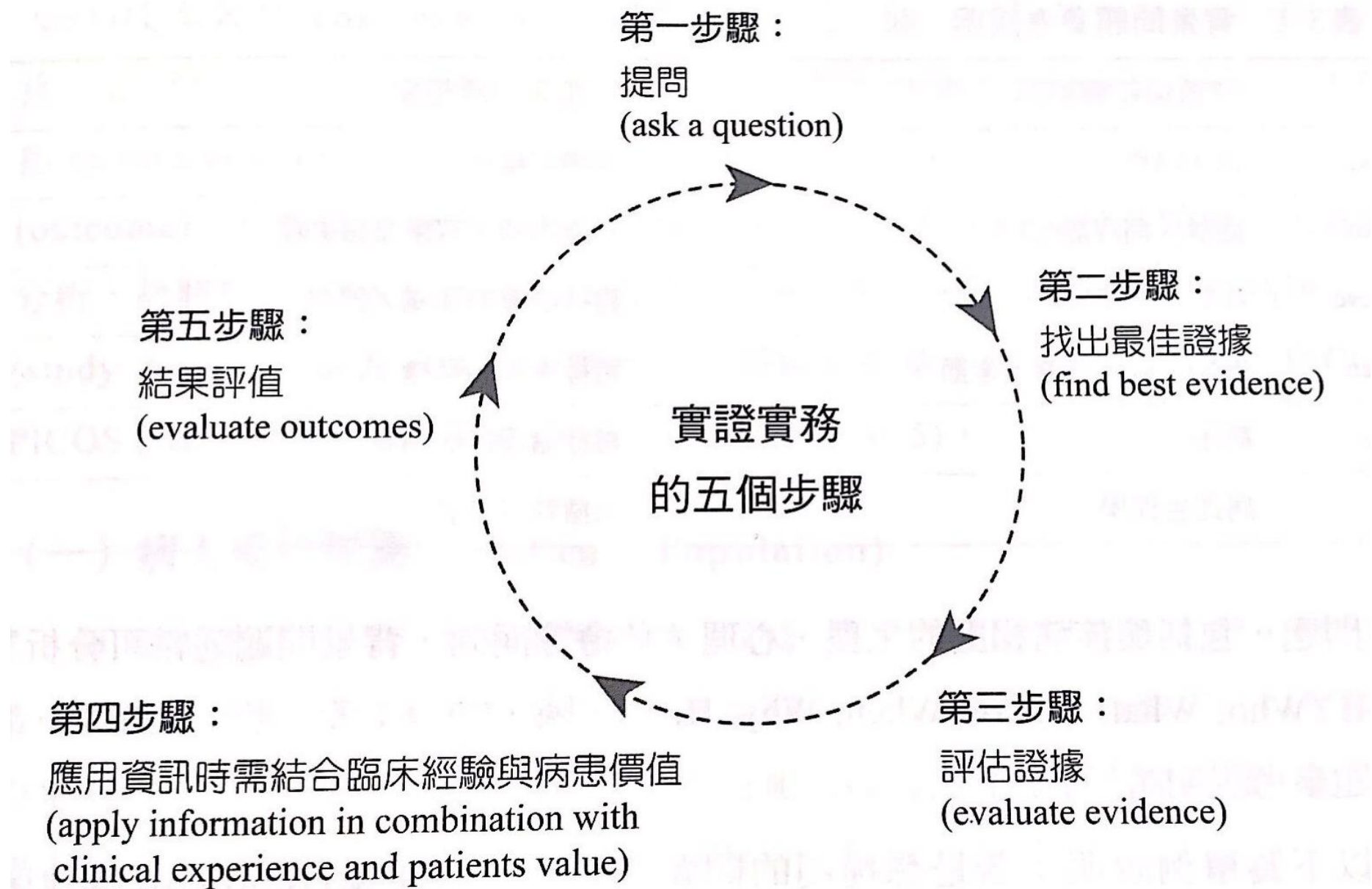
## ▶ 實證醫學步驟 - 從問題到應用

提出問題 → 將臨床問題改寫為「PICO」 → 檢索資料庫/挑選文獻 → 嚴格評讀文獻 → 決定證據等級/做出建議



估價

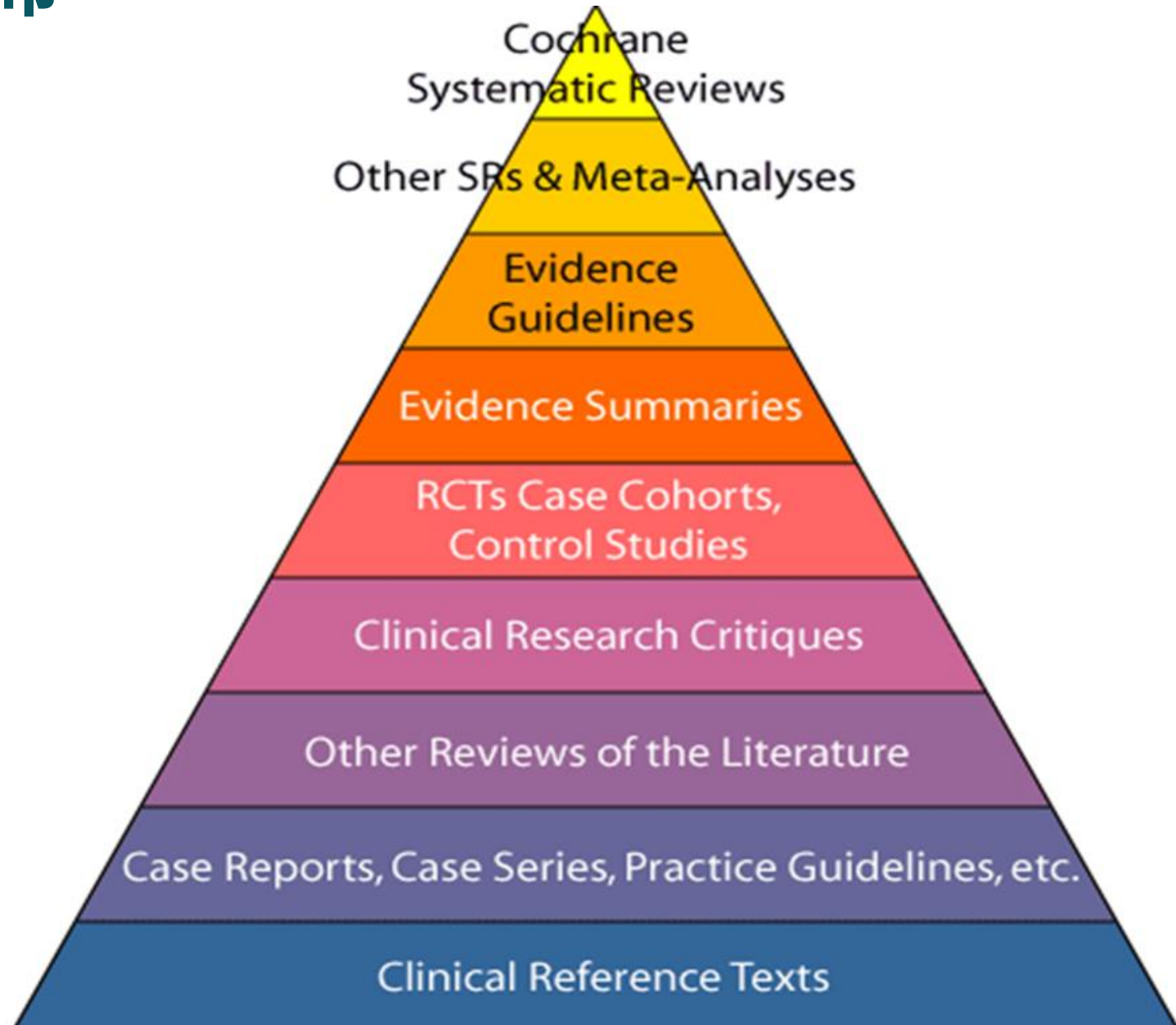
Appraise



○ 圖3-1 不斷循環的實證五大步驟

資料來源：Johnson, 2008

# 證據等級分佈



# Oxford CEBM

## Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence

Question	Step 1 (Level 1*)	Step 2 (Level 2*)	Step 3 (Level 3*)	Step 4 (Level 4*)	Step 5 (Level 5)
<b>How common is the problem?</b>	Local and current random sample surveys (or censuses)	Systematic review of surveys that allow matching to local circumstances**	Local non-random sample**	Case-series**	n/a
<b>Is this diagnostic or monitoring test accurate?</b> (Diagnosis)	Systematic review of cross sectional studies with consistently applied reference standard and blinding	Individual cross sectional studies with consistently applied reference standard and blinding	Non-consecutive studies, or studies without consistently applied reference standards**	Case-control studies, or "poor or non-independent reference standard**	Mechanism-based reasoning
<b>What will happen if we do not add a therapy?</b> (Prognosis)	Systematic review of inception cohort studies	Inception cohort studies	Cohort study or control arm of randomized trial*	Case-series or case-control studies, or poor quality prognostic cohort study**	n/a
<b>Does this intervention help?</b> (Treatment Benefits)	Systematic review of randomized trials or <i>n</i> -of-1 trials	Randomized trial or observational study with dramatic effect	Non-randomized controlled cohort/follow-up study**	Case-series, case-control studies, or historically controlled studies**	Mechanism-based reasoning
<b>What are the COMMON harms?</b> (Treatment Harms)	Systematic review of randomized trials, systematic review of nested case-control studies, <i>n</i> -of-1 trial with the patient you are raising the question about, or observational study with dramatic effect	Individual randomized trial or (exceptionally) observational study with dramatic effect	Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning
<b>What are the RARE harms?</b> (Treatment Harms)	Systematic review of randomized trials or <i>n</i> -of-1 trial	Randomized trial or (exceptionally) observational study with dramatic effect			
<b>Is this (early detection) test worthwhile?</b> (Screening)	Systematic review of randomized trials	Randomized trial	Non-randomized controlled cohort/follow-up study**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning

\* Level may be graded down on the basis of study quality, imprecision, indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size.

\*\* As always, a systematic review is generally better than an individual study.

### How to cite the Levels of Evidence Table

OCEBM Levels of Evidence Working Group\*. "The Oxford 2011 Levels of Evidence".

Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?o=5653>

\* OCEBM Table of Evidence Working Group = Jeremy Howick, Iain Chalmers (James Lind Library), Paul Glasziou, Trish Greenhalgh, Carl Heneghan, Alessandro Liberati, Ivan Moschetti, Bob Phillips, Hazel Thornton, Olive Goddard and Mary Hodgkinson



# 牛津實證醫學中心證據等級表 2011 (翻譯 by EGRD Office)

臨床問題	建議等級 A	建議等級 B		建議等級 C	建議等級 D
	證據等級 1*	證據等級 2*	證據等級 3*	證據等級 4*	證據等級 5*
此臨床問題多常見?	區域及當代隨機樣本的(人口)普查	切合當地環境普查的系統性回顧**	區域性非隨機樣本調查**	系列病例報告**	n/a
診斷或監測工具正確嗎? (診斷)	使用一致的參考(黃金)標準及盲化的斷代研究的系統性回顧	使用一致的參考(黃金)標準及盲化的斷代研究	非連續性研究 或 沒有使用一致的參考(黃金)標準的研究**	病例對照試驗, 低品質或 無獨立參考標準的研究**	基於基本原理的推斷
如果不治療會有什麼後果? (預後)	初期(inception)世代研究的系統性回顧	初期(inception)世代研究	世代研究 或 隨機對照研究的控制組**	系列病例報告, 病例對照試驗 或 低品質的預後型世代研究**	n/a
這個介入治療有幫助嗎? (治療益處)	隨機對照試驗(RCT) 或 隨機序列治療試驗(n-of-1) <sup>(a)</sup> 的系統性回顧	隨機對照試驗 或 有顯著療效的觀察型研究	非隨機控制的世代/追蹤研究**	系列病例報告, 病例對照試驗 或 歷史對照試驗**	基於基本原理的推斷
介入治療的常見害處? (治療傷害)	隨機對照試驗(RCT), 巢式病例對照試驗, 隨機序列治療試驗(n-of-1) <sup>(a)</sup> 或 有顯著療效的觀察型研究的系統性回顧	隨機對照試驗 或 有顯著傷害 <sup>(b)</sup> 的異常觀察型研究	非隨機控制的世代/追蹤研究 (上市後監測) 提供足夠的樣本數去排除常見之危害 (關於長期危害, 需有足夠的追蹤時間)**	系列病例報告, 病例對照試驗 或 歷史對照試驗**	基於基本原理的推斷
介入治療的罕見害處? (治療傷害)	隨機對照試驗(RCT) 或 隨機序列治療試驗(n-of-1) <sup>(a)</sup> 的系統性回顧	隨機對照試驗 或 有顯著傷害的異常觀察型研究			
值得早期診斷嗎? (篩檢)	隨機對照試驗(RCT)的系統性回顧	隨機對照試驗	非隨機控制的世代/追蹤研究**	系列病例報告, 病例對照試驗 或 歷史對照試驗**	基於基本原理的推斷

# 牛津實證醫學中心2011年 實證證據等級

	Level 1	Level 2	Level 3	Level 4	Level 5
這個介入助 嗎治療有幫 (治療益處)	RCT或隨 機 序列治 療試 驗之 SR	RCT或有 顯 著療效 的觀 察型 研究	非隨機控 制的世代/ 追 蹤研究	系列病例報告 或病例對照試 驗或歷史對照 試驗	基於基本 原理的推 斷



# 步驟三 嚴格評讀文獻



- 文獻的品質等級：研究方法的評價
- 如為系統性回顧文獻，應視其使用的文獻等級及個別文獻之特性。
- 與我的病人或欲探討之情境相同否？
- 所有的結果指標利弊分析
- 措施成本效益

# 文獻評讀三步驟

**Validity (Reliability) 效度/信度**  
Can we believe it ?  
(研究方法的探討)

有效?

**Importance (Impact) 重要性**  
We believe it ! But does it matter?  
(研究結果的分析)

可信?

**Practice (Applicability) 適用性**  
If we believe it -  
does it apply to our patients?  
(如何在臨床運用)

可用?

# 效度(Validity)

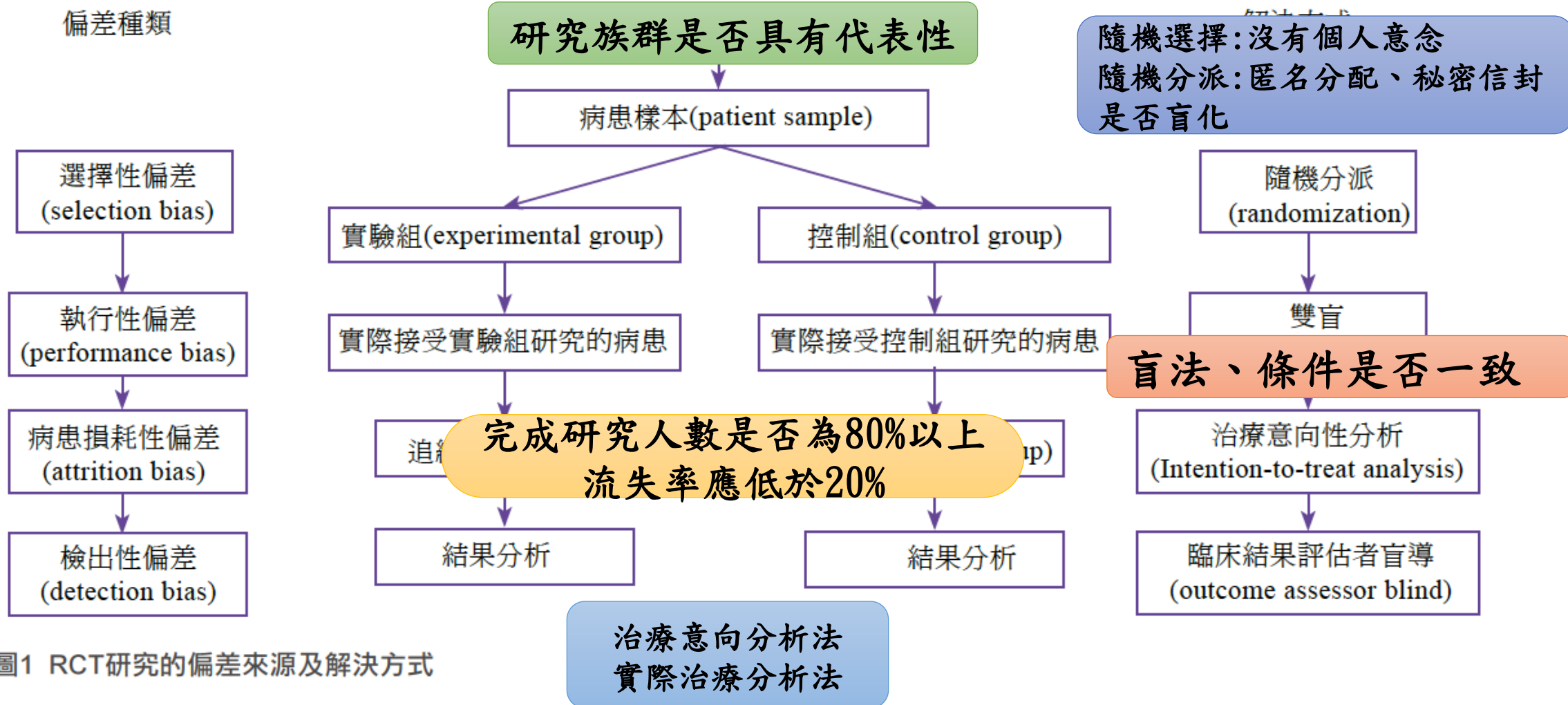


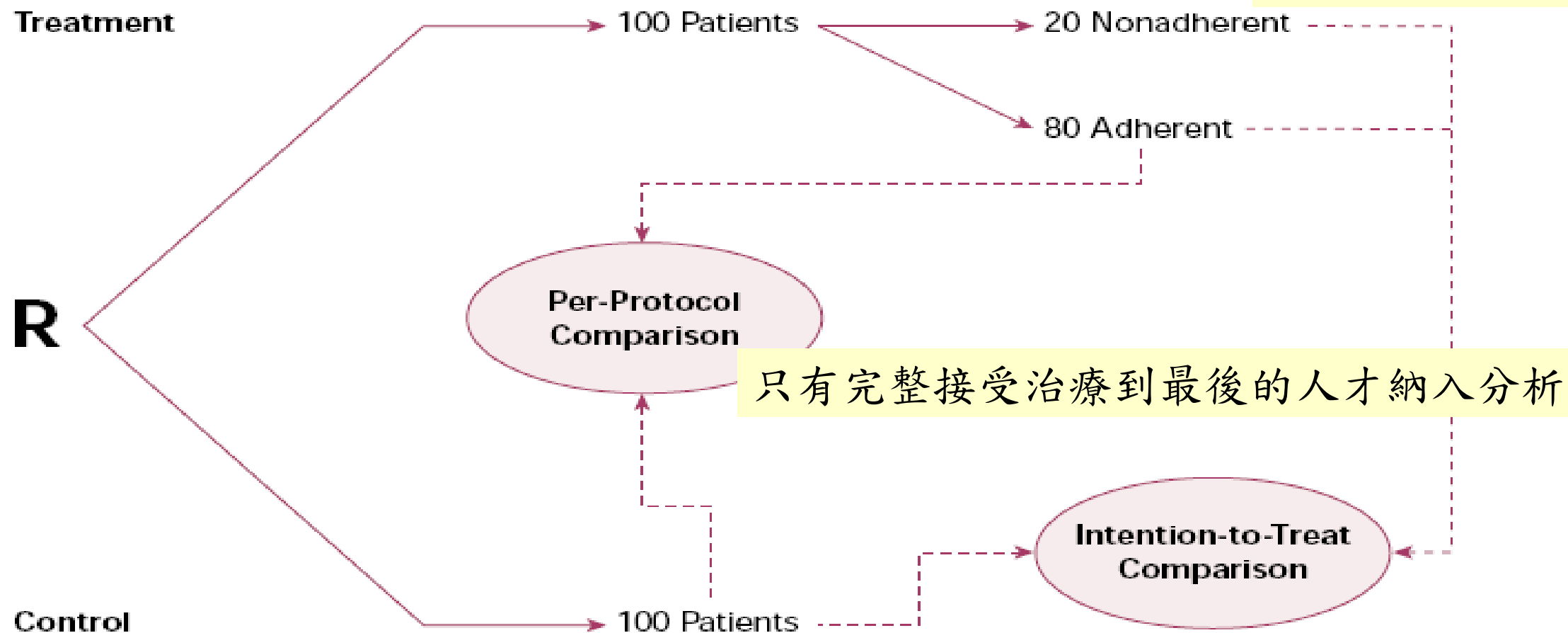
圖1 RCT研究的偏差來源及解決方式

# Intention to Treat Analysis

FIGURE 2B1-2

A Schematic View of Per-Protocol and Intention-to-Treat Comparisons

(1) 違反計畫書  
(2) 接受與預定不同治療



# 重要性(Importance)

- “ 我們相信它，但這個結論重要嗎? ”
- How large was the intervention effect? 成效大小
  - 若在臨床執行，是否可造成重要性的差異
  - 比較兩組發生率差異或比值
- How precise was the estimate of the intervention effect? 成效精確度

p 值 < 0.05 其實只代表該論文的研究結果具有統計學上意義，但是並不一定具有臨床意義(或臨床價值)；統計學上的有意義並不代表臨床上的有意義。

– 休司的右倫比且，世間已指上別衣小拍加兼双

# VIP簡介

- 勝算比(Odds Ratio)/危險比(Risk Ratio)

勝算比

- 發生某件事的人數與未發生該事件人數的比值

危險比  
(相對風險)

- 實驗組和對照組間產生的風險比率

相對危險比

- 發生某事件的風險在實驗組與對照組的比值

勝算比大於1，表示可能是危險因子；小於1則可能是有益的；若等於1則表示此因素在兩組間沒有影響。

# VIP簡介

- 例:

抽菸是否會增加肺癌的風險?抽菸的人調查10,000人發現有50人得肺癌，不抽菸的人亦調查10,000人發現得肺癌為10人。

	肺癌/事件(有病)	無肺癌/無事件(沒病)
抽菸(實驗組)	50/a	9050/b
不抽菸(對照組)	10/c	9990/d

事件的勝算比： $(a/b)/(c/d) = (50/9050)/(10/9990) = 5.5$

事件的危險比： $EER/CER = (a/a+b)/(c/c+d) = 5.0$

解釋：勝算及危險比均大於1，所以抽菸是罹患肺癌的危險因子。

# VIP簡介

- **益一需治數(Number Needed to Treat, NNT)**

- ◆簡單說:平均治療多少人，就會有一個有效。
- ◆ $NNT = 1$ 時，表示治療組的每一個病人都有效。
- ◆NNT越大，治療效果越差(必須把時間帶入)。
- ◆公式 $= 1/ARR(CER - EER)$ 。

- **害一需治數(Number Needed to Harm, NNH)**

- ◆治療進行中，導致一人死亡或不良反應，所需治療病人數
- ◆NNH數目越大越好。
- ◆公式 $= 1/ARI(CER - EER)$ 。



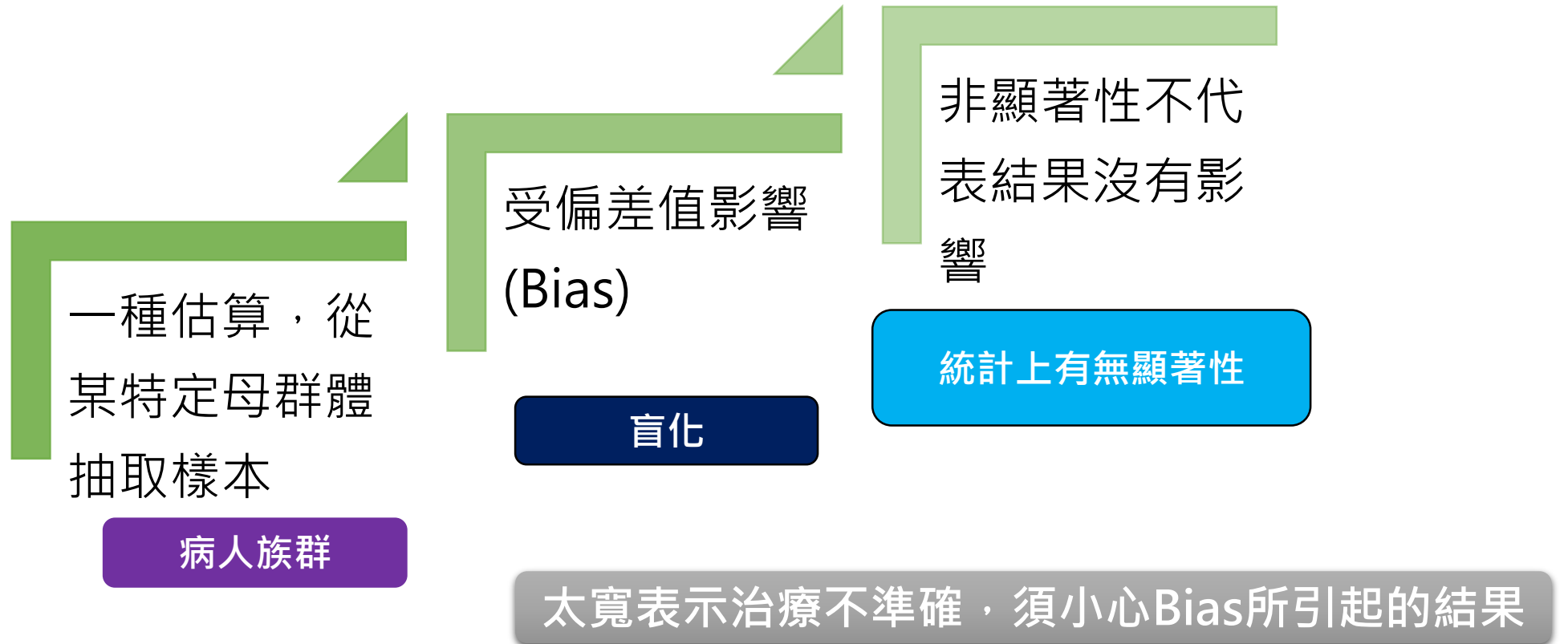
# VIP簡介

- 範例

- ◆ 胃潰瘍分別用Omeprazole(實驗組)及Cimetidine(對照組)治療，發現一年後復發率假設各為1%及10%，因此計算NNT為 $10\% - 1\% = 9\%$ (絕對風險降低度，ARR, Absolute risk reduction)， $NNT = 1/ARR = 1/0.09 = 11.1$ ，所以NNT為12，因此解釋為必須12個病人使用Omeprazole一年後，才會有一人不會復發胃潰瘍，相反若這12個病人使用Cimetidine一年後，都會復發胃潰瘍。

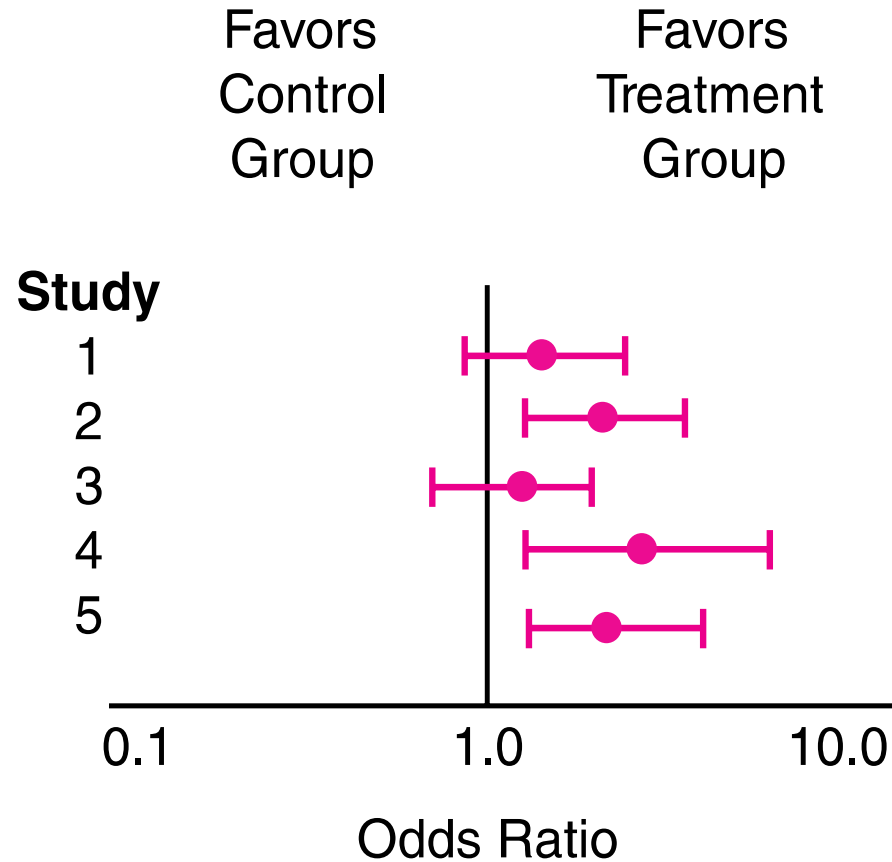
# 重要性(Importance)

- 信賴區間(Confidence Interval, CI)

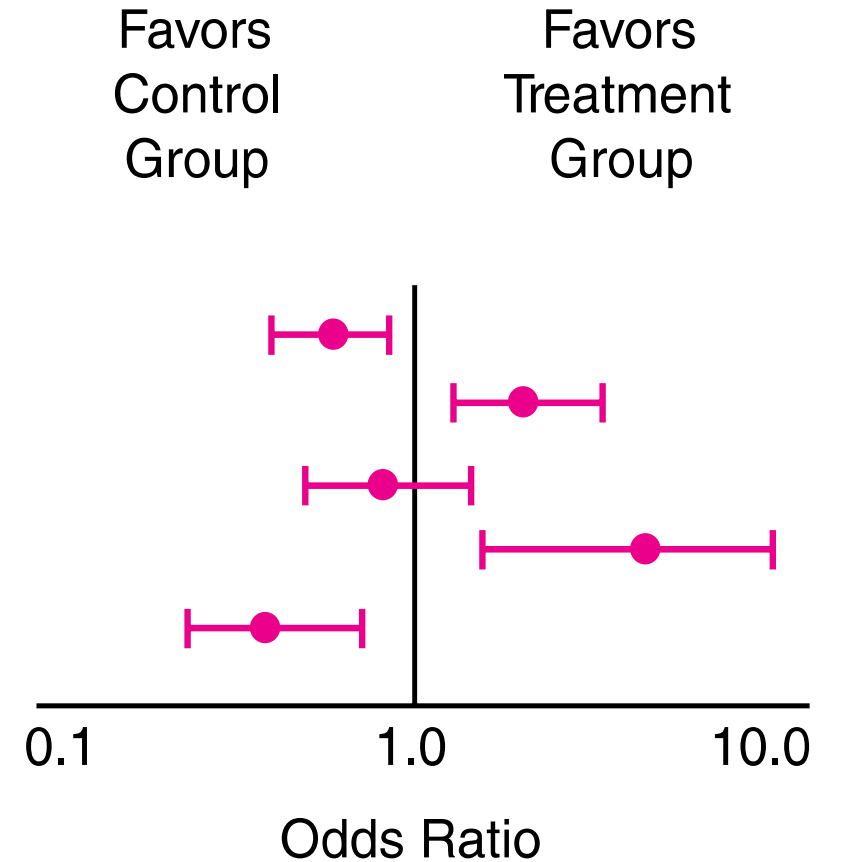


# forest plot

## A. Low Heterogeneity



## B. High Heterogeneity



● = Estimated effect size (OR)

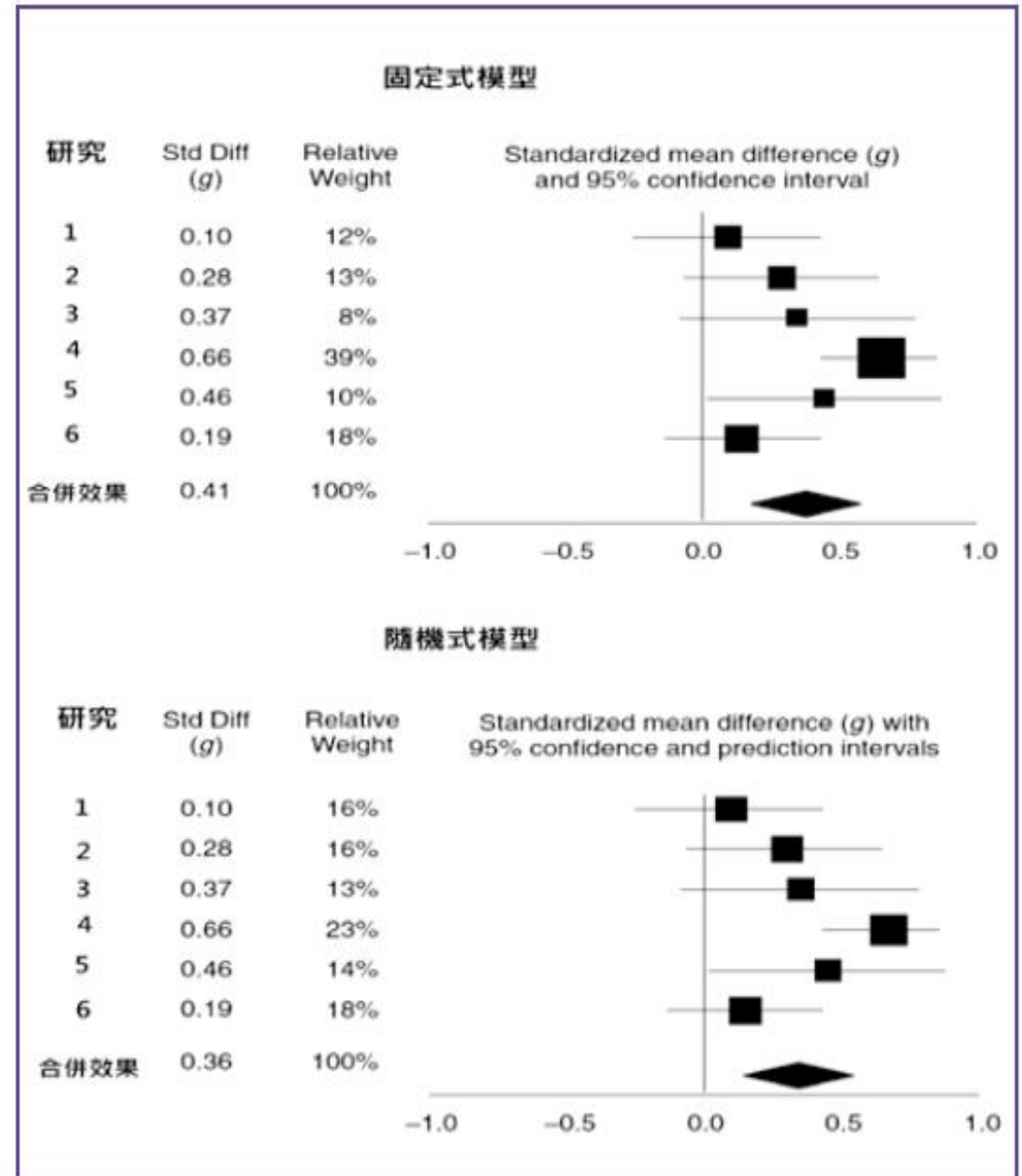
— = 95% CI

# Fixed effects model

- Assumption is that a single true effect size.
- The error term in a fixed effects model represents only within-study variation, and between-study variation is ignored.

# Random effects model

- A random effects model, assumes effects that each study estimates different.
- True relative effects and that the various effects are normally distributed around a mean effect size value.
- Takes both within- study and between - study variations into account.



# VIP簡介

統合計算後的研究總人數及事件發生總

各文章的作者及發表年代

各個文章的總人數及事件發生人數

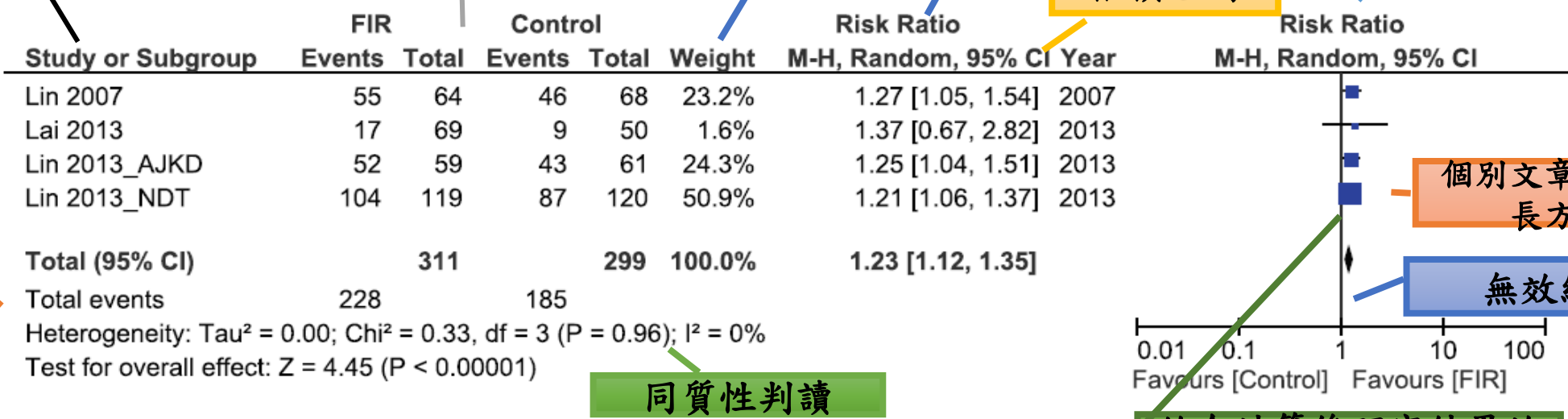
權重

各研究的勝算比或危險比

Far Infra-Red Therapy Role in Maturat

固定或隨機模式

信賴區間



個別文章研究的長方圖

無效線

統合計算後研究結果的菱形圖

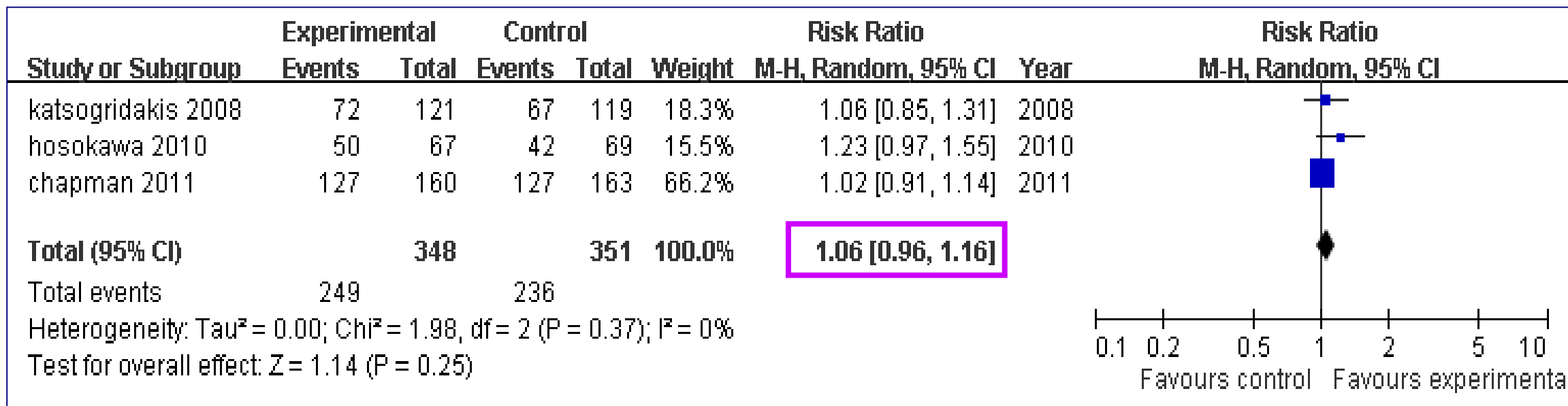
同質性判讀

Figure 2. Forest Plot showing Primary AVFs patency at 12 months. doi:10.1371/journal.pone.0104931.g002

判讀：  
 同質性表示文章可信度非常高。  
 I<sup>2</sup>可以另外證明，當I<sup>2</sup> < 25%表異質性低，若 > 50%表高異質性。

# 範例：靜脈探照燈輔助第一次注射成功之成效

(隨機控制研究)



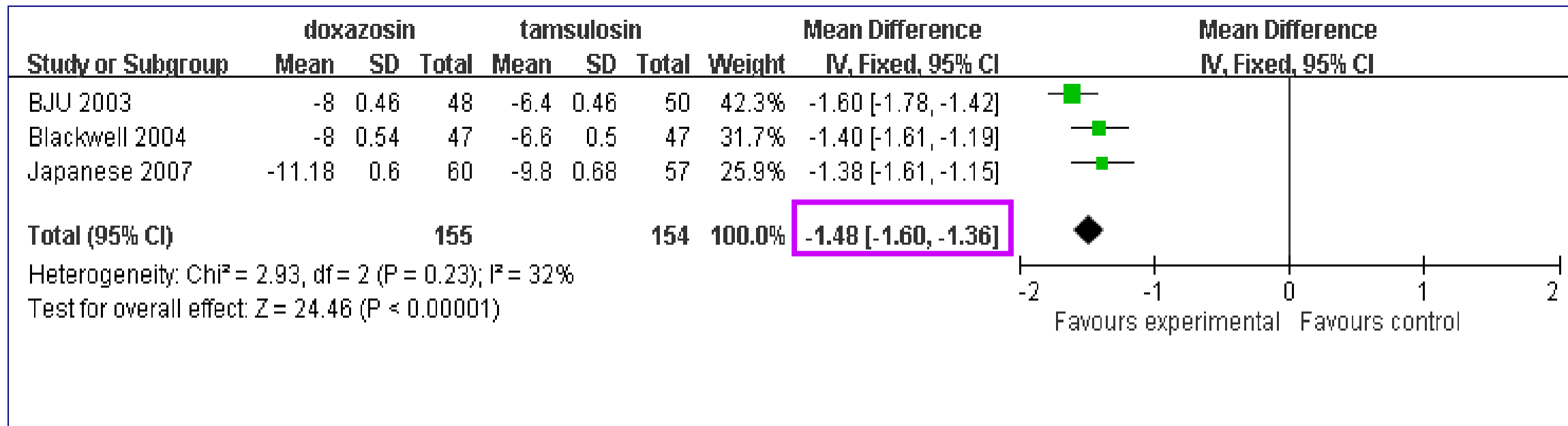
**解釋：**以靜脈探照燈輔助而第一次注射即成功之比率是未使用者的1.06倍 (95% CI 0.96-1.16)

由結果可知，

靜脈探照燈輔助執行靜脈注射的效益並不顯著

範例:

## 比較攝護腺肥大病人以doxazosin及tamsulosin之成效差異



- 以攝護腺症狀評分表評分，用藥8週後追蹤結果
- 此三篇文獻整合結果為服用doxazosin可降低攝護腺症狀1.48分(95% CI -1.60 ~ -1.36)，因信賴區間並未包含0，故可知doxazosin效果較tamsulosin佳

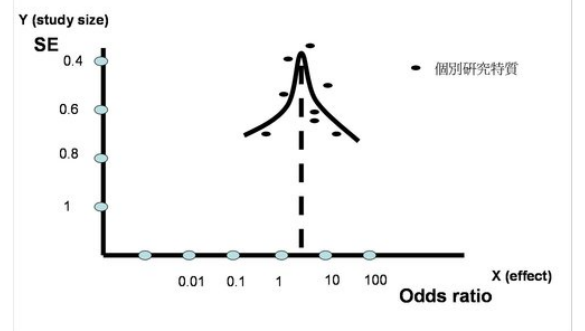
# VIP簡介

- 漏斗圖(Funnel plot)
- ◆ Publication bias
- ◆ Selective bias
- ◆ Reporting bias
- ◆ Study size
- ◆ Effect
- ◆ Heterogeneity .....



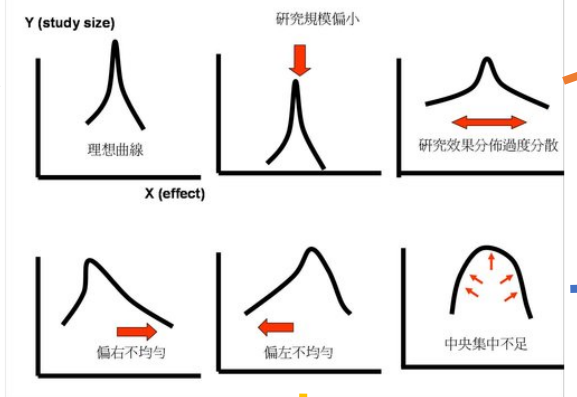
# VIP簡介

研究規模偏小



理想曲線

研究效果分佈過度分散



偏右不均

中央集中不足

偏左不均

# 運用(Practice/Applicability)

“ 這個結論可應用我們的病患嗎 ”

- ◆ 介入成效是否有臨床意義？
- ◆ 研究過程及對象是否與我們所關心的議題相近？
- ◆ 這個結果可以運用到我們的病人嗎？
- ◆ 病人會從此處理得到多少好處及壞處？
- ◆ 病人價值的傾向與病人的優先考量？
- ◆ 倫理？
- ◆ 經濟影響或資源？

# 評讀工具的選擇



# CEBM (Centre for Evidence-Based Medicine)

- 牛津大學實證醫學中心 2010 發展
- SR、RCT、Diagnostic Study、Prognosis Study Checklist
- 已有中文版(楊宗翰，邵時傑譯，2017)
- <http://www.cebm.net/critical-appraisal/>
- 免費下載

# CEBM RCT Checklist

8 questions to help you make sense of a RCT 是、否、不清楚

## A) 研究結果是否有效？（內部效度）

1. 患者是否採隨機分派進入治療組？
2. 各組在試驗開始時是否相似？
3. 除了被分派的治療外，各組是否接受相同的處置？
4. 是否所有參加試驗的患者都列入計算？且依原隨機分派的組別分析？
5. 測量是否客觀？或病患和醫師不知道接受的治療為何？

## B) 研究結果為何？

6. 治療效果有多大？
7. 治療效果估計值的精確性如何？

## C) 適用性（外部效度 / 應用）

8. 這個治療對我的患者有益嗎？

# CASP RCT Checklist

11 questions to help you make sense of a RCT 是、不明確、否

## A) 研究結果可信嗎

1. 研究問題是否清楚且聚焦？
2. 病人被隨機分派到不同組別？
3. 是否所有的病患都有納入結果中去分析？
4. 病人、(給藥、測量結果的)醫療照護者、分析數據人員是否都採盲化？
5. 研究各組族群相似嗎？
6. 除了實驗的介入措施之外，各組是否接受相同對待？

## B) 研究結果為何

7. 介入措施的效果有多大？
8. 治療效果是否精確？

## C) 研究結果對於當地病人有幫助嗎

9. 此研究是否可應用到你的病患？
10. 是否所有重要的臨床結果都被考量到？
11. 結果的益處是否勝於傷害與成本？

# 參考資料

- 刁茂盟、郭耀仁(2014)·實證醫學功夫談(初版)·合記·ISBN:978-986-126-956-6
- 劉芝妤、王瑋、郭素娥、劉靜如、周芳如、黃秋雯(2021)·ABCDE bundle介入對加護病房病人譫妄之成效—系統性回顧暨統合分析·台灣公共衛生雜誌,40(1),71-82。  
[http://dx.doi.org/10.6288/TJPH.202102\\_40\(1\).109104](http://dx.doi.org/10.6288/TJPH.202102_40(1).109104)
- Straus, S. E., Glasziou, P., Richardson, W. S., & Haynes, R. B. (2019)·臨床實務與教學指引(沈英琪、林佩姿譯;五版)·愛思唯爾。(原著出版於2019)。  
ISBN:978-986-98014-0-9
- Kim, H. J., & Kim, H. J. (2018). Effects of Application of Elastic Compression Stockings on Blood Pressure, Pulse Rate, and Hypotensive Symptoms in Patients with Intradialytic Hypotension. *Nephrology Nursing Journal*, 45(5), 455-461.

評讀工具	SR/Meta	RCT	Cohort	Case control	Observation study	Diagnostic test	Prognosis
CASP	V	V	V	V		V	
CEBM	V	V				V	V
PRISMA	V						
STROBE			V	V	V		
CONSORT		V					
Jadad scale		V					
JBI	V	V	V	V	V	V	



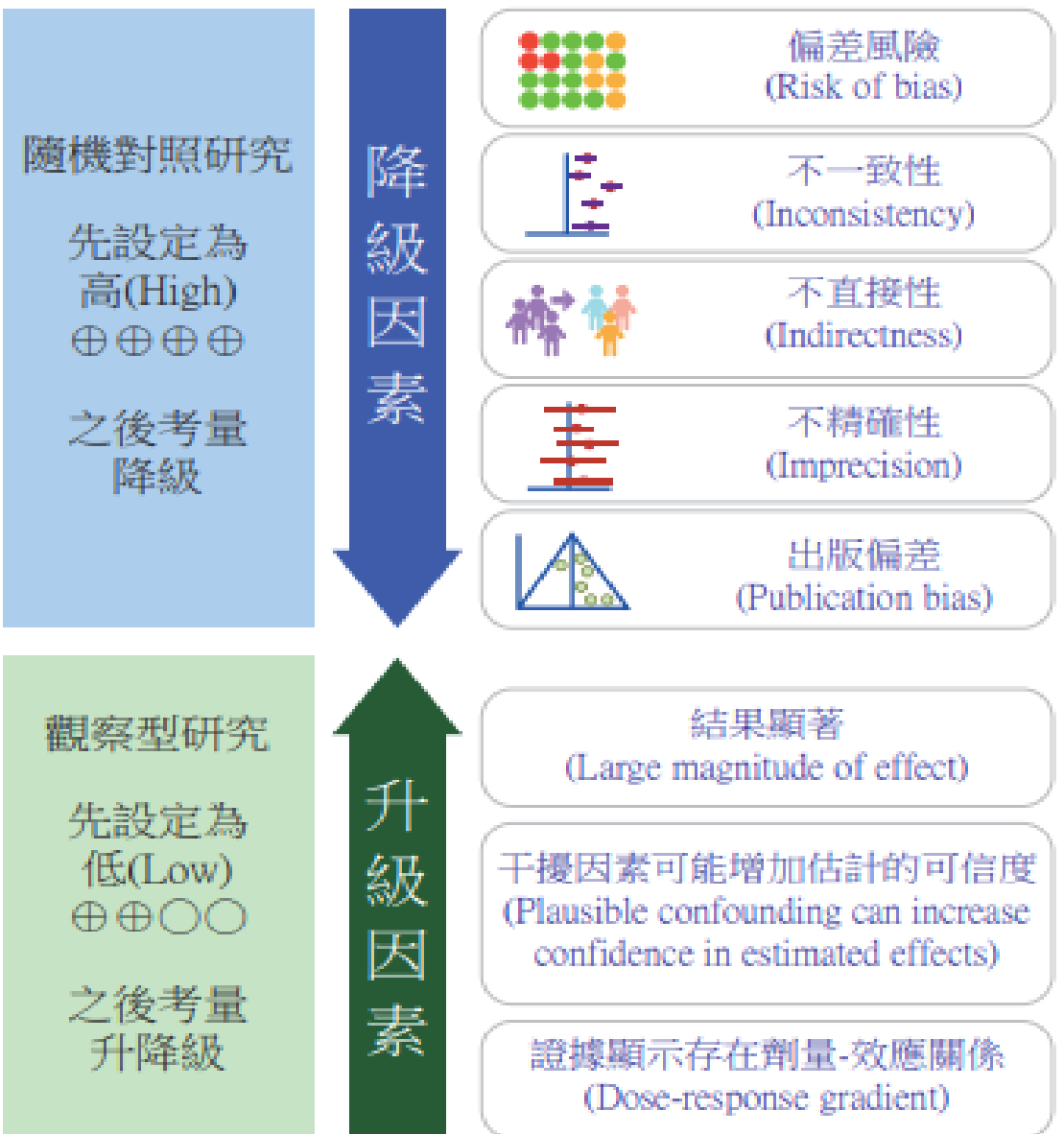
# Jadad Scale-判斷研究品質

評估項目	分數	說明
1.是否隨機分派 ( randomized )	2	詳細說明如何進行隨機分派且正確
	1	提及隨機分派，但未說明方式
	0	未採隨機方式
2.是否為雙盲 ( double-blind )	2	具體說明如何進行雙盲，且被認為恰當
	1	提及採雙盲試驗，但未說明如何進行
	0	使用單盲或未採盲化
3.對失聯及退出 樣本的追蹤	1	清楚說明個案退出及失聯原因
	0	未說明個案退出及失聯原因

# Modified Jadad Score

評估項目	分數	說明
1.是否清楚描述研究收案與排除條件	+1/0	是/否
2.是否隨機分派	+1/0	是/否
3.隨機分派方式是否合適	+1 -1 0	是 否 未說明方式 採雙盲 採單
4.是否採盲化	+1 +0.5 0	盲 未採盲化 是/否 是/ 否
5.盲化方式是否合適	+1/0	是/否 是/
6.是否清楚說明樣本失聯及退出的原因	+1/0	否
7.是否描述有評估不良反應	+1/0	
8.是否清楚描述統計方法	+1/0	

# 形成臨床建議 (GRADE 評核系統)



表三 影響證據等級的因素

## 可能降低隨機對照試驗證據等級的因素及其解釋

不直接性	不直接性可分兩類：一是比較兩種介入措施的療效時，沒有單獨的研究直接比較二者的隨機對照試驗，但可能存在每種介入與安慰劑比較的多個隨機對照試驗，這些試驗可用於進行二者之間療效的間接比較，但提供的證據等級比單獨的研究直接比較的隨機對照試驗要低。二是研究中所報告的族群、介入措施、對照措施、預期結果等，與實際應用時存在重要差異。
不精確性	當研究納入的病人和觀察事件相對較少，而導致信賴區間較寬時，需降低其證據等級。

## 可能提高觀察性研究證據等級的因素及其解釋

結果顯著	當方法學嚴謹的觀察性研究顯示療效顯著、或非常顯著，且結果高度一致時，可提高其證據等級。
干擾因素可能增加估計的可信度	當影響觀察性研究的偏差不是誇大其效果，而可能是低估效果時，可提高其證據等級。
證據顯示存在劑量-效應關係	當介入的劑量和產生的效應大小之間有明顯關聯時，即存在劑量-效應關係時，可提高其證據等級。
升級標準：以上三個因素中任意一個因素，可根據其大小或強度，將證據等級升1級（如相對危險度 [relative risk, RR] 大於 2 或小於 0.5）、或升2級（如相對危險度大於 5 或小於 0.2），證據等級最高可升級到高證據等級（A 級）。	

# 形成臨床建議 (GRADE 評核系統)

證據品質

GRADEpro GDT (<https://gradepro.org>)



表二 證據等級與建議強度分級

證據等級分級	具體描述
高 (A)	非常有把握觀察值接近真實值
中 (B)	對觀察值有中等把握：觀察值有可能接近真實值，但也有可能差距很大
低 (C)	觀察值的把握有限：觀察值可能與真實值有很大差別
極低 (D)	對觀察值幾乎沒有把握：觀察值與真實值可能有極大差別
建議強度分級	具體描述
強 (1)	明確顯示介入措施利大於弊、或弊大於利
弱 (2)	利弊不確定或無論品質高低的證據均顯示利弊相當

表四 GRADE 中建議強度的含義

### 強烈建議的含義

對病人	幾乎所有病人均會接受所建議的方案；此時若未接受建議，則應說明。
對臨床醫師	應對幾乎所有病人都建議該方案；此時若未給予建議，則應說明。
對政策制定者	該建議方案一般會被直接採納到政策制定中。

### 弱建議的含義

對病人	多數病人會採納建議方案，但仍有不少病人可能因不同的偏好與價值觀而不採用。
對臨床醫師	應該認識到不同病人有各自適合的選擇，幫助每個病人做出符合他偏好與價值觀的決定。
對政策制定者	制定政策時需要充分討論，並需要眾多權益關係人參與。

用

Apply

# 步驟四 應用於病人身上

- 將以實證為基礎的訊息傳達給病人、藉由雙向討論，共同做出最適合病人的臨床決定
- 應用於臨床照顧的對象

# 步驟4：應用於病人身上

- 為有效做出臨床決定，護理人員可
  - 列出照護的選擇
  - 各選擇的好處及危險性
  - 選項之綜合評估顯示好處高於危險性，則傾向執行該措施
- 在機構或單位中如何執行改變
  - 臨床運用是否困難（人力、政策、經費）
  - 是否考量成本效益
- 實證人員應藉由溝通，將以實證為基礎的訊息傳達給病人、藉由雙向討論，共同做出最適合病人的臨床決定



審

Assess

# 步驟五 對過程進行稽核

依病人結果評值成效，決定是否繼續執行此措施或加以修正

- 結果指標—健康照顧最終成果
  - 如病人住院天數、感染率
  - 可觀察個案症狀、行為、感受或評價
  - 治療副作用之降低、身體功能之恢復
  - 可反應醫療費用的降低
  
- 發展或修改臨床照護準則及指引
- 修改常規活動

# 練習評讀一下！！

- 白噪音對於成人加護病房病人睡眠促進之成效：隨機對照試驗

魏可婷, 吳珮菁, 吳欣玫, & 謝伶瑜. (2021). 白噪音對於成人加護病房病人睡眠促進之成效: 隨機對照試驗. *志為護理-慈濟護理雜誌*, 20(3), 71-82.

# 參考資料

- 台灣護理學會實證健康照護知識館(2021) · 文獻評讀與檢核表 · <https://www.ebhc.e-tna.org.tw/1132/1177/1264/>
- 考科藍臺灣研究中心 指引發展工作小組 ( 2020 ) · 台灣實證臨床指引發展及更新手冊(初版) · 考科藍臺灣研究中心。
- 宋慧娟(2018) · 嚴格評讀證據之效度與重要性 · 於穆佩芬總校閱，實證護理(初版，319-326頁) · 新文京。
- 簡莉盈、劉影梅(2017) · 實證護理學導論(三版) · 華杏。
- Kim, H. J., & Kim, H. J. (2018). Effects of Application of Elastic Compression Stockings on Blood Pressure, Pulse Rate, and Hypotensive Symptoms in Patients with Intradialytic Hypotension. *Nephrology Nursing Journal*, 45(5), 455-461.
- Polit, D. F., & Beck, C.T. (2015). *Nursing Research: Generating and Assessing Evidence for Nursing Practice*. (10th ed.) Philadelphia, PA: Wolters Kluwer Lippincott Williams & Wilkins.