流行病學方法論及實驗

The Methods of Epidemiology and Practices

授課教師：陳秀熙 教授
許辰陽 助理教授

流行病學方法論及實驗

(The Methods of Epidemiology and Practices)

一、 牛頓與南丁格爾流行病學故事

二、 辛普森矛盾故事與護理流行病學

三、 解決辛普森矛盾之護理魔術

四、 觸媒流行病學故事與護理照護

五、 偏好選擇流行病學故事之護理觀

六、 護理人如何校正偏好選擇之實例

七、 貝氏柯南護理流行病學

八、 賽先生護理流行病學

九、 信校度之護理流行病學

十、 傳染病流行病學之護理觀

十一、 基因與護理行為科學之流行病學應用

參考書籍

1. Beth Dawson, Robert G. Trapp *Basic & clinical biostatistics* New York : Lange Medical Books-McGraw-Hill, Medical Pub. Division, c2004
2. Robert C. Elston, William D. Johnson *Basic biostatistics for geneticists and epidemiologists : a practical approach* Chichester, U.K. : John Wiley & Sons, 2008
3. Stephen C. Newman *Biostatistical methods in epidemiology* New York : John Wiley & Sons, c2001
4. Breslow NE, Day NE. *Statistical methods in cancer research. Volume I - the analysis of case-control studies* IARC Sci Publ, 1980; (32):5-338.

一、 牛頓與南丁格爾流行病學故事-流行病學盛行率發生率

1. 盛行率 (Prevalence) **(**e-book: 01. Basic & Clinical Biostatistics Ch2. p30-32; 02. Basic Biostatistics for Geneticists and Epidemiologists Ch3. p55-58; 04. Statistical Methods in Cancer Research Ch2. p42-49) :

   代表某狀態 (如染病與否) 在人群中的占比

   (1) 定義: 該指標可用來量測在人群中在某特定時間點上, 某一相對上恆定的特徵 (如近視) 或疾病 (如肥胖症) 的比例. 在所附文章的表二中 (Prevalence, incidence, and mortality of PD, 2001), 盛行率用以描述巴金森氏症在人群中所占的比例.

   (2) 量測該指標所需的研究設計: 橫斷式研究設計 (Cross-sectional survey) 即可測得

   (3) 估計方法: 盛行率為比例的量測

$$\text{盛行率(Prevalence)} = \frac{m}{N} \qquad \text{-- (1)}$$

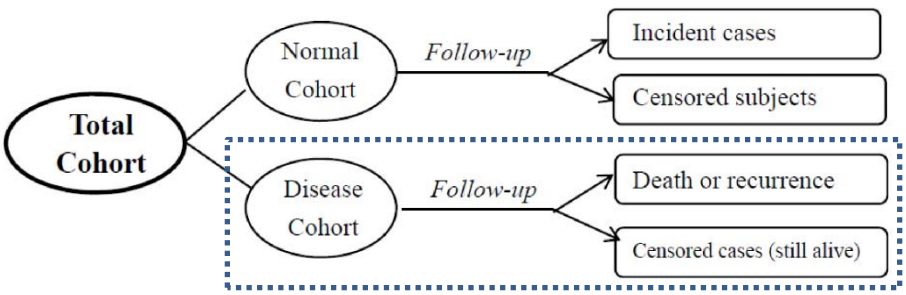
   理論上,這個指標是沒有單位; 以所附文章的巴金森氏症為例: 其中 m 為巴金森氏症個案數, 而 N 為調查族群的總人數, m/N 即為在該觀測時間, 巴金森氏症在所有調查族群中佔的比例.

(4) 生物醫學應用：盛行率代表了某疾病狀態在人群中的占比，故可用來衡量該疾病對於研究群體的罹病之負擔 (disease burden)，以利於評估其對於醫療資源以及人力的耗用情形，因此該指標可據以進行衛生行政措施規劃. 由於盛行率僅為比例，並不加入時間因素的考量，故不適合用作病因 (etiology)探討.

2. 發生率 (incidence) **(**e-book: 01. Basic & Clinical Biostatistics Ch2. 36-39; 02. Basic Biostatistics for Geneticists and Epidemiologists Ch3. p55-58;04. Statistical Methods in Cancer Research Ch2. p42-49)

(1) 發生率概念與牛頓力學之瞬時速度之概念相同. 例如隨附文章 (Prevalence, incidence, and mortality of PD, 2001) 的表二中，發生率的表示帶有時間單位 (no./100,000 person-yr; 每 100,000 人-年所發生的事件數)，此與力學中的加速度(a = $\frac{F(力)}{m(質量)}$) 相呼應。

(2) 量測所需的研究設計：世代研究 (Cohort follow-up study)



將上圖中之所有世代個體分為兩部分：無病世代 (Normal cohort) 以及罹病世代 (Disease cohort)，藉由追蹤無病世代，得到所追蹤時間的新

發個案數以計算發生率；而另一部分則是有關罹病世代，藉由追蹤染病世代，得到所追蹤時間中的死亡人數或復發個案數以計算及疾病死亡率或復發率. 兩者之概念相同。

(3) 估計方法:

$$\text{發生率(Incidence)} = \frac{\text{總事件數 (如巴金森氏症事件數)}}{\text{總觀察 人–時數 (如總人年數)}} \quad \text{-- (2)}$$

發生率代表了在單位時間 (年) 內發生事件 (例如: 巴金森氏症) 的相對瞬時速度. 發生率越高，則該族群發生該特定事件的風險越高

(4) 生物醫學應用: 發生率可以作為找出疾病病因 (etiology) 的推論基礎

為了理解上述發生率之估計方法，我們必須對速度以及瞬時速度的概念有所了解。

(5) 速度之概念: 存活函數可以作為評估相關事件 (例如: 巴金森氏症) 或者死亡發生速度的基礎. 存活函數 (S(t)) 的正式定義如下(e-book 03. Epidemiology and Biostatistics Ch20 p215-218):

$$S(t) = P(\text{在 } 0 \text{ 時到 } t \text{ 時存活})$$

$$= P(\text{在 } [0,t] \text{ 時間區間內存活}) \quad \text{-- (3)}$$

此一定義相當於如下之表示:

S(t) = P(存活超過 t 時)

=P(T≥t)，$0 \le S(t) \le 1$　(T 代表存活時間之變數)

舉例說明如下: 若有一存活函數為 $S(t) = e^{-0.04t}$，式中之 0.04，如同上述發生率。

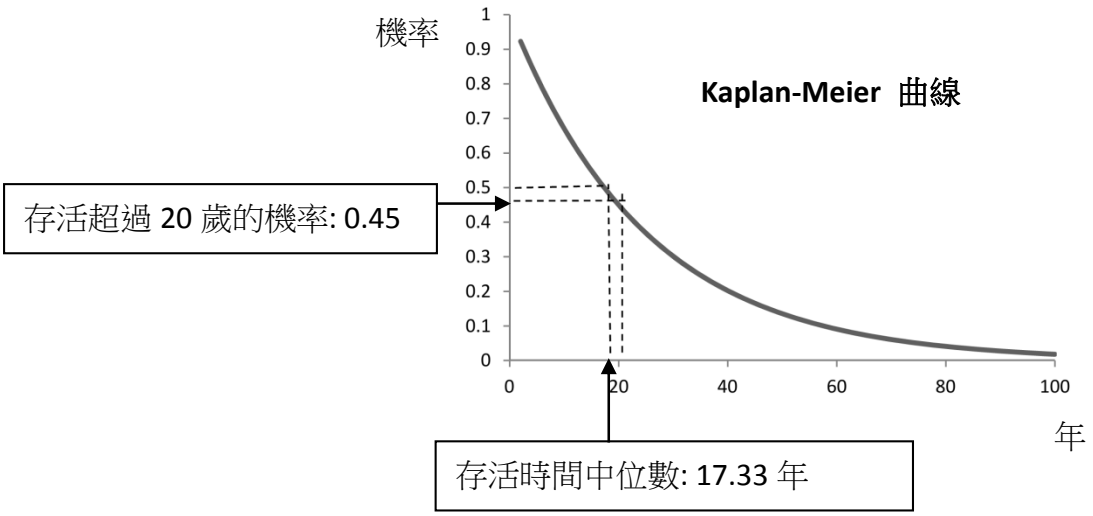
(1) 存活超過 t (年齡)=20 歲之機率為

$$S(20) = P(T \ge 20) = e^{-0.04\times 20} = 0.45$$

(2) 存活時間中位數 (即半衰期) 為

$$e^{-0.04\times t_M} = 0.5,\ t_M = 17.33 \text{ 年}$$

將該存活函數繪圖如下



(6) 瞬時速度 (instantaneous rate)

牛頓將瞬時速度定義為: 在極短的時間區段(Δt)中 S(t)的變化, 可以寫成

S(t) 的微分: $\frac{d}{dt}S(t)$

,或者以

S(t) 到 $S(t+\Delta t)$ 間的斜率: $\frac{S(t+\Delta t)-S(t)}{\Delta t}$

得到近似值

(7) 相對瞬時速度 (instantaneous relative rate，風險率)

由於瞬時速度無法完全適切的反應疾病風險的分母概念,因此上述牛頓所定義的瞬時速度絕少應用於描述死亡或者疾病之發生, 例如:

同樣觀察到瞬時速度為每月 10 個死亡事件, 但觀察族群為 1,000 人與另一觀察族群為 10,000 人, 事件發生的風險雖在兩族群有所不同, 以瞬時速度卻無法表現在單純的速度量測中 (皆為每月 10 個事件).

因此, 瞬時相對速度便用以描述風險的概念:

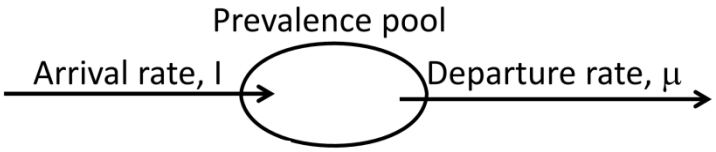
瞬時相對速度 (Instantaneous relative rate)=

$$h(t) = \frac{\frac{-d}{dt}S(t)}{S(t)} \qquad \text{-- (4)}$$

其中, 上式分子為原來的瞬時速度的概念, 而分母的 S(t) 表示事件之發生是基於具有此事件風險之族群下產生的。

3. 盛行率與發生率之比 (P/I ratio) **(**03. Epidemiology and Biostatistics Ch1. p5-11)

(1) 可用下圖的 prevalence pool 說明此一概念



進入疾病狀態的速度（例如：發生率）

離開疾病狀態的速度（例如：致死速率）

若某一受觀測的族群大小為 N，其中有 m 個巴金森氏症的盛行個案. 在橫斷式研究中可以估計得到盛行率 P = $\frac{m}{N}$ (式 (1)).

假定該族群為一穩定族群 (人口恆定,移出與移入相等), 則在極短的時間區段(Δt) 內進入 prevalence pool 的人數等於離開 prevalence pool 的人數, 故其平衡式如下:

$$I \times (N - m) \times \Delta t = \mu \times m \times \Delta t$$

$$\frac{m}{(N - m)} = \frac{I}{\mu}$$

若 N 遠大於 m, 則上式近似於

$$\frac{m}{N} = \frac{I}{\mu}$$

$$P\,(prevalence) = \frac{I}{\mu}$$

$$\frac{P}{I} = \frac{1}{\mu} = \bar{D}\,(average\ duration)$$

(2) 生物醫學應用: P/I ratio 代表了處於疾病狀態的平均時間. 若該平均時間 $\bar{D}$可得到, 則 μ (departure rate) 即為 $\frac{1}{\bar{D}}$ 。

在基於指數分布的假設之下,存活函數即為

$$S(t)=e^{-\mu t} \qquad \text{-- (6)}$$

根據此參數 (μ) 舉例說明如下:

若某一對某族群進行觀測並估計得到該族群第二型糖尿病的平均時間 $\bar{D}$為 8 年, 則 $\mu=\frac{1}{\bar{D}}=0.125$

故存活函數為 S(t)= $e^{-\mu t}=e^{-0.125\times t}$, 依據此存活函數,可以推估

(A)該族群中第二型糖尿病患者的 5 年存活機率為

S(5)= $e^{-0.125\times 5}=0.535$

(B) 平均半存活時間(t_M)為

0.5= $e^{-0.125\times t_M}$

$t_M=\frac{0.693}{0.125}=5.544$ 年

流行病學方法論及實驗

(The Methods of Epidemiology and Practices)

授課教師：陳秀熙 教授

台灣大學流行病學與預防醫學研究所

學習宗旨

1. 南丁格爾故事及牛頓原理闡述社會物理 科學之重要

2.流行病學盛行率(prevalence)及 發生率(incidence)之應用

「手持燈籠的女士」（The lady with a lamp)
護理鼻祖南丁格爾（Florence Nightingale）
護理學的先驅及著名的應用統計學家

- 弗羅倫斯**應用**統計分析護理學專。
- 圓形圖 (Pie Chart)**應用**鑒於她在統計學方面的傑出成就
- 1858年被選為英國皇家統計學會會員 (第一位女性會員)，後來又選為美國統計學會榮譽會員。社會影響遍及印度，澳大利亞等國
- 發明雞冠花圖

，「南丁格爾—威廉法雷同盟」社會物理-Newton

威廉法雷 (William Far) :「現代流行病學之父」(另一位是約翰斯諾 John Snow)-

運用統計數字追查病源和改善國民健康，「國際疾病分類法」（International Classification of Disease）便是以法雷的方法為基礎。

南丁格爾: 護士之母

The wrong denominator in Epidemiology

法雷計算英國醫院死亡率:

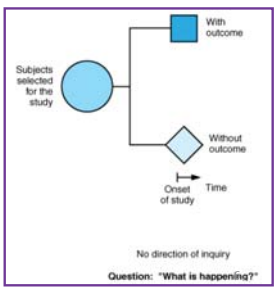
醫院**一年**死亡人數/醫院**一天**之內的住院人數

南丁格爾根據法雷的統計報告，宣稱英國二十四間醫院的死亡率高達百分之九十。

Prevalence and Study Design

e-book: **01**. Basic & Clinical Biostatistics Ch2. p30-32;
02. Basic Biostatistics for Geneticists and Epidemiologists Ch3. p55-58
04. Statistical Methods in Cancer Research Ch2. p42-49

- **Definition**:
  - This indicator is tailored for measuring the constant and existing frequencies of characteristics or disease among the number of population at a given time point (period). In our example, we are interested in Parkinson disease.
- **Study Design**:
  - Cross-sectional survey.



Prevalence

- **Estimate**:
  - It is a proportion.
  - $Prevalence = \frac{m}{N}$

    m number of Parkinson disease among the underlying population size equal to N.
- **Usefulness**:
  - Prevalence is used for measuring disease burden in community in order to elucidate health demand, health planning for clinical manpower, and resources allocation for Parkinson disease. All these aspects are related to health administration.

Original Research

Prevalence, incidence, and mortality of PD

A door-to-door survey in Ilan County, Taiwan

R.C. Chen, MD; S.F. Chang, MD, MPH; C.L. Su, MD; T.H.H. Chen, DDS, PhD; M.F. Yen, MS; H.M. Wu, BS; Z.Y. Chen, MD; and H.H. Liou, MD, PhD

Article abstract— *Background:* The reported prevalence and incidence rates of PD were significantly lower in China than those in Western countries. People in China and Taiwan have a similar ethnic background. *Objective:* To investigate the prevalence, incidence, and mortality rate of PD in Taiwan. *Methods:* The authors conducted a population-based survey using a two-stage door-to-door approach for patients aged 40 years or older in Ilan, Taiwan. Patients were diagnosed with PD by having at least two of the four cardinal signs of parkinsonism and exclusion of seconddary parkinsonism. To identify new cases of PD after the survey, patients with negative results of parkinsonism in the first stage were matched to the information on clinical diagnosis of PD from the Bureau of National Health Insurance toward the end of December 31, 1997. All cases of PD were linked to the Taiwan mortality registration to ascertain causes of deaths until December 31, 1999. *Results:* The participation rate was 88.1% among the 11,411 contacted individuals. Thirty-seven cases of PD were identified. The age-adjusted prevalence rate of PD for all age groups was 130.1 per 100,000 population after being adjusted to the 1970 US census, assuming no cases of PD would be found among those younger than 40 years of age. Of 9972 non-PD subjects in the first screen, 15 new cases of PD were ascertained. The age-adjusted incidence rate was 10.4 per 100,000 population for all age groups. The case fatality rate of PD after a 7-year follow-up was 40.4% (21 deaths in 52 patients with PD). The relative risk of death for PD cases versus non-PD cases was 3.38 (95% CI: 2.05–4.34). The 5-year cumulative survival rate in PD cases (78.85%) was statistically lower than that in non-PD cases (92.84%). *Conclusion:* The prevalence and incidence rates of PD in Taiwan were much higher than those reported in China, but closer to those in Western countries. These results suggest that environmental factors may be more important than racial factors in the pathogenesis of PD.

NEUROLOGY 2001;57:1679–1686

Original Research

Background: The reported prevalence and incidence rates of PD were significantly lower in China than those in Western countries. People in China and Taiwan have a similar ethnic background.

Objective: To investigate the prevalence, incidence, and mortality rate of PD in Taiwan.

Methods:

population-based survey (two-stage door-to-door approach) ⇨ patients with PD aged 40 years or older in Ilan, Taiwan

Original Research

A map of the survey area in Ilan, Taiwan

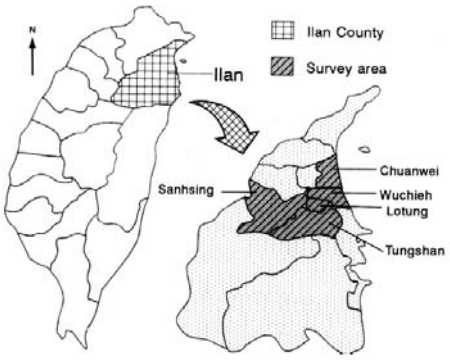


Figure 1. A map of the survey area in Ilan, Taiwan.

Original Research

Result

Table 1 *Age and sex distribution of the total population, targeted population, and participation rates in Ilan on January 1, 1993*

Age group, y	Men	Women	Both sexes
No. of total population			
0–39	32,184	29,929	62,113
40–49	2096	1876	3972
50–59	2108	1957	4065
60–69	1738	1501	3239
70–79	812	821	1633
80+	218	339	557
Total	39,156	36,423	75,579
Total (40+)	6972	6494	13,466
No. of target population			
⋮			
No. of participants			
⋮			
Participation rate, %			
⋮			

Original Research

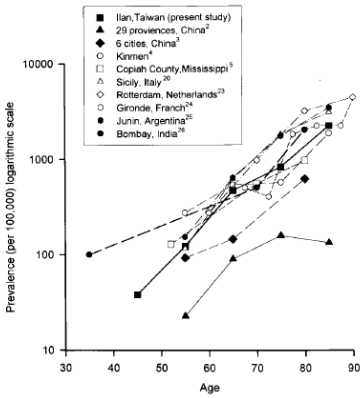
The prevalence per 100,000 of PD in Ilan

Table 2 *The prevalence per 100,000 and incidence per 100,000 of PD in Ilan*

	Men		Women		Both sexes	
Age, y	No. of cases	Prevalence per 100,000	No. of cases	Prevalence per 100,000	No. of cases	Prevalence per 100,000
40–49	1	78.3	0	0.0	1	37.8
50–59	4	252.5	0	0.0	4	122.5
60–69	3	224.9	11	896.5	14	546.7
70–79	4	645.2	6	1000.0	10	819.7
80+	3	2013.4	5	2325.6	8	2197.8
Total (40+)	15	302.2	22	431.9	37	367.9
Age-adjusted (40+)		299.2		423.7		357.9
Age-adjusted (for all age)		108.7		154.0		130.1

Original Research

Figure 3. Comparison of age-specific prevalence rates of PD obtained through door-to-door studies. The prevalence rates in Taiwan are much higher than those reported in mainland China but closer to those in Kinmen and Western countries.



Prevalence of Pressure Ulcers

Point Prevalence: A measure of the number of cases of pressure ulcers at a specific time.

Period prevalence: A measure of the number of cases of pressure ulcers over a prolonged time period such as the entire hospitalization.

Reflecting the total burden of the disease and providing insights into the magnitude of the pressure ulcer problem for the planning for health resource needs but provide fewer insights into the quality of care being delivered.

14

Sir Isaac Newton

- 1669 -- Lucasian Professor of Mathematics at Cambridge (a position now held by Stephen Hawking)
- 1689 -- Member of Parliament representing Cambridge
- 1699 -- Master of the Mint
- 1701 to 1702 -- Member of Parliament for the second time
- 1703 -- President of the Royal Society of London, the United Kingdom's national academy of science
- 1705 -- Knighted

牛頓第二運動定律與罹病後存活狀況

Newton's Second Law and Survival of Disease

- Survival vs Distance
  (存活) (距離)
- Hazard Rate (Incidence) vs Velocity
  (風險率) (發生率) (速度)

16

速度 (Velocity) 與 率 (Rate)

- 平均速度 Average Velocity

$$\bar{V} = \frac{x_j - x_i}{t_j - t_i} = \frac{\Delta x_{ij}}{\Delta t_{ij}}$$ x: 位置 t: 時間

- 瞬間速度 Instantaneous Velocity

$$\tilde{V} = \lim_{\Delta t \to \infty} \frac{\Delta x_{ij}}{\Delta t_{ij}}$$

17

Incidence and survival (1)

e-book **03**. Epidemiology and Biostatistics Ch20 p215-218

- **Concept of Rate**:
  - To measure rates of death or disease (such as Parkinson disease), we had better start from survival function.

$$S(t) = P(Surviving\ from\ time = 0\ to\ time = t)$$
$$= P(Surviving\ during\ [0, t])$$

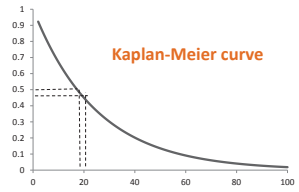
  - Equivalently,

$$S(t) = P(Surviving\ beyound\ time\ t)$$
$$= P(T \geq t) \qquad 0 \leq S(t) \leq 1$$

18

Incidence Rate (Hazard Rate) and survival (2)

- **Ex.** Suppose we have $S(t) = e^{-0.04t}$
  - The probability of surviving beyond t (age)=20 year is
    $S(20) = P(T \geq 20) = e^{-0.04 \cdot 20} = 0.45$
  - Median survival time (Half of life)
    $e^{-0.04 \cdot t_M} = 0.5, t_M = 17.33\ year$



Effect of Pressure Ulcers on the Survival of Long-Term Care Residents

- Background. Past studies have emphasized that patients with pressure ulcers are at high risk of dying. However, it remains unclear whether this increased risk is related to the ulcer or to coexisting conditions. In this study we examined the independent effect of pressure ulcers on the survival of long-term care residents.
- Methods. We evaluated all 19,981 long-term care residents institutionalized in Department of Veterans Affairs (VA) long-term care facilities as of April 1, 1993. Baseline resident characteristics and survival status were obtained by merging data from five existing VA data bases. Survival experience over a 6-month period was described using a proportional hazards model.
- Results., Pressure ulcers were present in 1,539 (7.7%) long-term care residents. Residents with pressure ulcers had a relative risk of 2.37 (95% CI = 2.13, 2.64) for dying as compared to those without ulcers. After adjusting for 16 other measures of clinical and functional status, the relative risk associated with pressure ulcers decreased to 1.45 (95% CI = 1.30, 1.65). No increased risk of death was noted for residents with deeper ulcers.
- Conclusions. Pressure ulcers are a significant marker for long-term care residents at risk of dying. After adjusting for clinical and functional status, however, the independent risk associated with pressure ulcers declines considerably. The fact that larger ulcers are not associated with greater risk suggests that other unmeasured clinical conditions may also be contributing to the increased mortality associated with pressure ulcers.

Effect of Pressure on Survival of Patients in Long-term Care by Stage

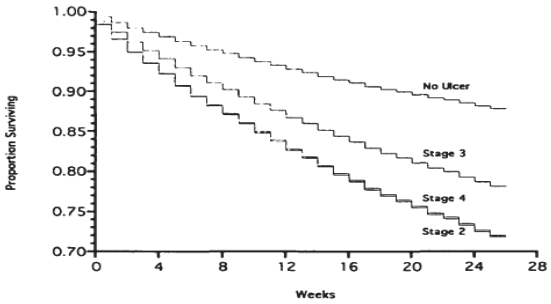


Figure 1. Proportional hazards estimates of the proportion of patients surviving, by weeks from baseline, for the no pressure ulcer and the ulcer stages 2, 3, and 4 subject groups, *without* adjustment for other factors.

Incidence and instantaneous rate

- **Instantaneous rate:**
  - Newton defined an instantaneous rate as the change in S(t) as the length of the time interval (Δt) becomes infinitesimally small.
  - The derivation of S(t)= $\frac{d}{dt}S(t)$

    $\frac{d}{dt}S(t) \approx \frac{S(t+\Delta t)-S(t)}{\Delta t}$

    = slope of a straight line between S(t) and S(t+Δt)

Incidence and instantaneous relative rate

- **Instantaneous relative rate**
  - Newton's instantaneous rate is rarely used to describe mortality or disease data, because it does not reflect risk.
  - A rate of 10 deaths per month in a community of 1,000 individuals indicates an entirely different risk than the same rate in a community of 100,000.
  - To measure risk, a relative rate is created, where

    Instantaneous relative rate = $h(t) = {-\frac{d}{dt}S(t)}\Big/{S(t)}$

Applications with Incidence

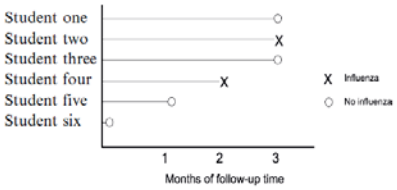
- An instantaneous relative rate is called a hazard rate in human populations, sometimes called the force of mortality (an instantaneous rate of death) or relative velocity from the viewpoint of physics.
- To estimate the instantaneous rate of $h(t)$, we must know the exact form of the survival function $S(t)$.

Incidence-Hazard in Public Health
(Newton's Second Law)

F=ma
=m ($\Delta v/\Delta t$)

$$Total\ Person-time$$

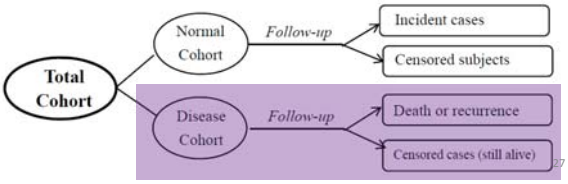
Fig. 1.1 Diagram of individual risk time and disease status



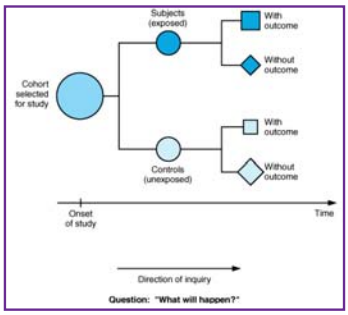
Incidence and Study Design

e-book: **01**. Basic & Clinical Biostatistics Ch2. 36-39;
02. Basic Biostatistics for Geneticists and Epidemiologists Ch3. p55-58
04. Statistical Methods in Cancer Research Ch2. p42-49

- The concept of incidence is highly related to relative instantaneous rate (proposed since Newton in 17th Century)
- **Study Design**:



2. Incidence



- **Estimate**:

$$= \frac{Total\ Events\ (i.e.\ Parkinson\ Dx)}{Total\ Person-time}$$

  - It represents relative instantaneous rate (force of morbidity) of yielding events (Parkinson disease) per unit time (years). The higher the incidence, the higher the risk for a specific population.

- **Usefulness**:
  - The incidence is a measure for elucidating the etiological aspect of disease of interest.

Original Research

Prevalence, incidence, and mortality of PD

A door-to-door survey in Ilan County, Taiwan

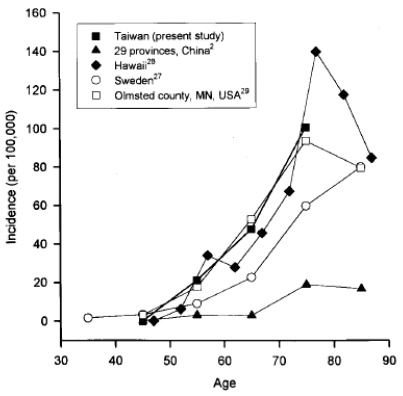
R.C. Chen, MD; S.F. Chang, MD, MPH; C.L. Su, MD; T.H.H. Chen, DDS, PhD; M.F. Yen, MS; H.M. Wu, BS; Z.Y. Chen, MD; and H.H. Liou, MD, PhD

The incidence per 100,000 of PD in Ilan

Age, y	Men		Women		Both sexes	
	No. of cases	Incidence per 100,000	No. of cases	Incidence per 100,000	No. of cases	Incidence per 100,000
40–49	0	0.0	0	0.0	0	0.0
50–59	2	25.4	1	12.0	3	18.5
60–69	3	45.4	3	49.6	6	47.4
70–79	3	98.1	3	102.5	6	100.2
80+	0	0.0	0	0.0	0	0.0
Total (40+)	8	32.5	7	27.8	15	30.1
Age-adjusted (40+)		30.5		27.0		28.7
Age-adjusted (for all age)		11.1		9.8		10.4

Original Research

Figure 4. Comparison of age-specific incidence rates of PD reported in China, Sweden, Hawaii, Olmsted County, Minnesota, and Taiwan. The incidence rates in Taiwan are much higher than those reported in mainland China but closer to those in Western countries.



Incidence of Pressure Ulcers

The number of new pressure ulcers in people without an ulcer at baseline

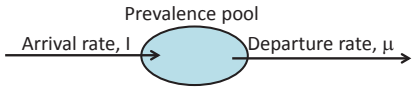
Direct measure of quality of care and the identification of causative factors for pressure ulcer development.

31

3. P/I (Prevalence/Incidence) ratio

e-book **03**. Epidemiology and Biostatistics Ch1 p5-11

- **Prevalence Pool**:



  - Suppose a population with size N consists of *m* prevalent Parkinson disease cases. In a cross-sectional survey, prevalence (*P*) is estimate as $P = \frac{m}{N}$
  - In a steady population (i.e. inflow = outflow), we have the following balance equation in a small time inter (Δt)

$I \times (N - m) \times \Delta t = \mu \times m \times \Delta t$ ➔ $\frac{m}{N-m} = \frac{I}{\mu}$

If $N \gg m, N - m \cong N$

$P(Prevalence) = \frac{I(Incidence)}{\mu}$ ➔ $\frac{P}{I} = \frac{1}{\mu} = \overline{D}$ (Average Duration)

32

3. P/I (Prevalence/Incidence) ratio

Usefulness: This indicator is used to denote the average duration of disease.

It has several applications. In our example of Parkinson disease, the P/I is used to reflect the quality of treatments or therapies.

33

Original Research

Table 4 *Prevalence, incidence, and ratio of PD in community-based studies*

Location	Prevalence per 100,000	Incidence per 100,000	Prevalence/ incidence
Rochester, MN[34]	187.0	20.0	9.4
Carlisle, England[35]	113.0	12.0	9.4
Iceland[36]	162.0	16.0	10.1
Turku, Finland[37]	120.1	15.0	8.0
Yonago, Japan[38]	80.6	10.0	8.6
Sardinia, Italy[39]	65.6	4.9	13.4
Benghazi, Libya[40]	31.4	4.5	7.0
Ferrara, Italy[41]	164.7	10.0	16.5
China[2]	18.0	1.9	9.5
Östergötland, Sweden[27]	115.0	11	10.5
Ilan, Taiwan (current study)	130.1	10.4	12.5

The Relationships Among Incidence, Prevalence, and Duration of Disease: Asthma

Age	Annual Incidence	Prevalence	Duration= Prevalence/ Annual Incidence
0-5	6/1,000	29/1,000	4.8 years
6-16	3/1,000	32/1,000	10.7 years
17-44	2/1,000	26/1,000	13.0 years
45-64	1/1,000	33/1,000	33.0 years
65+	0	36/1,000	33.0 years
Total	3/1,000	30/1,000	10.0 years

35

二、 辛普森矛盾故事與護理流行病學

1. 辛普森矛盾 (Simpson's Paradox) 及干擾

	昨天				今天	
	A 桌		B 桌		A、B 混合	
	黑	灰	黑	灰	黑	灰
適合	9	17	3	1	12	18
不適合	1	3	17	9	18	12
	10	20	20	10	30	30
	90%	85%	15%	10%	40%	60%

上表之情境說明如下:

假如一個人走進一賣帽子之商店,店中的帽子分置於兩個桌子: A桌放著30頂帽子,其中20頂是灰色而10頂是黑色;著他試戴A桌的帽子之後發現10頂黑色有9頂可以戴，20頂灰色中17頂灰色可以戴，也就是90%黑色，85%黑色及10%灰色可以戴. B桌也有30頂帽子, 有20頂為黑色,10頂為灰色, 試戴之後黑色合適的比例為15%, 灰色合適的比例為10%. 但在他決定買哪一頂之前這個店已打烊，於是隔天這個客戶再到這家商店來買，這次他發現所有帽子全部放在同一桌上，現在有30頂黑色及30頂灰色之帽子，試戴之後他發現合適的比例黑色為40%(12/30)，灰色為60%(18/30)

(1) 矛盾之處為何?

第一天不論A、B桌分開看都是黑色比較適合(A桌: 黑色比上灰色為90%比上85%; B桌:黑色比上灰色為15%比上10%), 但第二天混合時(A+B)卻是灰色適合的比例較高 (黑色比上灰色為40%比上60%)。

帽子店將帽子分桌自然有其意涵, 我們可以合理的認為, 為了方便選

擇, 店中將不同大小的帽子分類並分開放置, 因此A桌以及B桌其實代表了帽子大小的不同, 若A桌代表了大號帽子而B桌代表了小號帽子, 該名客戶適合戴大號帽子,因此A桌合適比例高於B桌; 但A桌中又以灰帽為多,B桌以黑帽為多, 所以大號帽子多是灰色. 如此一來, 帽子的大小與合適與否有關也與顏色有關, 因在上例中可以說帽子的大小干擾了顏色對於合適比例的影響.

此矛盾就是所謂的干擾因子造成的。例如1980~1990年代尚未知道HPV為子宮頸癌的危險因子時，認為接觸性伴侶人數、抽菸、服用避孕藥等等是子宮頸癌的危險因子. 後來發現HPV時,將HPV放進模式中分析發現有HPV成為子宮頸癌的可能性是沒有HPV的20倍.當考慮到HPV時,之前認為的危險因子(接觸性伴侶人數或服用避孕藥等)對於成為子宮頸癌的可能性都變成無顯著影響.也就是說當我們找到真正的原因時,這些原本有影響的因子影響減少了，這因子就是所謂的干擾因子。

(2) 自變數(X), 干擾因子(Z), 依變數 (Y: 結果)

桌子只是一個象徵，在統計裡是所謂的潛在變項，以A、B桌來講可能A桌的大尺寸帽子較多,剛好進來的客人頭較大，便會是A桌的適合度較高；而混合來看,灰色的大尺寸帽子較多,混和時便會是灰色較適合。在辛普森矛盾的例子中,帽子尺寸會影響帽子顏色的適合度，帽子尺寸便是所謂的干擾因子. 我們可以表示成:

Y：帽子是否合適

X：帽子的顏色

Z：帽子的尺寸(桌子)

(3) 利用甚麼樣的測量來描述辛普森矛盾？

百分比、勝算比(Odds Ratio, OR)、相對危險性(Relative Risk, RR)等皆可. 針對各因子分類比較這些測量, 例如在普森矛盾的例子中, 分桌別描述合適的比例相當於依照尺寸大小比較帽子顏色對於適合與否的影響為何.

2. 實證醫學之干擾因子的概念

(參考文獻：Grimes DA & Schulz KF: Bias and casual association in observational research. Lancet 2002; 359:248-252. 為例, 對於研究子宮內避孕器是否與輸卵管發炎有關)

(1) 若我們想研究醫療人員在兩組病人(Group A和Group B)打針成功率是否.

(A) 變項的操作型定義

在一先導性研究使用非隨機分派對照試驗設計，其三個變項如下：

Y: 成功 (Y=1) 與 失敗(Y=0)

X: Group A 與 Group B

Z: 護士與實習生

(B) 資料及結果

	Group A	Group B	
成功	30	24	54
失敗	30	16	46
	60	40	100

成功率: Group B (60%) > Group A (40%)

由於護理與實習生對於打針成功率也可能有影響, 因此研究者對於不同的護理實習生的治療成功率作分析:

	護士		實習生	
	Group A	Group B	Group A	Group B
成功	8	21	12	3
失敗	2	9	18	7
	10	30	30	10

Group A (80%) > Group B(70%), Group A (40%) > Group B (30%)

(C) 解釋

1. 實習生的成功率低於護士。

2. Group A中實習生的比例較Group B中來的高。

3. 醫療人員型態干擾病人組別的成功率。

(D) 在考慮醫療人員型態的情況下，Group A有較佳的成功率。

(E) 如同在健康促進研究一例, 醫療人員型態對於打針成功產生干擾, 是因為醫療人員型態在Group A與Group B的比例不同, 而醫療人員型態本身也會影響到打針成功的成功率, 因此可以利用隨機分派對照試驗設計(randomized controlled trial, RCT)達到平衡醫療人員型態及其他可能的干擾因子在兩個組別之間的分佈.

(2) 若我們想研究: 在考慮到年齡的影響下,某健康促進計畫能否提升慢跑效能？

(A) 變項之操作型定義 (operational definition)

X: 介入組 (有參加健康促進計畫, intervention, I) vs.
非介入組 (未參加健康促進計畫, No intervention, I)

Y: 結果: 是/否 能在30分鐘跑完5000公尺

Z: 干擾因子:年齡

(B) 資料及結果

	失敗	成功	總計	
$\bar{I}$	45	955	1000	4.5%
I	15	985	1000	1.5%
				RR=0.33

因介入組的失敗風險為非介入組的0.33倍 (相對風險比, RR=0.33); 若看百分比的絕對差異也可知道有訓練失敗率降低3%(4.5%降至1.5%)

我們知道年齡對於是否能在 30 分鐘內跑完 5000 公尺也會有影響, 因此我們

把受試者依據年齡的不同進行比較, 結果如下表:

(a) Age 30-39

	Fail	Success	Total		
$\bar{I}$	3	297	300	1%	
I	9	891	900	1%	RR=1

(b) Age 40-49

	Fail	Success	Total		
$\bar{I}$	42	658	700	6%	
I	6	94	100	6%	RR=1

在各別的年齡層中 (30-39 歲 以及 40-49 歲), 介入組與非介入組的失敗比例都相同, 相對風險比也都是 1, 所以在考慮年齡的不同下, 介入 (參加健康促進計畫)並無法達到降低失敗 (無法在30分鐘內跑完5000公尺) 的目的.

(C) 解釋

1. 有介入組的年齡多較輕 (較多人年齡為30-39歲), 30-39 歲參與健康促進計畫的比例比上 40-49 歲參與健康促進計畫的比例為 90% 比上 30%.
2. 較年輕的組別失敗率較低(1% vs 6%)
3. 年齡會干擾訓練對於30分鐘跑完5000公尺的失敗率

(D) 結論

在考慮年齡的情況下，健康促進計畫無法提升慢跑的效能。

(E) 隨機控制試驗設計

上述干擾作用的產生, 是因為年齡 (干擾因子) 在介入組與非介入組的比例不同, 而年齡本身也會影響到評估結果,所以可以利用隨機分派對照試驗設計 (randomized controlled trial, RCT) 達到平衡兩組間年齡分佈的目的.

*辛普森矛盾的條件：干擾因子必定與研究中有興趣的因、果皆有關係。

(3) 因用藥或住院造成之干擾 (Confounding by indication)

“Confounding-by-indication” 是在觀察性研究探討藥物或疫苗的效益時經常遭遇的一種干擾因子型態。此種型態之意義在於比較使用藥物或非使用藥物之結果差異常源自於: 醫師決定開此藥物或施打疫苗之其他原因(如: 開藥或疫苗施打前病人之基本特徵(如是否有慢性病存在))而非藥物或疫苗本身所造成。

在 1994 年由 Nichol KL 等人發表在 NEJM 文章之表一可見: 注射疫苗組在 1992 到 1993 年間之罹患慢性病之情形(例如冠心症)高於不注射疫苗組。因此，在進行兩組間由於鬱血性心臟病而住院之比較時須納入兩組在此基礎條件上不一致的調整。在調整此一不一致的情形後，兩組間的差異由原來的 1.8/1,000 降到 0.2/1,000。(Reference: Nichol KL et al, NEJM 1994 (22) 778-784).

Table 1. Base-line characteristics of the study subjects, according to study period and vaccination status.

Characteristic	*1992-1993*		
	vaccine	**no vaccine**	**p value**
Outpatient diagnosis during previous 12 mo			
Coronary heart disease	17.1	11.5	<0.001
Chronic lung disease	10.1	6.4	<0.001
Diabetes	11.6	7.9	<0.001
Vaculitis or rheumatologic disaese	2.1	1.3	<0.001
Demantia or stroke	2.4	4.5	<0.001

Table 2. Hospitalizations per 1000 elderly enrollees for Pneumonia and Influenza, all acute and chronic respiratory conditions, and congestive heart failre among vaccine recipients and nonrecipients, according to Influenza season.

Cause of hospitalization	*1992-1993*	
	vaccine	**novaccine**
Congestive heart failure		
unadjusted	4.9	3.1
adjustedd	4.1	3.9
difference (95% CI)	0.2 (-1.9 to +2.2)	
P value	0.88	

3. 干擾因子的影響

(1) 高估 (overestimation)

Z(+)			Z(-)			總和 (aggregated)		
	X(+)	X(-)		X(+)	X(-)		X(+)	X(-)
Y(+)	80	20	Y(+)	10	30	Y(+)	90	50
Y(-)	80	20	Y(-)	40	120	Y(-)	120	140
OR=1			OR=1			OR=2.1(粗率,crude)		

原本兩組沒差異(OR=1)，合在一起後變有差異(OR≠1)，稱為正向干擾(positive confounding)

(2) 低估 (underestimation)

Z(+)			Z(-)			總和 (aggregated)		
	X(+)	X(-)		X(+)	X(-)		X(+)	X(-)
Y(+)	120	10	Y(+)	80	190	Y(+)	200	200
Y(-)	160	30	Y(-)	40	170	Y(-)	200	200
OR=2.25			OR=1.79			OR=1 (粗率,crude)		

原本兩組有差異(OR≠1)，合在一起後變沒差異(OR=1)，稱做負干擾 (negative confounding)

正干擾 (positive confounding) 的例子如同之前健康促進計畫的研究一般: 介入(參與健康促進計畫) 實際上對於結果並無影響, 但由於干擾因子 (年齡) 的存在, 使的總合後所計算的粗率表現出有差異的結果, 而負干擾 (negative confounding) 的例子則相反: 利用總和所計算的粗率看似介入對於結果並無影

響, 可是一旦針對干擾因子進行分層分析, 介入因子的效用就顯現出來; 對於負干擾的解釋為, 當干擾因子 (Z) 存在的情況下,介入因子 (X) 才會對結果 (Y) 發生影響, 在免疫學中常有類似的情形: T 細胞的作用機轉 (Y)是否會被某特定抗原 (X) 所啟動端看 MHC (major histocompatibility complex) 是否存在並完成抗原處理得前置作業並將該抗原訊息傳遞給 T 細胞 (Z) 而定.

(3) 無干擾因子

	Z(+)		Z(-)		總合 (aggregated)	
	X(+)	X(-)	X(+)	X(-)	X(+)	X(-)
Y(+)	50	30	30	20	80	50
Y(-)	40	60	90	130	130	190
	OR=2.5		OR=2.17		OR=2.34	

相對風險比 (relative risk, RR) 與勝算比 (odds ratio, OR)

假設研究中收集到資料整理如下

	X	$\bar{X}$
Y	a	b
$\bar{Y}$	c	d

上表中, X 代表暴露於某一因子 (暴露組, 如使用子宮內避孕器,參與健康促進計畫,新治療等), 而 $\bar{X}$ 即代表未暴露於該因子 (非暴露組, 如未使用子宮內避孕器, 未參與健康促進計畫或一般治療); Y 代表發生定義的結果 (如發生輸卵軟管發炎,無法在 30 分鐘內跑完 5000 公尺或治療成功), 而 $\bar{Y}$ 代表沒有發生定義的結果 (如未發生輸卵軟管發炎,在 30 分鐘內跑完 5000 公尺或治療失敗)

勝算比計算如下:

$$\text{OR}=\frac{a/\text{c}}{\text{b}/d}=\frac{ad}{bc}$$

$$\text{疾病勝算比 (disease OR)}=\frac{\text{disease odds for X}}{\text{disease odds for }\bar{\text{X}}}=\frac{a/c}{b/d}=\frac{ad}{bc}$$

$$\text{暴露勝算比 (exposure OR)}=\frac{\text{exposure odds for Y}}{\text{exposure odds for }\bar{\text{Y}}}=\frac{a/b}{c/d}=\frac{ad}{bc}$$

若 X 與 Y 無關, 則暴露組與非暴露組對於發生事件與不發生事件的勝算都相同, 則勝算比為 1, 因此可以設立虛無假說以及對立假說如下:

H_0：OR=1　v.s H_1:OR≠1$\begin{cases}\text{OR}>1\\ \text{OR}<1\end{cases}$

對立假說的部分可能大於 1(暴露組發生事件的勝算高於非暴露組), 也可能小於 1 (暴露組發生事件的勝算低於非暴露組), 須依照實際情況設立.

暴露組 (X) 與非暴露組 ($\bar{X}$) 是否發生事件 (Y) 可以視為兩個二項分佈:

$$\text{暴露組: } n_X=a+c,\ A\sim\text{Bino}(n_X, p_X)$$

$$\text{非暴露組: } n_{\bar{X}}=b+d,\ B\sim\text{Bino}(n_{\bar{X}}, p_{\bar{X}})$$

因此, 暴露組與非暴露組的勝算如下:

$$\text{暴露組勝算(odds for X): } p_X/1-p_X$$

$$\text{非暴露組勝算(odds for } \bar{X}\text{): } p_{\bar{X}}/1-p_{\bar{X}}$$

則勝算比 (odds ratio, OR) 為:

$$\text{OR}=\frac{p_X/1-p_X}{p_{\bar{X}}/1-p_{\bar{X}}}$$

將兩個參數(p_X、$p_{\bar{X}}$)轉換成一個參數(OR、RR)稱作"再參數化"(reparametrization), 藉由再參數化的過程, 可以免於估計 nuisance (未暴露組發生事件的機率或勝算) 只須得到相對的大小 (相對危險比 RR, 或勝算比, OR) .

註：nuisance, 討厭鬼

流行病學方法論及實驗 (The Methods of Epidemiology and Practices): 辛普森矛盾故事與護理流行病學

授課教師：陳秀熙 教授

台灣大學流行病學與預防醫學研究所

學習宗旨

Concept of Simpson paradox- 辛普森矛盾

Concept of Confounding-干擾因子

Applications to Nursing Epidemiology

辛普森矛盾與棒球統計

	Max			John		
	成功	出擊	平均	成功	出擊	平均
所有打點	90	200	0.45	18	38	0.47

Max < John

	Max			John		
	成功	出擊	平均	成功	出擊	平均
兩分打點	60	100	0.6	16	30	0.53
參分打點	30	100	0.3	2	8	0.25

Max > John

所謂「辛普森矛盾」(Simpson's Paradox)

Simpson, Edward H. (1951). "The Interpretation of Interaction in Contingency Tables". *Journal of the Royal Statistical Society, Series B* 13: 238–241.

Simpson's Paradox

	Yesterday				Today	
	Table A		Table B		A+B	
	Black	Gray	Black	Gray	Black	Gray
fit	9	17	3	1	12	18
ill-fit	1	3	17	9	18	12
	10	20	20	10	30	30
	90%	85%	15%	10%	40%	60%

(1) What is the paradox?

A + B: Black < Gray

A or B: Black > Gray

Simpson's Paradox and Confounding

	Yesterday				Today	
	Table A		Table B		A+B	
	Black	Gray	Black	Gray	Black	Gray
fit	9	17	3	1	12	18
ill-fit	1	3	17	9	18	12
	10	20	20	10	30	30
	90%	85%	15%	10%	40%	60%

(2) Independent variable (X), confounding factor (Z), and dependent variable (Y: outcome)

- X: **color** (Black vs Gray)
- Y: **fit** vs **ill-fit**
- Z: table ? -> **size**

(3) What's kind of measure for describing Simpson's paradox?

- Proportion of fit

Simpson's Paradox and Confounding: Successful injection rate

	Group A	Group B	
Success	20	24	44
Failure	20	16	36
	40	40	80

Success rate: Group B > Group A

Successful rate of Group A:
Successful rate of Group B:
Relative Risk=

Odds of Group A:
Odds of Group B:
Odds Ratio=

Y: **Success** (Y=1), failure (Y=0)
X: Group A and B
Z: **Type of medical personnel**

Simpson's Paradox and Confounding: Successful Injection Rate

Nurses	Group A	Group B
Success	8	21
Failure	2	9
	10	30

Group A (80%) > Group B (70%)

Relative risk : $0.8/0.7 = 1.14$

$Odds\ ratio: \frac{8\times 9}{2\times 21} = 1.71$

Interns	Group A	Group B
Success	12	3
Failure	18	7
	30	10

Group A (40%) > Group B (30%)

Relative risk : $0.4/0.3 = 1.33$

$Odds\ ratio: \frac{12\times 7}{18\times 3} = 1.56$

Y: Success (Y=1), failure (Y=0)
X: Group A and Group B
Z: Type of personnel

7

Simpson's Paradox and Confounding: Efficacy of new treatment

(C) Interpretation

1. Interns had poorer outcomes than nurses
2. The proportion of interns in the Group A is larger than that in the Group B
3. Type of medical personnel confounds the successful rate of the two groups

8

Simpson's Paradox and Confounding: Efficacy of new treatment

(D) The Group A has better outcome than Group B after type of medical personnel is taken into account

(E) A randomized controlled trial is adopted to balance the distribution of type of medical personnel other possible confounders.

9

Simpson's Paradox and Confounding-Evaluation of Health Promotion Programme

- Can certain health promotion program improve the performance of jogging after making allowance for age?

		Performance: Fail	Performance: Success	
Intervention	No(B)	45	955	1000
	Yes(A)	15	985	1000
		60	1940	2000

$RR = \frac{P_A}{P_B} = 0.33$

Relative Risk
危險對比值

10

Strata 1 30-39		Performance: Fail	Performance: Success	
Inter-vention	No (B)	3	297	300
	Yes (A)	9	891	900
		12	1188	1200

$RR = \frac{P_A}{P_B} = 1$

Strata 2 40-49		Performance: Fail	Performance: Success	
	No (B)	42	658	700
	Yes (A)	6	94	100
		48	752	800

$RR = \frac{P_A}{P_B} = 1$

11

Simpson's Paradox and Confounding: Postoperative Infections

(C) Interpretation

1. Invention group (type A) has more participant aged 30-39 (90% vs 30%)
2. The younger group has lower fail rate (1% vs 6%)
3. **Age confounds** the effect of intervention on the reduction of failure rate

12

Simpson's Paradox and Confounding: Postoperative Infections

(D) Conclusion

Intervention of health promotion do not reduce failure rate after taking into account age

(E) **Randomized controlled trial design**

Using randomization design can balance age distribution between the two groups.

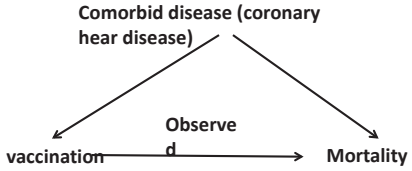
13

Confounding by indication

SPECIAL ARTICLE

THE EFFICACY AND COST EFFECTIVENESS OF VACCINATION AGAINST INFLUENZA AMONG ELDERLY PERSONS LIVING IN THE COMMUNITY

K.L. Nichol, M.D., M.P.H., K.L. Margolis, M.D., M.P.H., J. Wuorenma, R.N., B.S.N., and T. Von Sternberg, M.D.



Reference: Nichol KL et al, NEJM 1994 (22) 778-784

14

Confounding by indication

Table 1. Base-line characteristics of the study subjects, according to study period and vaccination status.

Characteristic	1992-1993		
	vaccine	no vaccine	p value
Outpatient diagnosis during previous 12 mo			
Coronary heart disease	17.1	11.5	<0.001
Chronic lung disease	10.1	6.4	<0.001
Diabetes	11.6	7.9	<0.001
Vaculitis or rheumatologic disease	2.1	1.3	<0.001
Dementia or stroke	2.4	4.5	<0.001

Table 3. Hospitalizations per 1000 elderly enrollees for Pneumonia and Influenza, all acute and chronic respiratory conditions, and congestive heart failre among vaccine recipients and nonrecipients, according to Influenza season.

Cause of hospitalization	1992-1993	
	vaccine	novaccine
Congestive heart failure		
unadjusted	4.9	3.1
adjustedd	4.1	3.9
difference (95% CI)	0.2 (-1.9 to +2.2)	
P value	0.88	

Reference: Nichol KL et al, NEJM 1994 (22) 778-784

15

Influence of confounding

(1) Overestimation

	Z(+)		Z(-)		Aggregated	
	X(+)	X(-)	X(+)	X(-)	X(+)	X(-)
Y(+)	80	20	10	30	90	50
Y(-)	80	20	40	120	120	140
	OR=1		OR=1		OR=2.1(Crude)	

(2) Underestimation

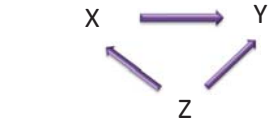
	Z(+)		Z(-)		Aggregated	
	X(+)	X(-)	X(+)	X(-)	X(+)	X(-)
Y(+)	120	10	80	190	200	200
Y(-)	160	30	40	170	200	200
	OR=2.25		OR=1.79		OR=1(Crude)	

(3) No confounding

	Z(+)		Z(-)		Aggregated	
	X(+)	X(-)	X(+)	X(-)	X(+)	X(-)
Y(+)	50	30	30	20	80	50
Y(-)	40	60	90	130	130	190
	OR=2.5		OR=2.17		OR=2.34	

16

Simpson's Paradox: Solutions



消除 Z 之干擾

1. 研究設計使用**隨機分配(Randomization)試驗**
2. 研究設計使用**限制 (Restriction)或配對 (Matching)**
3. 統計資料分析使用**分層分析(Stratified Analysis) 或數學迴歸模式 (Regression Model)**

17

Overall Results

	失敗	成功		
非介入(X=0)	45	955	1000	4.5%
介入(X=1)	15	985	1000	1.5%

3倍

$$logit[P(Y=1)] = \beta_0 + \beta_1 X$$

$$\beta_0 = -3.055, \beta_1 = 1.1294$$

$\exp(\beta_1)=\exp(1.13)=3.09$

18

Stratified Analysis: Age 30-39

	失敗	成功		
非介入	3	297	300	1.0%
介入	9	891	900	1.0%

1倍

$logit[P(Y=1)] = \beta_0 + \beta_1 X$; $\beta_0 = -4.5951, \beta_1 = 0$, $\exp(\beta_1)=\exp(0)=1$

Stratified Analysis: Age 40-49

	失敗	成功		
非介入	42	658	700	6.0%
介入	6	94	100	6.0%

1倍

$logit[P(Y=1)] = \beta_0 + \beta_1 X$; $\beta_0 = -2.7515$, $\beta_1 = 0$, $\exp(\beta_1)=\exp(0)=1$

19

辛普森矛盾與孔子人生哲學

「三十而立」
「四十而不惑」 } Confusing（困惑）= Confounding（干擾）

「五十而知天命」

「六十而耳順」

「七十從心所欲，不踰矩」

20

三、 解決辛普森矛盾之護理魔術

一、控制干擾因子的方法 (參考 ebook_03, Ch 10, p.101-111)

控制干擾因子之最佳研究設計為隨機控制試驗(Randomized Controlled Trial, RCT) 若非 RCT 之設計即屬於觀察性研究,將會產生許多干擾因子及其他偏差，進而影響因果關係之推論，因此控制干擾因子方法變得相當複雜，其範圍從研究設計至資料分析。

1. 研究設計

(1) 隨機控制試驗(Randomized Controlled Trial, RCT)

(1) 配對 (matching)：類似實驗設計 RBD 中所使用之 Block Matching，但觀察性研究之配對沒有隨機分配(Randomization)之過程，因此仍需調整其他干擾因子。

通常用來配對的因子也是干擾因子之一，以課本為例，為研究新型抗生素 supramycin 和傳統使用之抗生素 amoxicillin 相比是否較容易導致患者長疹子，已知患者之濕疹病史為一干擾因子，故將暴露組(exposed)與非暴露組(unexposed)之個案，依照其濕疹病史進行配對，若為 1:1 (暴露組:非暴露組)配對，則配對方式如下表所示:

Supramycin 組(暴露組)	Amoxicillin 組(非暴露組)
個案 1: 使用 Supramycin, **無**濕疹病史	使用 Amoxicillin, **無**濕疹病史
個案 2: 使用 Supramycin, **有**濕疹病史	使用 Amoxicillin, **有**濕疹病史

個案 3: 使用 Supramycin, **無**濕疹病史	使用 Amoxicillin, **無**濕疹病史
個案 4: 使用 Supramycin, **無**濕疹病史	使用 Amoxicillin, **無**濕疹病史
……	……

重複此過程，盡可能的將暴露組中的每筆個案都進行配對。

分析前 1000 組配對結果，可發現暴露組中有 98 名個案有濕疹病史(占 9.8%)，相對的，非暴露組中也有 98 名個案(占 9.8%)有濕疹病史。利用配對的方式，使得配對因子-濕疹病史(干擾因子)在兩組中的分佈情形一致，達到類似隨機分派試驗的效果。利用配對過程消除了干擾因子(濕疹病史)的影響。

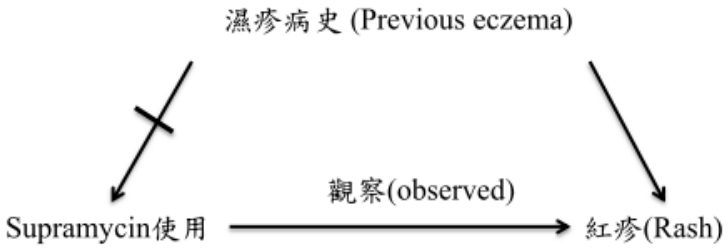
除了 1:1 配對外，亦可進行 1:n 配對，每個暴露組個案可配對 n 個非暴露組個案以增加統計檢定力。在干擾因子的數目方面，亦可針對多個干擾因子同時進行配對，以前述抗生素之研究為例，若除了濕疹病史外，用藥過敏史、家族濕疹病史亦為干擾因子，則配對時可同時考慮這三個干擾因子。若暴露組個案 1 為 使用 Supramycin, 無濕疹病史，無用藥過敏史、無家族濕疹病史，則與之配對的非暴露組個案為: 使用 Amoxicillin, 無濕疹病史，無用藥過敏史、無家族濕疹病史。多個干擾因子的配對方式，不限於 1:1 配對，亦可使用在 1:n 配對上。

除了上述優點之外，配對亦有其限制與缺點：a) 配對過程在一開始的研究設計階段即決定好，故若後續才想要增加配對因子(例如社經地位)有其困難度；b) 若同時針對多個因子配對，當樣本數不夠大時，可能不易配對到具有相同配對因子的個案； c) 用來進行配對的因子無法再用來作為風險因子分析其對結果的影響。

(2) 限制 (restriction, 例如: 若年齡是干擾因子, 則將收案的年齡限定在一定的範圍之內(如 35 歲以下，老人排除在外))，如同 RCT 之 Eligibility。

以前述新型抗生素 supramycin 和傳統使用之抗生素 amoxicillin 相比是否較容易導致患者長疹子之研究為例，最簡單的方法為將有濕疹病史的個案從

研究對象中去除。當研究對象中均為無濕疹病史之個案時，濕疹病史不會再干擾 supramycin 之使用情形(暴露因子)與起疹子之情形(結果)。



使用限制的方式雖然相當簡單且易於了解，並可消除特定干擾因子，但亦有其限制與缺點: a) 造成統計檢定力下降：將 1136 名有濕疹病史之個案去除後，樣本數隨之減少，若同時想要消除多個干擾因子，例如濕疹病史、用藥過敏史與家族濕疹病史，則可能由於去除過多個案數，而使得統計檢定力下降更多； b) 研究結果缺乏外推性 (generalizability)：在消除多個干擾因子後，即使剩餘之樣本數仍夠多，但此研究結果僅限於用來推論無濕疹病史、無用藥過敏史及無家族濕疹病史之個案，而不適合用在一般臨床個案上，因而缺乏外推性。

2. 資料分析

(1) 分層分析 (stratified analysis)

(2) 迴歸模式 (Regression model)

分層分析

1. 粗率(Crude Rate)分析

有關世代追蹤分析，若不考慮干擾因子，其分析方法和 RCT 兩組差異之分析相同，結果端視所選擇之結果(連續變項，二元變項，及存活時間)而定。在世

代追蹤研究中，追蹤暴露組以及未暴露組在一段時間下的得病狀況，因此可以使用相差風險性 (risk difference, RD)或是相對風險性(relative risk, RR)。

(1) 相差風險性(Risk Difference, RD)

	E	$\overline{E}$	
D	a	b	m_1
$\overline{D}$	c	d	m_2
	n_1	n_2	n

(A) 點估計值 (point estimate) $\hat{RD} = \hat{P}_1 - \hat{P}_0$

(B) 信賴區間 (confidence interval)

(a) 大樣本理論，利用泰勒分析(Taylor series method)

$$\hat{RD} \pm Z_{\frac{\alpha}{2}} \hat{SE}(\hat{RD}) = \hat{RD} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_0(1-\hat{P}_0)}{n_2}}$$

(b) Test-based method：推導過程見補充教材。

(C) 範例

在兒茶酚胺激素(catecholamine, CAT)與冠狀動脈血管疾病(coronary heart disease, CHD)之相關性研究中，資料結構如下，

	High CAT	*Low CAT*	
CHD	27	44	71
No CHD	95	443	538
Total	122	487	609

比較高濃度(P_1)與低濃度(P_0)的茶酚胺激素個體發生冠狀動脈血管疾病的風險：

$\hat{P}_1 = 0.221$ $\hat{P}_0 = 0.090$

$$\hat{RD} = \hat{P}_1 - \hat{P}_0 = 0.131$$

(a) 利用泰勒分析(Taylor series method)

$$S\hat{E}(\hat{RD}) = \sqrt{\frac{0.22(1-0.22)}{122} + \frac{0.090(1-0.090)}{487}} = 0.040$$

95%CI=(0.053-0.209)

(b) 利用 Test-based method (推導過程請參見補充教材)

$$\chi^2_{MH} = \frac{(609-1)(27\times 443 - 44\times 95)^2}{122\times 487\times 71\times 538} = 16.22$$

95% CI= $0.131 \pm 1.96\sqrt{(0.131)^2 / 16.22} = (0.067, 0.195)$

以上兩種方法 (Taylor series method 以及 test-based method), 若樣本數小, 則以 Taylor series method 較適合。

(2) 風險比值或相對風險 (Risk Ratio or Relative Risk)

風險比值或相對風險可以用來表現相關性的強度,因此為流行病學所關心。暴露組比上未暴露組發生疾病風險的風險程度可以用倍數來表示。

	E	$\overline{E}$	
D	a	b	m_1
$\overline{D}$	c	d	m_2
	n_1	n_2	n

(A) 點估計值 (point estimate)

$$\hat{RR} = \hat{P}_1 / \hat{P}_0$$

(B) 信賴區間 (confidence interval)

由於相對風險比呈現偏態分佈 (RR 不可能小於 0, 因此呈現右偏分佈 (right skewed) 因此會先取對數($\log RR$)求其信賴區間，然後再 Anti-log 轉換回來。

(a) 大樣本理論，利用泰勒分析(Taylor series method)

$$\hat{\mathrm{SE}}\{\log(\hat{\mathrm{RR}})\} \approx \sqrt{\frac{1-\hat{P}_1}{n_1\hat{P}_1}+\frac{1-\hat{P}_0}{n_2\hat{P}_0}}$$

$$\log(\hat{RR}) \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{1-\hat{P}_1}{n_1\hat{P}_1}+\frac{1-\hat{P}_0}{n_2\hat{P}_0}}$$

(b) Test-based method：推導過程見補充教材。

2. 分層分析 (stratified analysis) (參考 ebook_03, Ch 10, p.103-105)

(1) 2×2 聯列表之干擾因子調整

假定我們有一個 k 個層別 (level) 的干擾因子 F，以下依照 F 將資料分為 k 個 2×2 表，在每個分層中, F 的效果都相同 (對 F 而言達到立足點平等) 再評估暴露(E)的 RR; 其中每個分層的 RR 或 OR 理論上應該要很接近(若不接近則表示 F 與暴露 E 之間有交互作用, 為效應修飾 (effect modification))，此外我們希望知道:

- 綜合每個分層的 RR 為多少，稱做調整 F 因子之後的 RR。
- 各層中的 RR 是否夠接近 (RR 一致, 表示 F 只有干擾效應而無修飾效應。

F_1

	E	$\overline{E}$	
D	a_1	b_1	m_{11}
$\overline{D}$	c_1	d_1	m_{21}
	n_{11}	n_{21}	n_1

F_2

	E	$\overline{E}$	
D	a_2	b_2	m_{12}
$\overline{D}$	c_2	d_2	m_{22}
	n_{12}	n_{22}	n_2

		F_i			F_k		
		E	$\overline{E}$		E	$\overline{E}$	
……	D	a_i	b_i	m_{1i}	a_k	b_k	m_{1k}
	$\overline{D}$	c_i	d_i	m_{2i}	c_k	d_k	m_{2k}
		n_{1i}	n_{2i}	n_i	n_{1k}	n_{2k}	n_k

(2) 調整之風險比(Adjusted Risk Ratio)

(A) 合併加權估計式 (Pooled estimate)

以 F 的第 i 個分層所估計之相對風險比 (risk ratio, RR)為

$$\hat{RR}_i = \frac{a_i / n_{1i}}{b_i / n_{2i}}$$

欲知道每個分層的 RR 貢獻多少則須知道 $\hat{RR}_i$ 變異數, 利用該層的估計值之變異數倒數作為加權值

$\log\hat{RR}$ 的變異數估計為

$$\hat{Var}\ \{\log\hat{RR}_i\} = \frac{c_i}{a_i n_{1i}} + \frac{d_i}{b_i n_{2i}}$$

註: 由於 RR 必定大於 0, 呈現右偏態分佈(positive-skewed distribution)

因此將 RR 取對數後(log(RR))呈對稱分佈, 才能利用常態分佈近似。我們利用 logRR 的變異數的倒數作為加權值(W_i)以對每個分層的對數風險比 (log(RR_i))進行合併加權 (pool)

$$W_i = \frac{a_i b_i n_{1i} n_{2i}}{a_i d_i n_{1i} + b_i c_i n_{2i}}$$

利用加權計算得到合併加權(pooled)的 logRR 為

$$\log \hat{RR}_w = \frac{\sum_{i=1}^{k} W_i \log(RR_i)}{\sum_{i=1}^{k} W_i}$$

將求得的對數風險比($\log(\hat{RR}_w)$) 取指數得到合併加權後的風險比 (Pooled $\hat{RR}_w$):

$$\hat{RR}_w = \exp\left[\log \hat{RR}_w\right]$$

(B) Mantel-Haenszel 估計式 (Mantel-Haenszel estimator)

Mantel-Haenszel (1959) 提出了另一種方法利用分層加權[$\left(\frac{n_{1i}}{n_i}\right)b_i$]過後所計算的 RR ；在每一個分層知道暴露及非暴露的比例，各自與 a_i、b_i 相乘，將每一個分層的估計值進行加權平均得到 Mantel-Haensz 的調整後之總和風險比。

此方法較合併加權估計方法好的原因是因為能妥善處理 2×2 表中出現 0 的情況，因為分層後 2×2 表中很有可能會出現 0，而合併加權估計方法中計算 RR 的過程 ad/bc 若 0 出現在分母則會產生無窮的問題，而 Mantel-Haenszel 則利用"加總"的步驟解決了分母出現 0 的情況。

$$\hat{RR}_{MH} = \frac{\sum_{i=1}^{k}\left(\frac{n_{1i}}{n_i}\right)b_i\left(\frac{a_i n_{2i}}{b_i n_{1i}}\right)}{\sum_{i=1}^{k}\left(\frac{n_{1i}}{n_i}\right)b_i}$$

以上的式子可重新表示為

$$\hat{RR}_{MH} = \frac{\sum_{i=1}^{k}\left(\frac{n_{2i}}{n_i}\right)a_i}{\sum_{i=1}^{k}\left(\frac{n_{1i}}{n_i}\right)b_i}$$

(C) 信賴區間

(a) 合併加權估計方法

$$\log\left(\hat{RR}_w\right) \pm Z_{\alpha/2}\sqrt{\left\{\log\left(\hat{RR}_w\right)\right\}}$$

(b) Mantel-Haenzel 方法

$$\log\left(\hat{RR}_{MH}\right) \pm Z_{\alpha/2}\sqrt{\left\{\log\left(\hat{RR}_{MH}\right)\right\}}$$

(D) 範例

在研究 catecholamine (CAT) 與心臟疾病(CHD)之關係的研究中，年齡的大小 (Age)及心電圖是否異常(ECG)為兩個可能的干擾因子

特定分層之 RR:

年輕組($Age < 55$),
心電圖無異常($ECG = 0$)

	High CAT	*Low CAT*	*Total*
CHD	1	17	18
No CHD	7	257	264
Total	8	274	282

$\hat{RR}_1 = 2.015(0.30 \sim 13.34)$

年輕組($Age < 55$),
心電圖有異常($ECG = 1$)

	High CAT	*Low CAT*	*Total*
CHD	3	7	10
No CHD	14	52	66
Total	17	59	76

$\hat{RR}_2 = 1.49(0.43 \sim 5.14)$

年老組($Age \geq 55$),
心電圖無異常($ECG) = 0$

	High CAT	*Low CAT*	*Total*
CHD	9	15	24
No CHD	30	107	137
Total	39	122	161

$\hat{RR}_3 = 1.88(0.89 \sim 3.95)$

年老組($Age \geq 55$),
心電圖有異常 ($ECG = 1$)

	High CAT	*Low CAT*	*Total*
CHD	14	5	19
No CHD	44	27	71
Total	58	32	90

$\hat{RR}_4 = 1.54(0.61 \sim 3.90)$

(a) 合併加權方法 (pooled method) (logit method)

$$W_1 = \frac{a_1 b_1 n_{11} n_{21}}{b_1 c_1 n_{21} + a_1 d_1 n_{11}} = \frac{(1)(17)(8)(274)}{(17)(7)(274) + (1)(257)(8)} = 1.075$$

$W_2 = 2.497, W_3 = 6.947, W_4 = 4.486$

i	W_i	$\hat{RR}_i$	$\log \hat{RR}_i$	$W_i \log(\hat{RR}_i)$
1	1.075	2.015	0.701	0.753
2	2.497	1.487	0.397	0.991
3	6.947	1.877	0.630	4.374
4	4.486	1.545	0.435	1.952

$$\log \hat{RR}_w = \frac{8.070}{15.005} = 0.538 \qquad \hat{RR}_w = 1.713$$

$$\hat{Var}\left\{\log(\hat{RR}_w)\right\} = \frac{1}{15.005} = 0.067$$

$$95\%CI = \exp(0.538 \pm 1.96 \times \sqrt{0.067}) = (1.032 \sim 2.840)$$

若不調整的 RR 為 2.46 (crude RR，當調整了年齡及 ECG 後的 RR 為 1.713，其值變小是因為有部分相關性被年齡及 ECG 所解釋掉了。

(b) Mantel-Haenzel 方法

i	a_i	b_i	m_{1i}	n_{1i}	n_{2i}	n_i	$\left(\frac{n_{2i}}{n_i}\right)a_i$	$\left(\frac{n_{1i}}{n_i}\right)b_i$
1	1	17	18	8	274	282	0.972	0.482
2	3	7	10	17	59	76	2.329	1.566
3	9	15	24	39	122	161	6.820	3.634
4	14	5	19	58	32	90	4.978	3.222
							15.099	8.904

$$\therefore \hat{RR}_{MH} = \frac{15.099}{8.904} = 1.696$$

$$Var\left\{\log\left(\hat{RR}_{MH}\right)\right\} = \left[\frac{(18)(8)(274)-(1)(17)(282)}{(282)^2} + \frac{(10)(17)(59)-(3)(7)(76)}{(76)^2} + \frac{(24)(39)(122)\text{-}(9)(15)(161)}{(161)^2} + \frac{(19)(58)(32)-(14)(5)(90)}{(90)^2}\right] \div (15.099)(8.904) = \frac{9.039}{134.442} = 0.067$$

$$95\%CI = 1.696\times,\div \exp\left(1.96\times\sqrt{0.067}\right) = (1.02, 2.82)$$

(c) Test-based method (推導過程請參見補充教材)

$$X^2_{MH} = \frac{\left[\sum_{i=1}^{k} a_i - \sum_{i=1}^{k} \frac{m_{1i}n_{1i}}{n_i}\right]^2}{\sum_{i=1}^{k} \frac{m_{1i}m_{2i}n_{1i}n_{2i}}{n_i^2(n_i-1)}} = \frac{(27-20.806)^2}{9.239} = 4.153$$

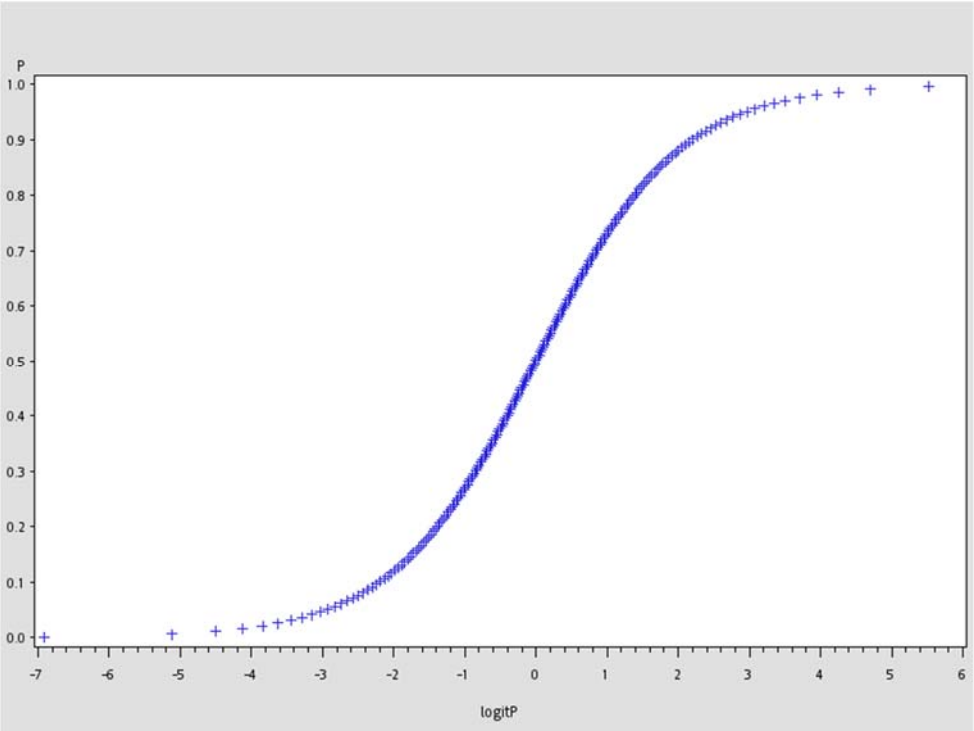
$$1.696\times,\div \exp\left(1.96\sqrt{\left(\log\left(\hat{RR}_{MH}\right)\right)^2 \Big/ X^2_{MH}}\right) = (1.021, 2.82)$$

(1) 羅吉斯迴歸(Logistic Regression Model) (參考 ebook_01, Ch 10, p.852-858; ebook 02, Ch 12, p.296-299; ebook 03, Ch 19, P.21-214)

在迴歸模型中連結 systematic component 與應變數，兩邊的值域需一致，由於二元變項(Y=1, 0)下，應變數(等式左側, $P(Y)$)的值域為 0 到 1，而等式右側(systematic component)的值域為負無窮到正無窮，故對應變數作 logit 變數轉換($\log(\frac{p}{1-p})$) 來連結等式右側(systematic component); 將 P 及 logitP 轉換列表如下:

P	logitP
0.001	-6.90675
0.006	-5.10998
0.011	-4.4988
0.016	-4.11904
0.021	-3.84201
0.026	-3.62331
0.031	-3.44228
0.036	-3.28757
⋮	⋮
0.981	3.94413
0.986	4.2546
0.991	4.70149
0.996	5.51745

將上表作圖成為:



這樣的曲線稱為"sigmoid curve", 廣泛運用在臨床研究上。 由於透過 logit 轉換連結等式兩側, 因此該模式稱為羅吉斯迴歸模式(Logistic Regression Model)，寫成：

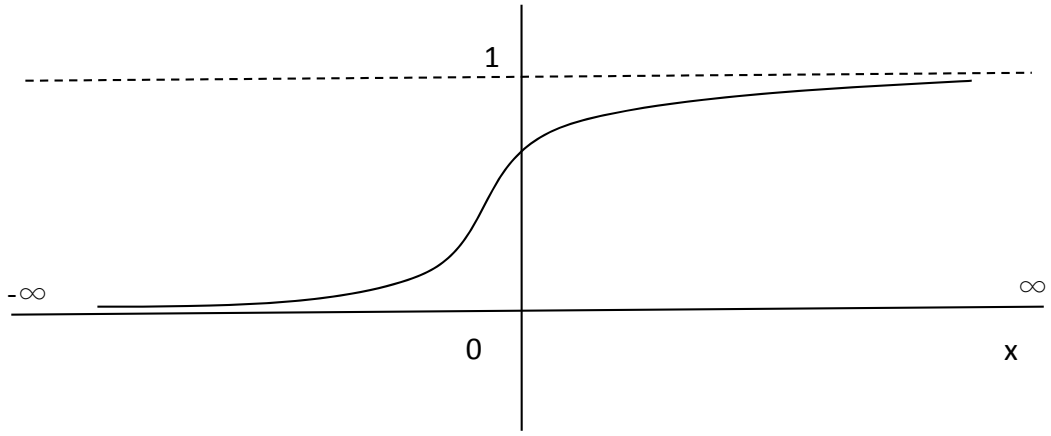
$$\log(\frac{P(Y=1)}{1-P(Y=1)}) = \text{logit}(\text{P}(\text{Y}=1)) = \text{logit}(\text{P}) = \beta_0 + \beta_1 X + \varepsilon$$

若為兩組比較, 則 X 代表了組別:

$$\begin{cases} x=1 & \log\frac{\hat{P}_1}{1-\hat{P}_1} = \hat{\beta}_0 + \hat{\beta}_1 \\ x=0 & \log\frac{\hat{P}_0}{1-\hat{P}_0} = \hat{\beta}_0 \end{cases}$$

同樣可以圖示為:

$$\text{logit(P)}, \qquad P = (\frac{\exp(\beta_0 + \beta_1 x)}{1+\exp(\beta_0 + \beta_1 x)})$$



模式中的 $\frac{P}{1-P}$ 即為勝算 (odds), 因此兩組的勝算比即為 $\frac{P_1/1-P_1}{P_0/1-P_0}$, 將迴歸模式帶入得到:

$$\frac{\hat{P}_1/1-\hat{P}_1}{\hat{P}_0/1-\hat{P}_0} = \frac{\exp\{\hat{\beta}_0+\hat{\beta}_1\}}{\exp\{\hat{\beta}_0\}},$$

因此

$$\log(\hat{OR}) = \log(\frac{\hat{P}_1 / 1-\hat{P}_1}{\hat{P}_0 / 1-\hat{P}_0}) = \hat{\beta}_1,$$

而勝算比為

$$\hat{OR} = \exp(\hat{\beta}_1).$$

比較兩組有無差異 ($H_0 : P_1 = P_0$) 等同於檢定 $H_0 : OR = 1$ 或 $H_0 : \beta_1 = 0$。

例：上述新治療方法對於 healing 及 infection 之影響，若所得到樣本如下：

	有療效	有療效	總和
新藥	40	20	60
舊藥	16	48	64
總和	56	68	124

$H_0 : \beta_1 = 0$ v.s. $H_1 : \beta_1 \neq 0$

利用羅吉斯迴歸得到 $\hat{\beta}_1$=1.7918，$\hat{OR} = \exp(1.7918) = 6$,

檢定統計值 $X^2 = 20.3$ 服從自由度為 1 的卡方分佈, P<0.0001.

三、迴歸模型(Regression model)(羅吉斯迴歸模型, logistic regression model)

探討 CAT 是否會跟 CHD 有相關，即為先前提到的兩樣本比例檢定 (proportional test, Y 的結果為二分法, binary outcome)，亦可利用羅吉斯迴歸的方式進行。

羅吉斯迴歸式:

Y=1: 發生心臟疾病(CHD) ,

Y=0 : 不發生心臟疾病(no CHD)

X: catecholamine (CAT) 之高低

$$\log it\ P(Y=1|x)=\log\frac{P(Y=1|X)}{1-P(Y=1|X)}=\alpha+\beta X,$$

$$P(Y=1|X)=\frac{\exp(\alpha+\beta X)}{1+\exp(\alpha+\beta X)}$$

若為世代研究 (cohort study)，除了 CAT 之外還要調整年齡(age)跟心電圖結果(ECG):

Y=1: 發生心臟疾病(CHD)

Y=0 : 不發生心臟疾病(no CHD)

X: CAT (高/低); Age (≥55 vs <55); ECG (不正常 vs 正常)

$$P(Y=1|\mathbf{X})=\frac{\exp(\alpha+\sum\beta_k x_k)}{1+\exp(\alpha+\sum\beta_k x_k)}$$

$$\text{logit}\,P(Y=1|\mathbf{X})=\alpha+\sum_{k=1}^{K}\beta_k x_k$$

模型 1: Log(P/(1-P))=α+β_1×(CAT=High)

模型 2: Log(P/(1-P))=α+β_1×(CAT=High)+ β_2×(ECG=Abnormal)

模型 3: Log (P/(1-P))=α+β_1×(CAT=High)+ β_2×(Age ≥55)

模型 4: Log(P/(1-P))=α+β_1×(CAT=High)+ β_2×(ECG=Abnormal)+β_3×(Age ≥55)

模型 5: Log(P/(1-P))=α+β_1×(CAT=High)+ β_2×(ECG=Abnormal)+β_3×(Age ≥55)

+γ_1×(CAT=High) ×(Age ≥55)

以分層分析的觀點來看，會分成年齡小於 55 下 ECG 正常及不正常 (age<55 ECG=0, age≥55 ECG=1)，以及年齡大於 55 下 ECG 正常及不正常(age<55 ECG=0, age≥55 ECG=1)，如此一來，隨著迴歸模型等式右側的解釋變項的不同構成了 2×2=4 ，四個類別的結果，故以羅吉斯迴歸模型來看即為模型 4 (Log(P/(1-P))=α+β_1×(CAT=High)+ β_2×(ECG=Abnormal)+β_3×(Age ≥55))。而迴歸模型中對應到的變項之係數β取指數後即為 OR:

模型 1 的β_1取指數即為 CAT 的粗勝算比 (crude odds ratio);

模型 2 的β_1取指數即為調整心電圖結果後的 CAT 勝算比 (odds ratio with the adjustment for ECG);

模型 3 的β_1取指數即為調整年齡後的 CAT 勝算比 (odds ratio with the adjustment for age);

模型 4 的β_1取指數即為同時調整年齡及心電圖結果後的 CAT 勝算比 (odds ratio with the adjustment for age and ECG);

利用羅吉斯迴歸得到對 CHD 風險之迴歸係數(標準差)估計結果列表如下:

變數	**分層**	***粗率*** **(模型 1)**	**調整心電圖** **(模型 2)**	**調整年齡** **(模型 3)**	**調整年齡及心電圖** **(模型 4)**	**年齡及 CAT 之交互作用** **(模型 5)**
CAT	High vs Low	0.8959	0.7601	0.6491	0.5160	0.6374
		(0.2646)	(0.2918)	(0.2912)	(0.3146)	(0.5971)
ECG	Abnormal vs		0.3191		0.3190	0.3156
	Normal		(0.2861)		(0.2858)	(0.2862)
Age	≥55 vs <55			0.5457	0.5451	0.5803
				(0.2824)	(0.2821)	(0.3184)
CAT*Age						-0.1583
						(0.6686)

交互作用: γ_1=-0.1583 (SE=0.6686), Wald χ^2=0.0561, P-value=0.8128

如同在分層分析時,進行調整前須先確定每個分層的 RR 是否相同 (同質)，即確定是否有交互作用，羅吉斯迴歸也可以檢定變項間是否存在交互作用, 以 CAT 及年齡為例, 檢定方法為在模型中加入一個變項為(CAT*年齡)表示 CAT 及年齡的交互作用(模型 5)檢定其對應的迴歸係數是否顯著 (γ_1)。 由以上報表看出 CAT 與年齡的交互作用是不顯著的 (γ_1=-0.1583 (SE=0.6686), Wald χ^2=0.0561, P-value=0.8128)。

利用上表中估計得到的迴歸係數可以計算風險比 (risk ratio, RR)如下(相對風險比(95% 信賴區間):

變數	分層	*粗率* (模型 1)	調整心電圖 (模型 2)	調整年齡 (模型 3)	調整年齡及心電圖 (模型 4)
CAT	High vs Low	2.45 (1.46, 4.12)	2.14 (1.21, 3.79)	1.91 (1.08, 3.39)	1.68 (0.90, 3.10)
ECG	Abnormal vs Normal	---	1.38 (0.79, 2.41)	---	1.38 (0.79, 2.41)
Age	≥55 vs <55	---	---	1.73 (0.99, 3.00)	1.73 (0.99, 3.00)

與分層分析的結果相比較, 利用分層分析中的合併加權方法 (pooled method) 得到的調整年齡及心電圖的風險比為 1.71, 利用 Mantel-Haenszel 方法得到的結果為 1.69, 而用羅吉斯迴歸得到的結果為 1.68, 三者相近。

以年齡標準化調整干擾因子 (Ebook 01. pp.137-140)

在觀察性研究中，控制干擾因子的方法除了利用分層分析以及迴歸模式外，還可以利用標準化的方式達到此一目的。

範例

為評估各區域間心血管疾病 (cardiovascular disease)發生的風險是否有所不同，研究者收集了兩個區域的心血管疾病資料以及人口資料如下：

標準人口		區域 1			區域 2			
年齡層	人口數	事件數	人口數	發生率	事件數	人口數	發生率	發生率之比
40-49	126,300	18	30,000	0.0006	45	90,000	0.0005	1.20
50-59	99,200	108	60,000	0.0018	300	150,000	0.002	0.90
60-69	66,800	1,050	210,000	0.005	360	60,000	0.006	0.83
總和	292,300	1,176	300,000	0.00392	705	300,000	0.00235	1.67

(1) 利用粗發生率之比 (crude risk ratio)比較兩區域心血管疾病的發生風險。

(2) 利用標準人口使用直接年齡調整方法後比較兩區域心血管疾病的發生風險。

1. 以年齡標準化方法調整干擾因子：

資料型式:

年齡組	標準族群			研究族群		
	人口數	事件數	事件比例	人口數	事件數	事件比例
1	N_1	R_1	P_1	n_1	r_1	p_1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
i	N_i	R_i	P_i	n_i	r_i	p_i
⋮	⋮	⋮	⋮	⋮	⋮	⋮
k	N_k	R_k	P_k	n_k	r_k	p_k

由於兩族群的人口年齡不同, 因此在比較時利用依標準人口, 將該標準人口某年齡層的人數 (N_i)乘上該區域對應的年齡層的發生率(p_i), 加總後得到加權平均的調整年齡結構之總和發生率(p'), 利用調整年齡結構後的總合發生率進行不同區域間的比較, 藉以避免區域間不同的年齡結構對粗發生率所造成的影響. 標準人口依進行比較的區域之不同可以是世界衛生組織(WHO)提供的標準人口結構, 或台灣地區的人口結構。

$$p' = \frac{\sum N_i p_i}{\sum N_i} \qquad (1)$$

2. 以直接標準化方法比較兩區域之心血管疾病發生的風險是否有所不同

(1) 利用粗發生率之比 (crude risk ratio)比較兩區域心血管疾病的發生風險。

粗發生率之比 (crude risk ratio): $\frac{1176/300000}{705/300000} = \frac{0.00392}{0.00235} = 1.67$ ，區域 1 之心血管疾病發生率高於區域 2 的心血管疾病發生率。

進一步觀察區域 1 與區域 2 的人口結構可以發現，區域 1 的老年(60-69 歲)人口較多，而區域 2 的中年(50-59 歲)人口較多。考慮到心

血管疾病在老年人口發生的風險可能較高，因此用粗發生率進行兩區域的比較可能不適合。

(2) 利用標準人口使用直接年齡調整方法後比較兩區域心血管疾病的發生風險。

利用標準人口之年齡結構進行直接年齡標準化方法：

區域 1：$$\frac{126300\times0.0006+99200\times0.0018+66800\times0.005}{292300}=0.00201$$

區域 2：$$\frac{126300\times0.0005+99200\times0.002+66800\times0.006}{292300}=0.00227$$

上式中，各年齡層中的人口數適用標準人口的人口數，而相對應的發生率是該特定區域的發生率，利用標準人口的年齡結構進行加權平均後得到總合發生率。利用直接方法調整年齡後的發生率比值為 0.89。因此區域 1 的心血管疾病發生率低於區域 2 的心血管疾病發生率。

直接用粗率進行比較由於受到年齡效果的干擾，因此得到相反的結果，在進行直接年齡標準化時，之所以可以利用標準人口的年齡結構對特定區域的發生率作加權平均得到總合發生率，是由於將年齡視為干擾因子 (confounder)，而不存在有交互作用的關係 (interaction, 即發生率之比不隨著年齡層的不同而有所不同)。因此可以透過加權平均的方式調整年齡層的干擾效果，得到總合發生率進行比較。這與 Cox 等比風險迴歸模型 (Cox PHREG model)中的等比風險 (proportional assumption)假設相同(風險比值不隨時間不同而有所不同)。若事件的發生率比值與年齡具有交互作用存在，就無法將年齡視為干擾因子進行調整。

流行病學方法論及實驗 (The Methods of Epidemiology and Practices): 解決辛普森矛盾之護理魔術

授課教師：陳秀熙 教授

台灣大學流行病學與預防醫學研究所

Simpson's Paradox in Nursing Care -Can Nightingale's Sanitary Improvement Reduce Death Rate ?

		Dead: Yes	Dead: No	
Intervention	No(B)	45	955	1000
	Yes(A)	15	985	1000
		60	1940	2000

$RR = \frac{P_A}{P_B} = 0.33$ Relative Risk 危險對比值

Non-slum Area　　Slum Area

		Non-slum Area: Dead Yes	Non-slum Area: Dead No		Slum Area: Dead Yes	Slum Area: Dead No	
Inter-vention	No (B)	3	297	300	42	658	700
	Yes (A)	9	891	900	6	94	100
		12	1188	1200	48	752	800

$RR = \frac{P_A}{P_B} = 1$　　$RR = \frac{P_A}{P_B} = 1$

Magic Method for Removing Confounding Factor (Socio-economic Status):Randomization

Non-slum Area　　Non-slum Area

		Non-slum Area: Dead Yes	Non-slum Area: Dead No	Non-slum Area: Dead Yes	Non-slum Area: Dead No
Inter-vention	No (B)				
	Yes (A)				

$RR = \frac{P_A}{P_B} =$ _____　　$RR = \frac{P_A}{P_B} =$ _____

觀察型研究控制干擾因子的方法

(1) 配對(matching)

Ex. 為研究新型抗生素supramycin和傳統使用之抗生素amoxicillin相比是否較容易導致患者長疹子:

Supramycin 組(暴露組)	Amoxicillin組(非暴露組)
個案1: 使用Supramycin, 無濕疹病史	使用Amoxicillin, 無濕疹病史
個案2: 使用Supramycin, 有濕疹病史	使用Amoxicillin, 有濕疹病史
個案3: 使用Supramycin, 無濕疹病史	使用Amoxicillin, 無濕疹病史
個案4: 使用Supramycin, 無濕疹病史	使用Amoxicillin, 無濕疹病史
.....	.....

暴露組: 非暴露組= 1:1
1:2
...
1:n

配對因子(干擾因子): 濕疹病史
(用藥過敏史、家族濕疹病史)

觀察型研究控制干擾因子的方法

(1) 配對(matching)

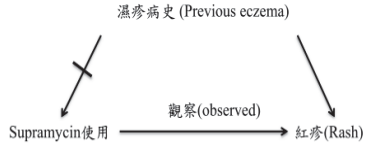
缺點:

1. 研究設計階段即決定配對因子→後續新增配對因子
2. 多個配對因子→不易配對
3. 配對因子→風險因子

觀察型研究控制干擾因子的方法

(2) 限制(restriction)

Ex.為研究新型抗生素supramycin和傳統使用之抗生素amoxicillin相比是否較容易導致患者長疹子→去除有濕疹病史之個案



限制與缺點：統計檢定力↓，缺乏外推性

觀察型研究控制干擾因子的方法

(2) 限制(restriction)

缺點:

1. 統計檢定力下降
2. 缺乏外推性

資料分析

1. 粗率(Crude rate)分析

(1)相差風險性(Risk difference, RD)

	E	$\bar{E}$	
D	a	b	m_1
$\bar{D}$	c	d	m_2
	n_1	n_2	n

$$\hat{RD} = \hat{P}_1 - \hat{P}_0 = 0.131 \quad \text{(點估計值)}$$

$$\hat{RD} \pm Z_{\frac{\alpha}{2}} \hat{SE}(\hat{RD}) = \hat{RD} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_0(1-\hat{P}_0)}{n_2}}$$

(大樣本理論-泰勒分析)

$$\hat{RD} \pm Z_{\frac{\alpha}{2}} \hat{SE}(\hat{RD}) = \hat{RD} \pm Z_{\frac{\alpha}{2}} \sqrt{(\hat{RD})^2 / \chi^2_{MH}}$$

(Test-based method)

其中 $\chi^2_{MH} = \frac{(n-1)(ad-bc)^2}{n_1 n_2 m_1 m_2}$

資料分析

1. 粗率(Crude rate)分析

(1)相差風險性(Risk difference, RD)

範例:在兒茶酚胺激素(catecholamine, CAT)與冠狀動脈血管疾病(coronary heart disease, CHD)之相關性研究中，資料結構如下，

	High CAT	*Low CAT*	
CHD	27	44	71
No CHD	95	443	538
Total	122	487	609

比較高濃度(P_1)與低濃度(P_0)的茶酚胺激素個體發生冠狀動脈血管疾病的風險：

$\hat{P}_1 = 0.221 \qquad \hat{P}_0 = 0.090 \qquad \hat{RD} = \hat{P}_1 - \hat{P}_0 = 0.131$

$\hat{SE}(\hat{RD}) = \sqrt{\frac{0.22(1-0.22)}{122} + \frac{0.090(1-0.090)}{487}} = 0.040$ 95%CI=(0.053-0.209) 利用泰勒分析

$\chi^2_{MH} = \frac{(609-1)(27\times 443 - 44\times 95)^2}{122\times 487\times 71\times 538} = 16.22$ 95%CI=$0.131 \pm 1.96\sqrt{(0.131)^2/16.22} = (0.067, 0.195)$ Test-based method

資料分析

1. 粗率(Crude rate)分析

(2)風險比值或相對風險(Risk Ratio or Relative Risk)

	E	$\bar{E}$	
D	a	b	m_1
$\bar{D}$	c	d	m_2
	n_1	n_2	n

$$\hat{RR} = \hat{P}_1 / \hat{P}_0 \quad \text{(點估計值)}$$

$$\hat{SE}\{\log(\hat{RR})\} \approx \sqrt{\frac{1-\hat{P}_1}{n_1\hat{P}_1} + \frac{1-\hat{P}_0}{n_2\hat{P}_0}} \quad \text{(大樣本理論-泰勒分析)}$$

$$\log(\hat{RR}) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{1-\hat{P}_1}{n_1\hat{P}_1} + \frac{1-\hat{P}_0}{n_2\hat{P}_0}}$$

資料分析

2. 分層分析(stratified analysis)分析

(1)2x2聯列表之干擾因子調整

干擾因子F有k個層別
Q1. 調整F因子後的RR=?
Q2. 各層中的RR是否夠接近?

F_1

	E	$\bar{E}$	
D	a_1	b_1	m_{11}
$\bar{D}$	c_1	d_1	m_{21}
	n_{11}	n_{21}	n_1

F_2

	E	$\bar{E}$	
D	a_2	b_2	m_{12}
$\bar{D}$	c_2	d_2	m_{22}
	n_{12}	n_{22}	n_2

.....

F_i

	E	$\bar{E}$	
D	a_i	b_i	m_{1i}
$\bar{D}$	c_i	d_i	m_{2i}
	n_{1i}	n_{2i}	n_i

F_k

	E	$\bar{E}$	
D	a_k	b_k	m_{1k}
$\bar{D}$	c_k	d_k	m_{2k}
	n_{1k}	n_{2k}	n_k

資料分析

2. 分層分析(stratified analysis)分析

(2)調整之風險比(Adjusted Risk Ratio)

(A)合併加權估計式(Pooled estimate)

F_1

	E	$\bar{E}$	
D	a_1	b_1	m_{11}
$\bar{D}$	c_1	d_1	m_{21}
	n_{11}	n_{21}	n_1

F_2

	E	$\bar{E}$	
D	a_2	b_2	m_{12}
$\bar{D}$	c_2	d_2	m_{22}
	n_{12}	n_{22}	n_2

..... F_i

	E	$\bar{E}$	
D	a_i	b_i	m_{1i}
$\bar{D}$	c_i	d_i	m_{2i}
	n_{1i}	n_{2i}	n_i

F_k

	E	$\bar{E}$	
D	a_k	b_k	m_{1k}
$\bar{D}$	c_k	d_k	m_{2k}
	n_{1k}	n_{2k}	n_k

F的第i個分層： $\hat{RR}_i = \frac{a_i/n_{1i}}{b_i/n_{2i}}$ $\quad \hat{Var}\{\log \hat{RR}_i\} = \frac{c_i}{a_i n_{1i}} + \frac{d_i}{b_i n_{2i}}$

$$\log \hat{RR}_w = \frac{\sum_{i=1}^{k} W_i \log(RR_i)}{\sum_{i=1}^{k} W_i} \text{(□□□□□ } \log RR\text{)}, \qquad W_i = \frac{a_i b_i n_{1i} n_{2i}}{a_i d_i n_{1i} + b_i c_i n_{2i}} \text{ (□log } RR \text{ □□□□□□□□□□)}$$

$$\hat{RR}_w = \exp(\log \hat{RR}_w)$$

資料分析

2. 分層分析(stratified analysis)分析

(2)調整之風險比(Adjusted Risk Ratio)

(B)Mantel-Haenszel估計式 (Mantel-Haenszel estimator)

F_1

	E	$\bar{E}$	
D	a_1	b_1	m_{11}
$\bar{D}$	c_1	d_1	m_{21}
	n_{11}	n_{21}	n_1

F_2

	E	$\bar{E}$	
D	a_2	b_2	m_{12}
$\bar{D}$	c_2	d_2	m_{22}
	n_{12}	n_{22}	n_2

..... F_i

	E	$\bar{E}$	
D	a_i	b_i	m_{1i}
$\bar{D}$	c_i	d_i	m_{2i}
	n_{1i}	n_{2i}	n_i

F_k

	E	$\bar{E}$	
D	a_k	b_k	m_{1k}
$\bar{D}$	c_k	d_k	m_{2k}
	n_{1k}	n_{2k}	n_k

□□□□□□$\left(\frac{n_{1i}}{n_i}\right)b_i$ □□: $\hat{RR}_{MH} = \frac{\sum_{i=1}^{k}\left(\frac{n_{1i}}{n_i}\right) b_i \left(\frac{a_i n_{2i}}{b_i n_{1i}}\right)}{\sum_{i=1}^{k}\left(\frac{n_{1i}}{n_i}\right) b_i} = \frac{\sum_{i=1}^{k}\left(\frac{n_{2i}}{n_i}\right) a_i}{\sum_{i=1}^{k}\left(\frac{n_{1i}}{n_i}\right) b_i}$

資料分析

2. 分層分析(stratified analysis)分析

(2)調整之風險比(Adjusted Risk Ratio)

(D)範例

在研究catecholamine (CAT) 與心臟疾病(CHD)之關係的研究中，年齡的大小(Age)及心電圖是否異常(ECG)為兩個可能的干擾因子

特定分層之RR

年輕組($Age < 55$)，心電圖無異常($ECG = 0$)

	High CAT	Low CAT	Total
CHD	1	17	18
No CHD	7	257	264
Total	8	274	282

$\hat{RR}_1 = 2.015(0.30 \sim 13.34)$

年輕組($Age < 55$)，心電圖有異常($ECG = 1$)

	High CAT	Low CAT	Total
CHD	3	7	10
No CHD	14	52	66
Total	17	59	76

$\hat{RR}_2 = 1.49(0.43 \sim 5.14)$

年老組($Age \geq 55$)，心電圖無異常(ECG) = 0

	High CAT	Low CAT	Total
CHD	9	15	24
No CHD	30	107	137
Total	39	122	161

$\hat{RR}_3 = 1.88(0.89 \sim 3.95)$

年老組($Age \geq 55$)，心電圖有異常($ECG = 1$)

	High CAT	Low CAT	Total
CHD	14	5	19
No CHD	44	27	71
Total	58	32	90

$\hat{RR}_4 = 1.54(0.61 \sim 3.90)$

Crude RR=0.221/0.09=2.46

年輕組($Age < 55$)，心電圖無異常($ECG = 0$)

	High CAT	Low CAT	Total
CHD	1	17	18
No CHD	7	257	264
Total	8	274	282

$\hat{RR}_1 = 2.015(0.30 \sim 13.34)$

年輕組($Age < 55$)，心電圖有異常($ECG = 1$)

	High CAT	Low CAT	Total
CHD	3	7	10
No CHD	14	52	66
Total	17	59	76

$\hat{RR}_2 = 1.49(0.43 \sim 5.14)$

年老組($Age \geq 55$)，心電圖無異常(ECG) = 0

	High CAT	Low CAT	Total
CHD	9	15	24
No CHD	30	107	137
Total	39	122	161

$\hat{RR}_3 = 1.88(0.89 \sim 3.95)$

年老組($Age \geq 55$)，心電圖有異常($ECG = 1$)

	High CAT	Low CAT	Total
CHD	14	5	19
No CHD	44	27	71
Total	58	32	90

$\hat{RR}_4 = 1.54(0.61 \sim 3.90)$

(調整前RR為2.46)

(a)合併加權方法 (pooled method) (logit method)

$$W_1 = \frac{a_1 b_1 n_{11} n_{21}}{b_1 c_1 n_{21} + a_1 d_1 n_{11}} = \frac{(1)(17)(8)(274)}{(17)(7)(274) + (1)(257)(8)} = 1.075$$

$W_2 = 2.497, W_3 = 6.947, W_4 = 4.486$

i	W_i	$\hat{RR}_i$	$\log \hat{RR}_i$	$W_i \log(\hat{RR}_i)$
1	1.075	2.015	0.701	0.753
2	2.497	1.487	0.397	0.991
3	6.947	1.877	0.630	4.374
4	4.486	1.545	0.435	1.952

$\log \hat{RR}_w = \frac{8.070}{15.005} = 0.538$ $\quad \hat{RR}_w = 1.713$ (調整後RR值↓，因部分相關性被年齡及ECG解釋掉了)

$\hat{Var}\{\log(\hat{RR}_w)\} = \frac{1}{15.005} = 0.067$

$95\%CI = \exp(0.538 \pm 1.96 \times \sqrt{0.067}) = (1.032 \sim 2.840)$

年輕組($Age < 55$)，心電圖無異常($ECG = 0$)

	High CAT	Low CAT	Total
CHD	1	17	18
No CHD	7	257	264
Total	8	274	282

$\hat{RR}_1 = 2.015(0.30 \sim 13.34)$

年輕組($Age < 55$)，心電圖有異常($ECG = 1$)

	High CAT	Low CAT	Total
CHD	3	7	10
No CHD	14	52	66
Total	17	59	76

$\hat{RR}_2 = 1.49(0.43 \sim 5.14)$

年老組($Age \geq 55$)，心電圖無異常(ECG) = 0

	High CAT	Low CAT	Total
CHD	9	15	24
No CHD	30	107	137
Total	39	122	161

$\hat{RR}_3 = 1.88(0.89 \sim 3.95)$

年老組($Age \geq 55$)，心電圖有異常($ECG = 1$)

	High CAT	Low CAT	Total
CHD	14	5	19
No CHD	44	27	71
Total	58	32	90

$\hat{RR}_4 = 1.54(0.61 \sim 3.90)$

(調整前RR為2.46)

(b)Mantel-Haenzel方法

i	a_i	b_i	m_{1i}	n_{1i}	n_{2i}	n_i	$\left(\frac{n_{2i}}{n_i}\right)a_i$	$\left(\frac{n_{1i}}{n_i}\right)b_i$
1	1	17	18	8	274	282	0.972	0.482
2	3	7	10	17	59	76	2.329	1.566
3	9	15	24	39	122	161	6.820	3.634
4	14	5	19	58	32	90	4.978	3.222
							15.099	8.904

$$\therefore \hat{RR}_{MH} = \frac{15.099}{8.904} = 1.696$$

$$Var\left\{\log\left(\hat{RR}_{MH}\right)\right\} = \left[\frac{(18)(8)(274)-(1)(17)(282)}{(282)^2} + \frac{(10)(17)(59)-(3)(7)(76)}{(76)^2} + \frac{(24)(39)(122)-(9)(15)(161)}{(161)^2} + \frac{(19)(58)(32)-(14)(5)(90)}{(90)^2}\right]$$

$$\div (15.099)(8.904) = \frac{9.039}{134.442} = 0.067$$

$95\%CI = 1.696 \times, \div \exp\left(1.96 \times \sqrt{0.067}\right) = (1.02, 2.82)$

統計分析之概念

	結果變項(Y)性質	樣本統計值平均值	非迴歸分析方法	迴歸分析方法
二組 (X=1/0)	(1)連續變項	$\bar{Y}$(平均值)	獨立 t-test $(\bar{Y}_1 - \bar{Y}_0)$	線性迴歸 Y=a+bx
	(2)二元變項 (Y=1/0)	$\hat{p}$(比例)=d/n d: 事件 n: 總數	χ^2 test(2×2 列聯表)及比例檢定 $\hat{p}_1 - \hat{p}_0$	羅吉斯迴歸 Logit(P(Y=1|X))=a+bx
	(3)存活時間 (Y=time to event)	存活函數 S(t) 風險函數 h(t)	Hazard Rate $=h_1(t)/h_0(t)$	Cox 迴歸模式(Cox proportional hazard regression model) $h(t)=h_0(t)\exp(a+bx)$
多組 (k=3) X=1/2/3	(1)連續變項	$\bar{Y}_1 - \bar{Y}_2$)/ $\bar{Y}_2 - \bar{Y}_3$)/ $\bar{Y}_1 - \bar{Y}_3$)	One-way Analysis of Variance(ANOVA)	$Y=a+b_1X_1+b_2X_2$ X=1 X=2 X=3; X_1 0 1 0; X_2 0 0 1

迴歸模式之想法

$Y=a+bx+e \quad e \sim N(0,\sigma^2)$

$X=1 \quad \bar{y}_1 = a+b$

$X=0 \quad \bar{y}_0 = a$

$\bar{y}_1 - \bar{y}_0 = b$ (迴歸係數也就是斜率)，b 越大，則兩組平均值差異越大

一、獨立樣本二元變項(Binary varivble)

(1) 兩組樣本比例檢定 (Two-sample proportion test) (參考ebook_01, Ch 6, p.476-484)

$$H_0: P_1 = P_0 \qquad H_1: P_1 \neq P_0 \quad \text{(雙尾檢定)}$$

$$Z = \frac{\left(\hat{P}_1 - \hat{P}_0\right) - 0}{\sqrt{Var\left(\hat{P}_1 - \hat{P}_0\right)}} \qquad Var\left(\hat{P}_1 - \hat{P}_0\right) = \sqrt{\frac{p'q'}{n_1} + \frac{p'q'}{n_2}}$$

當 $Z > Z_{1-\alpha/2}$ 或 $Z < Z_{\alpha/2}$，推翻虛無假說

$Z_{\alpha/2} \le Z \le Z_{1-\alpha/2}$，無法推翻虛無假說

$$p' = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_0}{n_1 + n_2} \quad (p_1 \text{與} p_0 \text{加權})$$

一、獨立樣本二元變項(Binary variable)

(2) 卡方檢定(Chi-square (χ^2) test) (參考ebook_01, Ch 6, p.484-494; ebook 02, Ch 9, p.209-219)

舉例:某臨床研究欲證明某種新藥是否比舊藥，其造成感染傷口癒合狀況較好。研究者利用隨機分派試驗得到下列結果：

	有療效	沒有療效	總和
新藥	40	20	60
舊藥	16	48	64
總和	56	68	124

E: Expectation under H_0 is true

H_0 : 是否有較好的治療效果和使用新藥與舊藥無關

a. 大樣本理論

$$E_{11} = 60 \times \frac{56}{124}$$

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$= \frac{(40-(56\times60)/124)^2}{56\times60/124} + \frac{(20-(60\times68)/124)^2}{60\times68/124} + \frac{(16-(56\times64)/124)^2}{56\times64/124} + \frac{(48-(64\times68)/124)^2}{64\times68/124}$$

$$= 21.71 \quad (p < 0.0001)$$

$$\sim \chi^2_{(1)} \qquad \chi^2_{(1)} \equiv Z^2 = \left(\frac{O-E}{\sqrt{V}}\right)^2$$

Table A–5. Percentage points or critical values for the χ^2 distribution corresponding to commonly used areas under the curve.

Degrees of Freedom	Area in Upper Tail 0.10	0.05	0.01	0.001
1	2.706	3.841	6.635	10.828
2	4.605	5.991	9.210	13.816
3	6.251	7.815	11.345	16.266
4	7.779	9.488	13.277	18.467
5	9.236	11.071	15.086	20.515
6	10.645	12.592	16.812	22.458
7	12.017	14.067	18.475	24.322
8	13.362	15.507	20.090	26.125
9	14.684	16.919	21.666	27.877
10	15.987	18.307	23.209	29.588
11	17.275	19.675	24.725	31.264
12	18.549	21.026	26.217	32.909
13	19.812	22.362	27.688	34.528
14	21.064	23.685	29.141	36.123
15	22.307	24.996	30.578	37.697
16	23.542	26.296	32.000	39.252
17	24.769	27.587	33.409	40.790
18	25.989	28.869	34.805	42.312
19	27.204	30.144	36.191	43.820
20	28.412	31.410	37.566	45.315

(2)卡方檢定(Chi-square (χ^2) test) (參考ebook_01, Ch 6, p.484-494; ebook 02, Ch 9, p.209-219)

	有療效	沒有療效	總和
新藥	10 (a)	2 (b)	12
舊藥	2 (c)	4 (d)	6
總和	12	6	18

b. 小樣本理論

Fisher's exact test

此表格之精算機率為:

$$P(a,b,c,d) = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}$$

(超幾何分布, hypergeometric distribution)

使用卡方檢定: $\chi^2_{(1)} = 4.50$, $p = 0.034$ (推翻虛無假說)

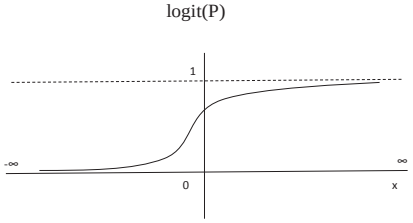
使用Fisher's Exact Test:

單尾檢定(One-Tail): p=0.057

雙尾檢定 (Two-Tail): p=0.107 (無法推翻虛無假說)

(3)羅吉斯迴歸(Logistic Regression Model) (參考ebook_01, Ch 10, p.852-858; ebook 02, Ch 12, p.296-299; ebook 03, Ch 19, P.21-214)

$$\log\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \text{logit}(P(Y=1)) = \text{logit}(P) = \beta_0 + \beta_1 X + \varepsilon$$



(3)羅吉斯迴歸(Logistic Regression Model) (參考ebook_01, Ch 10, p.852-858; ebook 02, Ch 12, p.296-299; ebook 03, Ch 19, P.21-214)

$$\log\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \text{logit}(P(Y=1)) = \text{logit}(P) = \beta_0 + \beta_1 X + \varepsilon$$

若為兩組比較, 則X代表了組別:

$$\begin{cases} X=1 & log\frac{\hat{P}_1}{1-\hat{P}_1} = \hat{\beta}_0 + \hat{\beta}_1 \quad (1) \\ X=0 & log\frac{\hat{P}_0}{1-\hat{P}_0} = \hat{\beta}_0 \quad (2) \end{cases}$$

(1)-(2) $\log(\hat{OR}) = \log\left(\frac{\hat{P}_1/1-\hat{P}_1}{\hat{P}_0/1-\hat{P}_0}\right) = \hat{\beta}_1$

勝算比 $\hat{OR} = \exp(\hat{\beta}_1)$

(3)羅吉斯迴歸(Logistic Regression Model) (參考 ebook_01, Ch 10, p.852-858; ebook 02, Ch 12, p.296-299; ebook 03, Ch 19, P.21-214)

比較兩組有無差異 ($H_0: P_1 = P_0$) 等同於檢定 $H_0: \text{OR} = 1$ 或 $H_0: \beta_1 = 0$

例: 上述新治療方法對於healing及infection之影響

	有療效	沒有療效	總和
新藥	40	20	60
舊藥	16	48	64
總和	56	68	124

$$H_0: \beta_1 = 0 \quad \text{vs.} \quad H_1: \beta_1 \neq 0$$

羅吉斯迴歸 ➔ $\hat{\beta}_1 = 1.7918$, $\hat{\text{OR}} = \exp(1.7918) = 6$

檢定統計值 $\chi^2 = 20.3 > 3.84$, $P < 0.0001$

三、迴歸模型(Regression model) (羅吉斯迴歸模型, logistic regression model)

探討CAT是否會和CHD有相關

羅吉斯迴歸式:

Y=1: 發生心臟疾病(CHD),

Y=0 : 不發生心臟疾病(no CHD)

X: catecholamine (CAT) 之高低

若為世代研究:
X: Age (≥55 vs <55);
ECG (不正常 vs 正常)

$$P(Y=1|\mathbf{X}) = \frac{\exp(\alpha + \sum \beta_k x_k)}{1 + \exp(\alpha + \sum \beta_k x_k)}$$

$$\text{logit } P(Y=1|\mathbf{X}) = \alpha + \sum_{k=1}^{K} \beta_k x_k$$

羅吉斯迴歸模型 (logistic regression model)

模型1: Log(P/(1-P))=α+β_1x(CAT=High)
模型2: Log(P/(1-P))=α+β_1x(CAT=High)+ β_2x(ECG=Abnormal)

利用羅吉斯迴歸得到對CHD風險之迴歸係數(標準差)估計結果列表如下

變數	分層	粗率 (模型 1)
CAT	High vs Low	0.8959 (0.2646)

羅吉斯迴歸模型 (logistic regression model)

模型1: Log(P/(1-P))=α+β_1x(CAT=High)
模型2: Log(P/(1-P))=α+β_1x(CAT=High)+ β_2x(ECG=Abnormal)
模型3: Log (P/(1-P))=α+β_1x(CAT=High)+ β_2x(Age ≥55)

利用羅吉斯迴歸得到對CHD風險之迴歸係數(標準差)估計結果列表如下

變數	分層	粗率 (模型 1)	調整 心電圖 (模型 2)
CAT	High vs Low	0.8959 (0.2646)	0.7601 (0.2918)
ECG	Abnormal vs Normal		0.3191 (0.2861)

羅吉斯迴歸模型 (logistic regression model)

模型1: Log(P/(1-P))=α+β_1x(CAT=High)
模型2: Log(P/(1-P))=α+β_1x(CAT=High)+ β_2x(ECG=Abnormal)
模型3: Log (P/(1-P))=α+β_1x(CAT=High)+ β_2x(Age ≥55)
模型4: Log(P/(1-P))=α+β_1x(CAT=High)+ β_2x(ECG=Abnormal)+β_3x(Age ≥55)

利用羅吉斯迴歸得到對CHD風險之迴歸係數(標準差)估計結果列表如下

變數	分層	粗率 (模型 1)	調整 心電圖 (模型 2)	調整年齡 (模型 3)
CAT	High vs Low	0.8959 (0.2646)	0.7601 (0.2918)	0.6491 (0.2912)
ECG	Abnormal vs Normal		0.3191 (0.2861)	
Age	≥55 vs <55			0.5457 (0.2824)

羅吉斯迴歸模型 (logistic regression model)

模型1: Log(P/(1-P))=α+β_1x(CAT=High)
模型2: Log(P/(1-P))=α+β_1x(CAT=High)+ β_2x(ECG=Abnormal)
模型3: Log (P/(1-P))=α+β_1x(CAT=High)+ β_2x(Age ≥55)
模型4: Log(P/(1-P))=α+β_1x(CAT=High)+ β_2x(ECG=Abnormal)+β_3x(Age ≥55)
模型5: Log(P/(1-P))=α+β_1x(CAT=High)+ β_2x(ECG=Abnormal)+β_3x(Age ≥55)
+γ_1x(CAT=High) x(Age ≥55)

利用羅吉斯迴歸得到對CHD風險之迴歸係數(標準差)估計結果列表如下

變數	分層	粗率 (模型 1)	調整 心電圖 (模型 2)	調整年齡 (模型 3)	調整年齡及 心電圖 (模型 4)
CAT	High vs Low	0.8959 (0.2646)	0.7601 (0.2918)	0.6491 (0.2912)	0.5160 (0.3146)
ECG	Abnormal vs Normal		0.3191 (0.2861)		0.3190 (0.2858)
Age	≥55 vs <55			0.5457 (0.2824)	0.5451 (0.2821)

羅吉斯迴歸模型 (logistic regression model)

利用上表中估計得到的迴歸係數可以計算風險比 (risk ratio, RR)如下(相對風險比(95% 信賴區間):

變數	分層	粗率 (模型 1)	調整心電圖 (模型 2)	調整年齡 (模型 3)	調整年齡及心電圖 (模型 4)
CAT	High vs Low	2.45 (1.46, 4.12)	2.14 (1.21, 3.79)	1.91 (1.08, 3.39)	1.68 (0.90, 3.10)
ECG	Abnormal vs Normal	---	1.38 (0.79, 2.41)	---	1.38 (0.79, 2.41)
Age	≥55 vs <55	---	---	1.73 (0.99, 3.00)	1.73 (0.99, 3.00)

調整後RR (aRR)= 1.71 (分層分析中的合併加權方法)
1.69 (Mantel-Haenszel方法)
1.68 (羅吉斯迴歸方法)

31

以年齡標準化調整干擾因子

兩區域之心血管疾病(cardiovascular disease)發生風險是否不同?

標準人口	區域 1			區域2			
年齡層	人口數	事件數	人口數	發生率	事件數	人口數	發生率	發生率之比
40-49	126,300	18	30,000	0.0006	45	90,000	0.0005	1.20
50-59	99,200	108	60,000	0.0018	300	150,000	0.002	0.90
60-69	66,800	1,050	210,000	0.005	360	60,000	0.006	0.83
總和	292,300	1,176	300,000	0.00392	705	300,000	0.00235	1.67

粗率 : 區域1 > 區域2
年齡別發生率: 區域1 ≈ 區域2

32

以年齡標準化方法調整干擾因子

年齡組	標準族群	研究族群		
	人口數	人口數	事件數	事件比例
1	N_1	n_1	r_1	p_1
⋮	⋮	⋮	⋮	⋮
i	N_i	n_i	r_i	p_i
⋮	⋮	⋮	⋮	⋮
k	N_k	n_k	r_k	p_k

$$p' = \frac{\sum N_i p_i}{\sum N_i}$$

33

粗發生率之比

標準人口	區域 1			區域2			
年齡層	人口數	事件數	人口數	發生率	事件數	人口數	發生率	發生率之比
40-49	126,300	18	30,000	0.0006	45	90,000	0.0005	1.20
50-59	99,200	108	60,000	0.0018	300	150,000	0.002	0.90
60-69	66,800	1,050	210,000	0.005	360	60,000	0.006	0.83
總和	292,300	1,176	300,000	0.00392	705	300,000	0.00235	1.67

粗發生率之比$(crude\ risk\ ratio)$:

$$\frac{1176/300000}{705/300000} = \frac{0.00392}{0.00235} = 1.67$$

34

直接標準化方法比較兩區域之心血管疾病發生的風險

標準人口	區域 1			區域2			
年齡層	人口數	事件數	人口數	發生率	事件數	人口數	發生率	發生率之比
40-49	126,300	18	30,000	0.0006	45	90,000	0.0005	1.20
50-59	99,200	108	60,000	0.0018	300	150,000	0.002	0.90
60-69	66,800	1,050	210,000	0.005	360	60,000	0.006	0.83
總和	292,300	1,176	300,000	0.00392	705	300,000	0.00235	1.67

$$\text{區域1}: \frac{126300\times0.0006 + 99200\times0.0018 + 66800\times0.005}{292300} = 0.00201$$

$$\text{區域2}: \frac{126300\times0.0005 + 99200\times0.002 + 66800\times0.006}{292300} = 0.00227$$

$$\text{年齡標準化發生率之比}: \frac{0.00201}{0.00227} = 0.89$$

35

四、 觸媒流行病學故事與護理照護

I 效應修飾因子之概念

1. 定義

若 X 與 Y 之間的關係會因 Z 不同而有異，則 Z 即為一可能的效應修飾因子，這代表 X 與 Z 之間的交互作用與 Y 的結果具相關性。

2. 資料與結果

(1) Modified effect due to Z

	Z(+)		Z(-)		Aggregated	
	X(+)	X(-)	X(+)	X(-)	X(+)	X(-)
Y(+)	80	20	50	30	130	50
Y(-)	60	20	40	90	100	110
	OR=1.33		OR=3.75		OR=2.86(Crude)	

(2) Opposite effects due to Z

	Z(+)		Z(-)		Aggregated	
	X(+)	X(-)	X(+)	X(-)	X(+)	X(-)
Y(+)	180	60	40	140	220	200
Y(-)	120	90	80	110	200	200
	OR=2.25		OR=0.39		OR=1.10(Crude)	

3. 結果解釋

(1) X 對 Y 的效應會因 Z 是否存在而異。當 Z 發生(Z(+) 時，X 對 Y 的勝算比大於 Z 不存在時(Z(-)) X 對 Y 的勝算比。

(2) X對Y的效應在Z是否發生的情況之下呈現相反的結果。

(3) 當修飾現象存在時，不應該報導合併分析結果。

(4) 多因子隨機臨床試驗(Factorial randomized controlled trial)設計是用來驗證 X 與 Z 之間是否存在交互作用。

II 評估該因子是否具有效應修飾作用

1. 分層分析以及同質性檢定

(1) 世代追蹤研究之例子

例：在一研究 CAT 與 CHA 之相關研究，研究者考慮 Age 與 ECG 可能是 confounders，研究者欲利用分層分析方法評估是否 Age 與 ECG 會 Confound CAT 與 CHD 之關係。

表一是不分層之 2×2 表

	High CAT	*Low* CAT
CHD	27	44	71
No CHD	95	443	538
Total	122	487	609

其 $\hat{RR}$ =2.456　95%CI(0.46-1.336)

如果將資料分層後可以得到下表

$Age < 55, ECG = 0$

	High CAT	*Low CAT*	*Total*
CHD	1	17	18
No CHD	7	257	264
Total	8	274	282

$\hat{RR}_1 = 2.015(0.30 \sim 13.34)$

$Age < 55, ECG = 1$

	High CAT	*Low CAT*	*Total*
CHD	3	7	10
No CHD	14	52	66
Total	17	59	76

$\hat{RR}_2 = 149(0.43 \sim 5.14)$

$Age \geq 55, ECG = 0$

	High CAT	*Low CAT*	*Total*
CHD	9	15	24
No CHD	30	107	137
Total	39	122	161

$\hat{RR}_3 = 188(0.89 \sim 3.95)$

$Age \geq 55, ECG = 1$

	High CAT	*Low CAT*	*Total*
CHD	14	5	19
No CHD	44	27	71
Total	58	32	90

$\hat{RR}_4 = 154(0.61 \sim 3.90)$

為了測定是否有 effect modification(也就是 E 和 D 之作用是否會因為干擾因子之不同而有異)通常在分層分析求 summary 之$\hat{RR}$之前會使用 X^2 test 檢視其各分層是否有 homogeneity，其使用之統計值為

$$\sum_{i=1}^{k} \frac{(\log \hat{RR}_i - \log RR_W)^2}{Var(\log RR_i)} \qquad \text{-------} \qquad (1)$$

在樣本數很大之情況下，式(1)是一個 X^2_{k-1} 分佈

以上述 CAT 及 CHD 為例，式(1)之統計值為

$$\frac{(0.701-0.538)^2}{0.9302}+\frac{(0.397-0.538)^2}{0.4005}+\frac{(0.630-0.538)^2}{0.1439}$$
$$+\frac{(0.435-0.538)^2}{0.2229}=0.185$$

$$P(X_3^2>0.185)=0.980$$

因此我們下結論：CAT 與 CHD 之 risk ratio 在 AGE 及 ECG 各分層並沒有顯著差異，也就是 CAT 與 CHD 之作用並不會因 AGE 及 ECG 之不同而有差異。

(2) 病例-對照研究之例子

以病例-對照試驗設計研究收縮壓與心肌梗塞之相關性 (Case-control study on SBP and MI)

1. 計算調整年齡後，收縮壓對心肌梗塞的勝算比，並與粗勝算比 (Crude odds ratio)以及年齡分層之勝算比相比較。

2. 檢驗年齡分層之勝算比是否同質 (Homogenous).

若有第 i 個分層的聯列表資料如下:

	E	$\overline{E}$	
D	a_i	b_i	m_{1i}
$\overline{D}$	c_i	d_i	m_{2i}
	n_{1i}	n_{2i}	n_i

(i) 合併加權估計式 (Pooled estimator)

$$\hat{OR}_i=\frac{a_i d_i}{b_i c_i}$$

$$Var\left\{\log(\hat{OR}_i)\right\}=\frac{1}{a_i}+\frac{1}{b_i}+\frac{1}{c_i}+\frac{1}{d_i}$$

$$\log\hat{OR}_W=\frac{\sum_{i=1}^{k}W_i\log\hat{OR}_i}{\sum_{i=1}^{k}W_i}$$

$$W_i = \frac{1}{Var\{\log(\hat{OR}_i)\}}$$

(ii) Mantel-Haenszel 方法

$$\hat{OR}_{MH} = \frac{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)\left(\frac{a_i d_i}{b_i c_i}\right)}{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)} = \frac{\sum_{i=1}^{k}\left(\frac{a_i d_i}{n_i}\right)}{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)}$$

$$Var_{MH}\left\{\log(\hat{OR}_{MH})\right\} = \frac{\sum_{i=1}^{k}\left(\frac{a_i d_i}{n_i}\right)^2\left[\frac{1}{a_i}+\frac{1}{b_i}+\frac{1}{c_i}+\frac{1}{d_i}\right]}{\left(\sum_{i=1}^{k}\frac{a_i d_i}{n_i}\right)^2}$$

3. 評估因子是否造成效應修飾 (Effect modification)

$\sum_{i=1}^{k}\frac{\left\{\log\hat{OR}_i - \log\hat{OR}\right\}^2}{Var\left\{\log\hat{OR}\right\}}$ 此統計檢定值服從卡方分佈(χ^2 distribution)

為了解收縮壓(Systolic blood pressure, SBP)與心肌梗塞(Myocardial infarction, MI)之相關性，研究者利用病例-對照試驗設計(Case-control study) 進行研究。已知年齡可能為干擾因子，因此研究以年齡(≥60 歲或 <60 歲) 作為分層。收集的資料如下：

	年齡≥60 歲 收縮壓 (SBP)			年齡<60 歲 收縮壓 (SBP)		
心肌梗塞	≥140	<140		≥140	<140	
發生(Case)	9	6	**15**	20	21	**41**
未發生 (Control)	115	73	**188**	596	1,171	**1,767**
總人數	**124**	**79**	**203**	**616**	**1,192**	**1,808**

(1) 粗率 (Crude odds ratio) : 1.88 (1.10 ~ 3.20)

(2) 各年齡分層之勝算比:

年齡≥60 歲: $\hat{OR}_1$=0.95 (0.33~2.79)

年齡<60 歲: $\hat{OR}_2$=1.87 (1.01 ~ 3.48)

(3) 調整年齡後的勝算比:

(i) 合併加權估計式 (Pooled estimator):

$$Var\left\{\log \hat{OR}_1\right\}=\frac{1}{9}+\frac{1}{6}+\frac{1}{115}+\frac{1}{73}=0.3002$$

$$Var\left\{\log \hat{OR}_2\right\}=\frac{1}{20}+\frac{1}{21}+\frac{1}{596}+\frac{1}{1171}=0.1002$$

$$\log 0.952=-0.049 \quad W_1=3.331$$

$$\log 1.871=0.627 \quad W_2=9.980$$

$$\log \hat{OR}_W=\frac{(3.331)(-0.049)+(9.980)(0.627)}{3.331+9.980}=0.457$$

調整年齡後的勝算比為: $\hat{OR}_W$=1.58

$$Var\left\{\log(\hat{OR}_w)\right\}=\frac{1}{3.331+9.980}=0.075$$

調整年齡後的勝算比的 95%信賴區間為:

$$95\%CI:\ \exp(0.458\pm 1.96\times\sqrt{0.075})=(0.92,2.70)$$

(ii) Mantel-Haenszel 方法:

調整年齡後的勝算比為:

$$\hat{OR}_{MH}=\frac{\{(9\times 73)/203\}+\{(20\times 1171)/1808\}}{\{(115\times 6)/203\}+\{(596\times 21)/1808\}}=1.569$$

$$Var_{MH}\left\{\log\left(\hat{OR}_{MH}\right)\right\}=0.0761$$

調整年齡後勝算比的 95% 信賴區間為:

$$95\%\text{CI: } \exp(0.4504\pm 1.96\times\sqrt{0.0761})=(0.91,2.69)$$

→調整年齡之效果後,收縮壓(SBP)與心肌梗塞 (MI) 之相關性不顯著。

(iii) 檢驗年齡分層之勝算比是否同質 (Homogenous):

$$\frac{(-0.049-0.458)^2}{0.3002}+\frac{(0.627-0.458)^2}{0.1002}=1.141$$

$$P(X^2>1.141)=0.285$$

→各年齡層間的勝算比具有同質性 (無法推翻各年齡層間的勝算比為同質之虛無假說)

2. 迴歸模式 Modelling Approach

(1) 世代追蹤研究之例子

探討 CAT 是否會跟 CHD 有相關，即為先前提到的兩樣本比例檢定(proportional test, Y 的結果為二分法, binary outcome)，亦可利用羅吉斯迴歸的方式進行。

羅吉斯迴歸式:

Y=1: 發生心臟疾病(CHD) ,

Y=0 : 不發生心臟疾病(no CHD)

X: catecholamine (CAT) 之高低

$$\log it\ P(Y=1|x)=\log\frac{P(Y=1|X)}{1-P(Y=1|X)}=\alpha+\beta X,$$

$$P(Y=1|X)=\frac{\exp(\alpha+\beta X)}{1+\exp(\alpha+\beta X)}$$

若為世代研究 (cohort study)，除了 CAT 之外還要調整年齡(age)跟心電圖結果(ECG):

Y=1: 發生心臟疾病(CHD)

Y=0 : 不發生心臟疾病(no CHD)

X: CAT (高/低); Age (≥55 vs <55); ECG (不正常 vs 正常)

$$P(Y=1|\mathbf{X})=\frac{\exp(\alpha+\sum\beta_k x_k)}{1+\exp(\alpha+\sum\beta_k x_k)}$$

$$\text{logit}\,P(Y=1|\mathbf{X})=\alpha+\sum_{k=1}^{K}\beta_k x_k$$

模型 1: Log(P/(1-P))=α+β_1×(CAT=High)

模型 2: Log(P/(1-P))=α+β_1×(CAT=High)+ β_2×(ECG=Abnormal)

模型 3: Log (P/(1-P))=α+β_1×(CAT=High)+ β_2×(Age ≥55)

模型 4: Log(P/(1-P))=α+β_1×(CAT=High)+ β_2×(ECG=Abnormal)+β_3×(Age ≥55)

模型 5: Log(P/(1-P))=α+β_1×(CAT=High)+ β_2×(ECG=Abnormal)+β_3×(Age ≥55)

+γ_1×(CAT=High) ×(Age ≥55)

以分層分析的觀點來看，會分成年齡小於 55 下 ECG 正常及不正常 (age<55 ECG=0, age≥55 ECG=1)，以及年齡大於 55 下 ECG 正常及不正常(age<55 ECG=0, age≥55 ECG=1)，如此一來，隨著迴歸模型等式右側的解釋變項的不同構成了 2×2=4 ，四個類別的結果，故以羅吉斯迴歸模型來看即為模型 4 (Log(P/(1-P))=α+β_1×(CAT=High)+ β_2×(ECG=Abnormal)+β_3×(Age ≥55))。而迴歸模型中對應到的變項之係數β取指數後即為 OR:

模型 1 的β_1取指數即為 CAT 的粗勝算比 (crude odds ratio);

模型 2 的β_1取指數即為調整心電圖結果後的 CAT 勝算比 (odds ratio with the adjustment for ECG);

模型 3 的β_1取指數即為調整年齡後的 CAT 勝算比 (odds ratio with the adjustment for age);

模型 4 的β_1取指數即為同時調整年齡及心電圖結果後的 CAT 勝算比 (odds ratio with the adjustment for age and ECG);

利用羅吉斯迴歸得到對 CHD 風險之迴歸係數(標準差)估計結果列表如下:

變數	分層	*粗率* (模型 1)	調整心電圖 (模型 2)	調整年齡 (模型 3)	調整年齡及心電圖 (模型 4)	年齡及 CAT 之交互作用 (模型 5)
CAT	High vs Low	0.8959	0.7601	0.6491	0.5160	0.6374
		(0.2646)	(0.2918)	(0.2912)	(0.3146)	(0.5971)
ECG	Abnormal vs		0.3191		0.3190	0.3156
	Normal		(0.2861)		(0.2858)	(0.2862)
Age	≥55 vs <55			0.5457	0.5451	0.5803
				(0.2824)	(0.2821)	(0.3184)
CAT*Age						-0.1583
						(0.6686)

交互作用: γ_1=-0.1583 (SE=0.6686), Wald χ^2=0.0561, P-value=0.8128

其他可能的交互作用

交互作用	迴歸係數	標準差	*p*
ECG*Age	-0.465	0.536	0.39
CAT*ECG	-0.378	0.569	0.51
Age*CAT*ECT	-0.367	0.507	0.47

如同在分層分析時,進行調整前須先確定每個分層的 RR 是否相同 (同質)，

即確定是否有交互作用，羅吉斯迴歸也可以檢定變項間是否存在交互作用, 以 CAT 及年齡為例, 檢定方法為在模型中加入一個變項為(CAT*年齡)表示 CAT 及年齡的交互作用(模型 5)檢定其對應的迴歸係數是否顯著 (γ_1)。 由以上報表看出 CAT 與年齡的交互作用是不顯著的 (γ_1=-0.1583 (SE=0.6686), Wald χ^2=0.0561, P-value=0.8128)。

(2) 病例-對照研究

羅吉斯迴歸分析病例-對照研究 (Modeling approach: logistic regression in case-control study, ebook 01. ch10.6, ebook 02. ch12 p.296)

以病例-對照試驗設計研究收縮壓與心肌梗塞之相關性 (Case-control study on SBP and MI))：利用羅吉斯迴歸方法估計調整年齡後收縮壓對心肌梗塞的勝算比

理論部分：

若有關於暴露與疾病之病例-對照研究資料如下：

	E (X=1)	$\overline{E}$ (X=0)
D	P_E	$P_{\overline{E}}$
$\overline{D}$	$1-P_E$	$1-P_{\overline{E}}$

則運用 logit 轉換 ($\log(\frac{P}{1-P})$) 可以連結發生疾病的勝算 (等式右側) 以及暴露 (等式右側) 如下:

$$\log(\frac{P}{1-P}) = \alpha + \beta X + \varepsilon$$

$$\log(\frac{P_E}{1-P_E}) = \hat{\alpha} + \hat{\beta} \qquad \text{(暴露組 (X=1))}$$

$$\log(\frac{P_{\bar{E}}}{1-P_{\bar{E}}}) = \hat{\alpha} \qquad (\text{非暴露組 (X=0)})$$

因此勝算比即為 $$\log(\frac{P_E / 1-P_E)}{P_{\bar{E}} / 1-P_{\bar{E}})} = \hat{\beta}, \quad \hat{OR} = \exp\{\hat{\beta}\}$$

所計算的是暴露勝算，由於利用病例-對照研究得到的暴露勝算比是世代研究中得到的疾病勝算比的不偏估計，故可以利用相同的羅吉斯迴歸式連結暴露與疾病，所得到的勝算比亦是疾病勝算比的不偏估計；但此時的截距項 α，由於病例-對照研究的無分母特性無法直接解釋為族群的疾病風險。

若利用羅吉斯迴歸模式分析如下：

反應變項 (dependent variable, outcome)

Y=1: 發生心肌梗塞(case, MI), Y=0:未發生心肌梗塞(control, no MI)

解釋變項 (independent variable)

X: 收縮壓(SBP : ≥140 vs. <140, 年齡(Age , ≥60 vs. <60)

(1) 檢驗收縮壓與年齡間是否有交互作用(interaction，各年齡組間的勝算比是否同質)

統計模式: $\log it(P(Y=1)) = \alpha + \beta_1 \times SBP + \beta_2 \times Age + r(SBP \times Age)$

假說檢定: $H_0 : r = 0 \quad vs. \quad r \neq 0$

估計結果:

$$\hat{r} = -0.676, \; sd(\hat{r}) = 0.633$$

$$(\frac{\hat{r}}{sd(\hat{r})})^2 = 1.14, \quad p = 0.29$$

Wald test
(大樣本理論，Asymptotic property)

→無法推翻虛無假說 (收縮壓與年齡不具交互作用，各年齡組間的勝算比具同質性)

統計報表：

```
proc logistic data=cscn;
   model mi(event='1')=age sbp age*sbp;
   weight n;
run;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.0211	0.2202	333.5706	<.0001
SBP	1	0.6266	0.3165	3.9201	0.0477
age	1	1.5224	0.4784	10.1290	0.0015
SBP*age	1	-0.6756	0.6327	1.1402	0.2856

(2) 估計調整年齡後收縮壓(SBP)的勝算比

統計模式: $\log it(P(Y=1)) = \alpha + \beta_1 \times SBP + \beta_2 \times Age$

估計結果:

$$\hat{\beta}_1 = 0.46,\ sd(\hat{\beta}_1) = 0.28$$
$$95\%\ CI\ of\ \hat{\beta}_1 = (-0.08, 1.01)$$
$$\hat{OR} = e^{0.46} = 1.59$$
$$95\ CI\ of\ \hat{OR} = (e^{-0.08}, e^{1.01}) = (0.92, 2.75)$$

統計報表:

```
proc logistic data=cscn;
   model mi(event='1')=age sbp;
   weight n;
run;
```

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.9458	0.2013	384.2516	<.0001
SBP	1	0.4645	0.2796	2.7603	0.0966
age	1	1.1124	0.3198	12.0992	0.0005

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
SBP	1.591	0.920	2.753
age	3.042	1.625	5.693

比較利用合併加權估計式 (pooled estimator), Mantel-Haenszel 方法以及羅吉斯迴歸方法所得到的調整年齡後收縮壓(SBP)對心肌梗塞(MI)之勝算比:

Method	Point estimate	95% CI
Pooled method	1.58	(0.92, 2.70)
Mantel-Haenszel	1.57	(0.91, 2.69)
Logistic regression	1.59	(0.92, 2.75)

補充教材

Wald test (大樣本理論，Asymptotic property)

在估計參數 θ 時，在大樣本理論下之進似的對數概似函數值表示為 $-\frac{1}{2}\left(\frac{M-\theta}{S}\right)^2$，其中 M 為最大概似函數估計值(Maximum likelihood estimate, MLE)，S 為 MLE 下之標準差，而此最大概似函數估計值具有較高之準確度，因此我們採用 $\left(\frac{M-\theta_\varphi}{S}\right)^2$ 代表在虛無假說下之參數值(null value) 的負兩倍對數概似函數比值(-2 log likelihood ratio)，我們以檢定 M 值來決定是否拒絕虛無假說。

$$\left(\frac{M-\theta_\varphi}{S}\right)^2 \sim \chi^2 \text{ 自由度 (d.f) =1}$$

上式可以解釋成 M 與 θ_φ(虛無假說)的距離有多遠，當距離很遠時，則推翻虛無假說，如距離不遠時，則接受虛無假說。

在上述以配對式病例-對照試驗設計研究收縮壓與心肌梗塞之相關性研究例子中：

$$\log \text{odds} = \log \frac{\text{D}}{\text{N - D}}$$

$$\text{M} = \log\left(\frac{\text{D}}{\text{N - D}}\right) \quad S = \sqrt{\frac{1}{D}+\frac{1}{N-D}}$$

$$\text{故 M} = \log\frac{13}{11} = 0.1672 \quad S = \sqrt{\frac{1}{13}+\frac{1}{11}} = 0.4097$$

虛無假說下 logΩ=0.0，log likelihood ratio 為

$$-\frac{1}{2}\left(\frac{0.1672-0.0}{0.4097}\right)^2 = -0.0833$$

故近似之-2 log likelihood 為 $\left(\frac{10.1672-0.0}{0.4097}\right)^2 = 0.1665$，P=0.6831，無法推翻

虛無假說。

流行病學方法論及實驗
(The Methods of Epidemiology and Practices):
觸媒流行病學故事與護理照護

授課教師：陳秀熙 教授

台灣大學流行病學與預防醫學研究所

Effect Modification- Definition

- If the relationship between X and Y varies with Z, Z is possibly an effect modifier. That means the interaction between X and Z in association with the outcome Y exists.
- Catalyst in Chemical Reaction

Effect Modification-Demonstration (1)

(1) Modified effect due to Z

	Z(+)			Z(-)			Aggregated	
	X(+)	X(-)		X(+)	X(-)		X(+)	X(-)
Y(+)	80	20		50	30		130	50
Y(-)	60	20		40	90		100	110
	OR=1.33			OR=3.75			OR=2.86 (Crude)	

Effect Modification Demonstration (2)

(2) Opposite effects due to Z

	Z(+)			Z(-)			Aggregated	
	X(+)	X(-)		X(+)	X(-)		X(+)	X(-)
Y(+)	180	60		40	140		220	200
Y(-)	120	90		80	110		200	200
	OR=2.25			OR=0.39			OR=1.10 (Crude)	

Effect Modification-Interpretation(1)

1. The effect of X on Y varies with the presence of Z. The odds ratio for the association between X and Y is larger in the absence of Z than that in the presence of Z.
2. The opposite effect is demonstrated by the presence of Z.

Effect Modification-Interpretation(2)

3. The presentation of aggregated results is not adequate when the effect modification exists.
4. Factorial randomized controlled trial design is needed to test interaction between X and Z.

Effect Modification with (stratified analysis)

在研究catecholamine（CAT）與心臟疾病(CHD)之關係的研究中，年齡的大小(Age)及心電圖是否異常(ECG)為兩個可能的干擾因子

特定分層之RR

年輕組($Age<55$),
心電圖無異常($ECG=0$)

	High CAT	Low CAT	Total
CHD	1	17	18
No CHD	7	257	264
Total	8	274	282

$\hat{RR}_1 = 2.015(0.30 \sim 13.34)$

年輕組($Age<55$),
心電圖有異常($ECG=1$)

	High CAT	Low CAT	Total
CHD	3	7	10
No CHD	14	52	66
Total	17	59	76

$\hat{RR}_2 = 1.49(0.43 \sim 5.14)$

年老組($Age\geq55$),
心電圖無異常($ECG=0$)

	High CAT	Low CAT	Total
CHD	9	15	24
No CHD	30	107	137
Total	39	122	161

$\hat{RR}_3 = 1.88(0.89 \sim 3.95)$

年老組($Age\geq55$),
心電圖有異常($ECG=1$)

	High CAT	Low CAT	Total
CHD	14	5	19
No CHD	44	27	71
Total	58	32	90

$\hat{RR}_4 = 1.54(0.61 \sim 3.90)$

Crude RR=0.221/0.09=2.46

7

評估因子是否造成效應修飾 (Effect modification)

$$\sum_{i=1}^{k} \frac{\left\{ \log \hat{RR}_i - \log \hat{RR} \right\}^2}{Var\left\{ \log \hat{RR} \right\}}$$

8

效應修飾之評估
(Evaluation of effect modification)

範例: 在研究catecholamine (CAT) 與心臟疾病(CHD)之關係的研究中，欲檢定各分層間RR之同質性:

□□□□□ $\sum_{i=1}^{k} \frac{(\log \hat{RR}_i - \log RR_W)^2}{Var(\log RR_i)}$

$$= \frac{(0.701-0.538)^2}{0.9302} + \frac{(0.397-0.538)^2}{0.4005} + \frac{(0.630-0.538)^2}{0.1439} + \frac{(0.435-0.538)^2}{0.2229} = 0.185$$

$P(X_3^2 > 0.185) = 0.980$

CAT及CHD間的關係不會被年齡及ECG所修飾(無效應修飾)，可繼續進行干擾因子的調整

9

以病例-對照試驗設計研究收縮壓與心肌梗塞之相關性
(Case-control study on SBP and MI)

	年齡≥60 歲 收縮壓 (SBP)			年齡<60 歲 收縮壓 (SBP)		
心肌梗塞	≥140	<140		≥140	<140	
發生 (Case)	9	6	15	20	21	41
未發生 (Control)	115	73	188	596	1,171	1,767
總人數	124	79	203	616	1,192	1,808

- 粗率 (Crude odds ratio) : 1.88 (1.10 ~ 3.20)
- 各年齡分層之勝算比:
  - 年齡≥60歲: 0.95 (0.33~2.79)
  - 年齡<60歲: 1.87 (1.01 ~ 3.48)
- 調整年齡後的勝算比為:
  - pooled: 1.58 (0.92,2.70)
  - Mantel-Haenszel: 1.57 (0.91~2.69)

→調整年齡之效果後，收縮壓(SBP)與心肌梗塞 (MI) 之相關性不顯著。

10

疾病勝算比與暴露勝算比
(disease OR and exposure OR)

E 代表暴露, $\bar{E}$代表未暴露; D代表罹病, $\bar{D}$代表未罹病

	E	$\bar{E}$
D	a	b
$\bar{D}$	c	d

世代研究

$$\frac{\text{暴露組的疾病勝算(disease odds for E)}}{\text{非暴露組的疾病勝算(disease odds for } \bar{E})} = \frac{\frac{P(D|E)}{P(\bar{D}|E)}}{\frac{P(D|\bar{E})}{P(\bar{D}|\bar{E})}} = \frac{\frac{a/(a+c)}{c/(a+c)}}{\frac{b/(b+d)}{d/(b+d)}} = \frac{a/c}{b/d} = \frac{ad}{bc}$$

病例-對照研究

$$\frac{\text{罹病組的暴露勝算(exposure odds for D)}}{\text{未罹病組的暴露勝算(exposure odds for } \bar{D})} = \frac{\frac{P(E|D)}{P(\bar{E}|D)}}{\frac{P(E|\bar{D})}{P(\bar{E}|\bar{D})}} = \frac{\frac{a/(a+b)}{b/(a+b)}}{\frac{c/(c+d)}{d/(c+d)}} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

11

調整干擾因子的方法

第 i 個分層的聯列表資料

	E	$\bar{E}$	
D	a_i	b_i	m_{1i}
$\bar{D}$	c_i	d_i	m_{2i}
	n_{1i}	n_{2i}	n_i

合併加權估計式 (Pooled estimator)

$$\hat{OR}_i = \frac{a_i d_i}{b_i c_i}$$

$$Var\left\{\log(\hat{OR}_i)\right\} = \frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i}$$

$$\log \hat{OR}_W = \frac{\sum_{i=1}^{k} W_i \log \hat{OR}_i}{\sum_{i=1}^{k} W_i}$$

$$W_i = \frac{1}{Var\{\log(\hat{OR}_i)\}}$$

Mantel-Haenszel 方法

$$\hat{OR}_{MH} = \frac{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)\left(\frac{a_i d_i}{b_i c_i}\right)}{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)} = \frac{\sum_{i=1}^{k}\left(\frac{a_i d_i}{n_i}\right)}{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)}$$

$$Var_{MH}\left\{\log(\hat{OR}_{MH})\right\} = \frac{\sum_{i=1}^{k}\left(\frac{a_i d_i}{n_i}\right)^2\left[\frac{1}{a_i}+\frac{1}{b_i}+\frac{1}{c_i}+\frac{1}{d_i}\right]}{\left(\sum_{i=1}^{k}\frac{a_i d_i}{n_i}\right)^2}$$

12

研究收縮壓與心肌梗塞之相關性

粗率 (Crude odds ratio) : 1.88 (1.10 ~ 3.20)
年齡≥60歲: 0.95 (0.33~2.79)
年齡<60歲: 1.87 (1.01 ~ 3.48)

已知年齡可能為干擾因子

	年齡≥60 歲 收縮壓 (SBP)			年齡<60 歲 收縮壓 (SBP)		
心肌梗塞	≥140	<140		≥140	<140	
發生(Case)	9	6	**15**	20	21	**41**
未發生 (Control)	115	73	**188**	596	1,171	**1,767**
總人數	**124**	**79**	**203**	**616**	**1,192**	**1,808**

合併加權估計式 (Pooled estimator)

$95\%CI$: $\exp(0.458 \pm 1.96 \times \sqrt{0.075}) = (0.92, 2.70)$

Mantel-Haenszel 方法

調整年齡後的勝算比為: $\hat{OR}_w = 1.569$

95%CI: $\exp(0.4504 \pm 1.96 \times \sqrt{0.0761}) = (0.91, 2.69)$

13

效應修飾之評估 (Evaluation of effect modification)

範例: 在研究收縮壓與心肌梗塞之關係的研究中，欲檢定不同年齡分層間OR之同質性:

□□□□□ $\sum_{i=1}^{k} \frac{\{\log \hat{OR}_i - \log \hat{OR}\}^2}{Var\{\log \hat{OR}\}}$

$$\frac{(-0.049 - 0.458)^2}{0.3002} + \frac{(0.627 - 0.458)^2}{0.1002} = 1.141$$

$$P(\chi_1^2 > 1.141) = 0.285$$

收縮壓及心肌梗塞間的關係不會被年齡所修飾(無效應修飾)，可進行干擾因子的調整

14

羅吉斯迴歸分析 -病例-對照研究 -

eBook 04, p. 293

暴露與疾病之病例-對照研究資料

	E	$\bar{E}$
	X=1	X=0
D	$P(E\|D)$	$P(\bar{E}\|D)$
$\bar{D}$	$P(E\|\bar{D})$	$P(\bar{E}\|\bar{D})$

$$OR = \frac{P(E|D,S)/P(\bar{E}|D,S)}{P(E|\bar{D},S)/P(\bar{E}|\bar{D},S)} = \frac{P(D,S|E)P(E)/P(D,S|\bar{E})P(\bar{E})}{P(\bar{D},S|E)P(E)/P(\bar{D},S|\bar{E})P(\bar{E})} = \frac{\pi_1.P(D|E)/\pi_1.P(D|\bar{E})}{\pi_0.P(\bar{D}|E)/\pi_0.P(\bar{D}|\bar{E})} = \frac{P_1/1-P_1}{P_0/1-P_0}$$

$P(D,S|E) = P(S|D,E)P(D|E)$

$P(S|D,E) = P(S|D,\bar{E}) = P(S|D) = \pi_1$
$P(S|\bar{D},E) = P(S|\bar{D},\bar{E}) = P(S|\bar{D}) = \pi_0$

15

羅吉斯迴歸分析 -病例-對照研究 -

暴露與疾病之世代研究資料

	E	$\bar{E}$
	X=1	X=0
D	P_1	P_0
$\bar{D}$	$1-P_1$	$1-P_0$

$$\log(\frac{P}{1-P}) = \alpha + \beta X + \varepsilon$$

X=1 ➔ $\log(\frac{P_1}{1-P_1}) = \hat{\alpha} + \hat{\beta}$

X=0 ➔ $\log(\frac{P_0}{1-P_0}) = \hat{\alpha}$

＊ 無分母研究

$$\log(\frac{P_1/1-P_1}{P_0/1-P_0}) = \hat{\beta}, \quad \hat{OR} = \exp\{\hat{\beta}\}$$

16

檢驗收縮壓與年齡間是否有交互作用

Y=1: 發生心肌梗塞(case, MI)
Y=0:未發生心肌梗塞(control, no MI)
X: 收縮壓 (SBP: ≥140 vs <140), 年齡 (Age: ≥60 vs <60)

$$\log it(P(Y=1)) = \alpha + \beta_1 \times SBP + \beta_2 \times Age + r(SBP \times Age)$$

$H_0: r = 0$ vs. $H_1: r \neq 0$

若存在交互作用 ($\gamma \neq 0$)

Age: ≥60 SBP=1 $\log(Odds_{Age\geq 60}) = \alpha + \beta_1 + \beta_2 + \gamma$
SBP=0 $\log(Odds_{Age\geq 60}) = \alpha + \beta_2$ ➔ $OR_{Age\geq 60} = \exp(\beta_1 + \gamma)$

Age: <60 SBP=1 $\log(Odds_{Age<60}) = \alpha + \beta_1$
SBP=0 $\log(Odds_{Age<60}) = \alpha$ ➔ $OR_{Age<60} = \exp(\beta_1)$

17

檢驗收縮壓與年齡間是否有交互作用

$$\log it(P(Y=1)) = \alpha + \beta_1 \times SBP + \beta_2 \times Age + r(SBP \times Age)$$

$H_0: r = 0$ vs. $H_1: r \neq 0$

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.0211	0.2202	333.5706	<.0001
SBP	1	0.6266	0.3165	3.9201	0.0477
age	1	1.5224	0.4784	10.1290	0.0015
SBP*age	1	-0.6756	0.6327	1.1402	0.2856

$\hat{r} = -0.676$, $sd(\hat{r}) = 0.633$

$(\frac{\hat{r}}{sd(\hat{r})})^2 = 1.14$, $p = 0.29$

Wald test
大樣本理論 (Asymptotic property)

➔無法推翻虛無假說 (收縮壓與年齡不具交互作用，各年齡組間的勝算比具同質性)

18

Genetic predisposition, Western dietary pattern, and the risk of type 2 diabetes in men[1–3]

Lu Qi, Marilyn C Cornelis, Cuilin Zhang, Rob M van Dam, and Frank B Hu

- 第二型糖尿病的發生風險與西式飲食以及基因的感受性（genetic susceptibility）有關
- 對於gene-lifestyle interaction的證據仍不足

19

Objective: The objective was to assess whether established genetic variants, mainly from genomewide association studies, modify dietary patterns in predicting diabetes risk.

Design: We determined 10 polymorphisms in a prospective, nested, case-control study of 1196 diabetic and 1337 nondiabetic men. A genetic risk score (GRS) was generated by using an allele counting method. Baseline dietary intakes were collected by using a semiquantitative food-frequency questionnaire. We used factor analysis to derive Western and "Prudent" dietary patterns from 40 food groups.

20

TABLE 1
Baseline characteristics of the diabetic patients and nondiabetic control subjects[1]

	Nondiabetic	Diabetic	*P*
No. of participants	1337	1196	—
Age (y)	55 ± 9[2]	56 ± 8	0.1
BMI (kg/m^2)	25 ± 2.8	27.8 ± 4.1	0.22
Obesity (%)[3]	5.3	25.5	<0.0001
Alcohol use (g/d)	12.2 ± 15.5	11.2 ± 16.6	0.01
Physical activity (MET/wk)	21.3 ± 27.4	14.6 ± 18.8	0.23
Current smoker (%)	7.0	11.3	0.0006
Family history of diabetes (%)	13.0	32.4	<0.0001
Total energy intake (kcal/d)	2039 ± 634	2031 ± 604	0.73

[1] MET, metabolic equivalent task. The geometric means of continuous variables were compared by using general linear models, and the proportions of categorical variables were compared by using chi-square tests.
[2] Mean ± SD (all such values).
[3] Defined as a BMI ≥ 30.

21

勝算比 (Odds Ratio)

- 勝算 (Odds) :某事件發生的機率與某事件沒發生機率的比值。= P/(1-P)
- 勝算比 (Odds Ratio, OR) : 實驗組中發生疾病的勝算與控制組中發生疾病的勝算比值, 或罹患疾病的病患暴露於某變因的勝算除以控制組暴露的勝算

22

TABLE 3
Interactions between dietary patterns and the genetic risk score in relation to diabetes risk[1]

	Dietary patterns[2]					
Genetic risk score[2]	Q1 (lowest)	Q2	Q3	Q4 (highest)	*P* for trend	*P* for interaction
Western dietary pattern						
<10 (*n* = 503)	1	0.79 (0.46, 1.38)	0.81 (0.48, 1.37)	1.07 (0.65, 1.76)	0.69	0.02
10–11 (*n* = 904)	1	0.98 (0.67, 1.44)	1.02 (0.69, 1.49)	1.40 (0.97, 2.01)	0.06	—
≥12 (*n* = 1126)	1	1.23 (0.88, 1.73)	1.49 (1.06, 2.09)	2.06 (1.48, 2.88)	0.01	—
Prudent dietary pattern						
<10 (*n* = 503)	1	0.85 (0.50, 1.44)	1.07 (0.65, 1.76)	1.29 (0.79, 2.11)	0.24	NS
10–11 (*n* = 904)	1	0.75 (0.52, 1.07)	0.81 (0.56, 1.18)	0.77 (0.53, 1.11)	0.21	—
≥12 (*n* = 1126)	1	0.81 (0.58, 1.14)	0.71 (0.51, 0.99)	0.81 (0.59, 1.13)	0.16	—

[1] The analyses were adjusted for age, BMI, smoking, alcohol consumption, physical activity, family history of diabetes, and total energy intakes. Q, quartile.

genetic risk score 較高者，西式飲食會增加糖尿病的風險

而genetic risk score 較低者（<10），西式飲食無顯著增加其罹患糖尿病風險

23

TABLE 5
Interactions between genetic risk score and individual foods and nutrients characterizing the Western dietary pattern[1]

	Western dietary patterns[2]					
Foods and genetic risk score	Q1 (lowest)	Q2	Q3	Q4 (highest)	*P* for trend	*P* for interaction
Red meat						
<10	1	0.54 (0.24, 1.23)	0.88 (0.39, 1.97)	0.81 (0.36, 1.80)	0.55	0.02
10–11	1	1.27 (0.75, 2.15)	1.16 (0.68, 1.97)	1.45 (0.86, 2.44)	0.23	—
≥12	1	1.03 (0.69, 1.54)	1.32 (0.87, 2.01)	2.42 (1.58, 3.70)	<0.0001	—
Processed meat						
<10	1	0.91 (0.45, 1.85)	0.70 (0.38, 1.28)	1.12 (0.66, 1.92)	0.76	0.029
10–11	1	0.98 (0.60, 1.61)	0.96 (0.62, 1.48)	1.47 (0.99, 2.20)	0.06	—
≥12	1	1.09 (0.72, 1.67)	1.88 (1.30, 2.71)	2.01 (1.41, 2.89)	<0.0001	—
Heme iron						
<10	1	1.06 (0.57, 1.96)	0.82 (0.45, 1.49)	0.87 (0.48, 1.56)	0.47	0.0004
10–11	1	1.19 (0.77, 1.85)	1.37 (0.89, 2.10)	1.85 (1.23, 2.77)	0.002	—
≥12	1	0.71 (0.49, 1.03)	1.24 (0.86, 1.80)	2.48 (1.72, 3.56)	<0.0001	—

[1] The analyses were adjusted for age, BMI, smoking, alcohol consumption, physical activity, family history of diabetes, and total energy intakes.
[2] Values are odds ratios (95% CIs) calculated by using an unconditional logistic regression model.

24

修飾因子(effect modifier)

Journal of the American College of Cardiology
© 2005 by the American College of Cardiology Foundation
Published by Elsevier Inc.

Vol. 46, No. 5, 2005
ISSN 0735-1097/05/$30.00
doi:10.1016/j.jacc.2005.05.060

The Effect of Losartan Versus Atenolol on Cardiovascular Morbidity and Mortality in Patients With Hypertension Taking Aspirin

The Losartan Intervention for Endpoint Reduction in Hypertension (LIFE) Study

高血壓病患常見併發症：
中風
心肌梗塞

OBJECTIVES We conducted a subgroup analysis in the Losartan Intervention For Endpoint reduction in hypertension (LIFE) study to determine whether aspirin interacted with the properties of losartan, an angiotensin-II receptor antagonist.

BACKGROUND Negative interactions between angiotensin-converting enzyme inhibitors and aspirin have been reported. There are no data reported from clinical trials about possible interactions between angiotensin-II receptor antagonists and aspirin.

METHODS The LIFE study assigned 9,193 patients with hypertension and left ventricular hypertrophy (LVH) to losartan- or atenolol-based therapy for a mean of 4.7 years, with 1,970 (21.4%) taking aspirin at baseline. The primary composite end point (CEP) included cardiovascular death, stroke, and myocardial infarction (MI). The present cohort was stratified by aspirin use at baseline.

RESULTS Blood pressures were reduced similarly in the losartan with aspirin (n = 1,004) and atenolol with aspirin (n = 966) groups. The CEP was reduced by 32% (95% confidence interval 0.55 to 0.86, p = 0.001) with losartan with aspirin compared to atenolol with aspirin, adjusted for Framingham risk score and LVH. The test for treatment versus aspirin interaction, excluding other covariates, was significant for the CEP (p = 0.016) and MI (p = 0.037).

CONCLUSIONS There was a statistical interaction between treatment and aspirin in the LIFE study, with significantly greater reductions for the CEP and MI with losartan in patients using aspirin than in patients not using aspirin at baseline. Further studies are needed to clarify whether this represents a pharmacologic interaction or a selection by aspirin use of patients more likely to respond to losartan treatment. (J Am Coll Cardiol 2005;46:770–5) © 2005 by the American College of Cardiology Foundation

有吃aspirin者

End Point	Losartan (n = 1,004) n	%	Rate†	Atenolol (n = 966) n	%	Rate†	Adjusted* Hazard Ratio (95% CI)	p Value
Primary composite end point‡	128	12.7	28.3	180	18.6	42.1	0.68 (0.55–0.86)	0.001
Cardiovascular mortality	56	5.6	11.8	76	7.9	16.7	0.73 (0.52–1.03)	0.074
Stroke	61	6.1	13.4	94	9.7	21.8	0.63 (0.45–0.86)	0.004
Myocardial infarction	44	4.4	9.6	58	6.0	13.1	0.75 (0.51–1.11)	0.16
Other prespecified end points								
Total mortality	106	10.6	22.4	121	12.5	26.6	0.86 (0.66–1.12)	0.26
Hospitalization for								
Angina pectoris	53	5.3	11.6	48	5.0	10.9	1.10 (0.74–1.62)	0.64
Heart failure	45	4.5	9.8	53	5.5	12.1	0.84 (0.56–1.25)	0.39
Revascularization	100	10.0	22.5	109	11.3	25.5	0.91 (0.70–1.20)	0.51
New-onset diabetes§	58	6.9	15.3	57	7.1	15.6	0.98 (0.68–1.41)	0.91

沒有吃aspirin者

End Point	Losartan (n = 3,601) n	%	Rate†	Atenolol (n = 3,622) n	%	Rate†	Adjusted* Hazard Ratio (95% CI)	p Value
Primary composite end point‡	380	10.6	22.6	408	11.3	24.3	0.95 (0.82–1.09)	0.46
Cardiovascular mortality	148	4.1	8.5	158	4.4	9.1	0.96 (0.77–1.20)	0.71
Stroke	171	4.7	10.1	215	5.9	12.7	0.80 (0.66–0.98)	0.034
Myocardial infarction	154	4.3	9.0	130	3.6	7.6	1.21 (0.96–1.53)	0.11
Other prespecified end points								
Total mortality	277	7.7	15.9	310	8.6	17.8	0.91 (0.77–1.07)	0.24
Hospitalization for								
Angina pectoris	107	3.0	6.3	93	2.6	5.4	1.18 (0.89–1.56)	0.25
Heart failure	108	3.0	6.3	108	3.0	6.3	1.02 (0.78–1.34)	0.86
Revascularization	161	4.5	9.5	175	4.8	10.3	0.94 (0.76–1.17)	0.58
New-onset diabetes§	184	5.8	12.4	263	8.3	17.9	0.70 (0.58–0.85)	<0.001

五、 偏好選擇流行病學故事之護理觀

(一) 造成觀察性研究中推論偏誤的可能原因

1. 選擇偏差(Selection Bias)

雌激素補充療法是否可以減少停經後女性心血管疾患 (CHD)的發生?

為了瞭解上述關於雌激素補充療法與停經後女性心血管疾患的相關性， 研究人員進行了觀察性研究:

A 組: 接受雌激素補充療法(postmenopasual estrogen)

B 組: 未接受雌激素補充療法

觀察結果:A 組發生心血管疾患(CHD) << B 組發生心血管疾患的比例

因此研究結論為: 停經後的雌激素補充療法可以有效降低心血管疾患的發生。

關於此研究的爭論:

接受激素補充療法的受試者可能本來就較健康， 因此發生心血管疾患的可能性本來就低於未接受雌激素補充療法的受試者。

這樣之誤差稱為選擇性偏差 (Selection bias):主要由於介入組與非介入組的特性本來就有差異， 因此兩組間缺乏可比較性 (comparability)，在一般流行病學，參與者通常教育水準較高、較健康，生活型態亦較佳，因此介入性所得到好之結果可能是介入組較健康而非介入影響。最近有關素食者較非素食者較容易得心血管及癌症之研究，也可能是選擇性偏差之結果。

2. 選擇性抽樣偏差(ebook 03. p52; ebook 04. p.22, p.83, p.93, p.112, p.114)

(1) 病例對照研究抽樣偏差

由於病例-對照研究屬於無分母的研究(並非觀察整個族群，而是對族群抽樣

進行分析)，因此病例-對照研究在計算暴露勝算比實際上用的是暴露人數比上非暴露人數 (a/b 或 c/d)，計算勝算比時則利用罹病組的暴露勝算比上非罹病組的暴露勝算 ($\frac{a/b}{c/d}$)，由此可知，對照組的選取對於最後所得到的暴露勝算比有很大的影響。

舉例而言，利用病例-對照試驗設計研究抽菸與肺癌的關係，若以慢性阻塞性肺疾 (chronic obstructive pulmonary disease, COPD) 作為對照組，由於慢性阻塞性肺疾患者有高比例的抽菸暴露，因此在計算暴露勝算比時 c 的人數就會上升，因此最後會得到低估的暴露勝算比，而抽菸與肺癌的相關性就會同樣的被低估。相反的，若使用參加健檢的族群作為對照組，由於參與健檢者的抽菸暴露比例可能低於一般族群，因此 d 的人數就會上升，所以最後得到的暴露勝算比會有高估的情形，而抽菸與肺癌的相關性就會同樣的被高估。綜合上述兩種狀況，在病例-對照研究中，對於對照組的選取，理想的狀況是個體能否成為研究的對照組，其條件是對照組之選取與有暴露最好是沒有相關，且對照組的暴露風險必須能夠代表病例組所來自的族群之暴露風險。由於病例對照研究所得到的是暴露比而非絕對值，因此無法得到真正兩組之發生率。

臨床實例 1: 使用手機是否會影響人類健康？

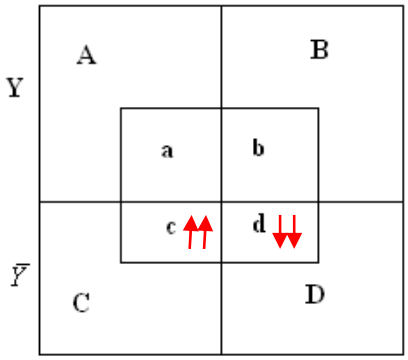
使用手機對於健康的影響一直以來都是備受爭議的話題。許多國家的研究結果指出曾經**頻繁使用手機**的人比較**不容易**得到某些腦部腫瘤。這樣的結果可以相信嗎？當然可能有問題，其原因大部份是因為參加對照組之民眾使用手機之習慣無法代表腦部腫瘤病患所來自族群。為了了解此種參加民眾與未參加者手機使用習慣是否不同，研究者另進行再調查，結果發現參與研究的健康(對照)民眾中有59%的人經常頻繁地使用手機，而拒絕參與研究的民眾經由再次調查後，發現只有 34%的人頻繁地使用手機。此結果表示參與研究的對照組使用手機的頻率比一般民眾高。

2X2 列聯表之推理

(A) 低估

利用 2X2 列聯表呈現此種選擇性偏差，解釋研究結果為何會呈現 "使用手機的人比較不容易得到腦部腫瘤" 的現象。若"對照組使用手機的頻率高於一般民眾"，參與研究的健康民眾 (對照組)的樣本分布情形如下：

X (頻繁使用)　$\bar{X}$ (較少使用)

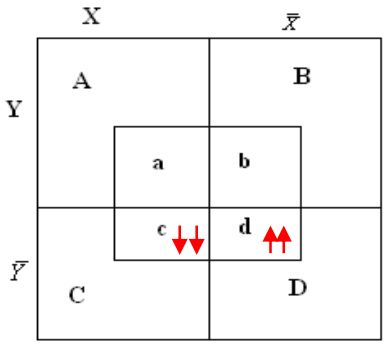


在上述關於是否頻繁使用手機與健康狀態關係研究之應用實例中，對照組(參與研究的健康民眾) 使用手機的比率高於一般群眾 (上圖中的 c (未罹病且使用手機)高估，d (未罹病且未使用手機) 減低，因此在計算相對勝算比(OR=ad/bc)時，會得到低估的 OR 估計值。

(B) 高估

假如調查發現參與研究的健康民眾(對照組)有 34%的人頻繁地使用手機。對於拒絕參與研究的民眾再次進行調查，發現有 59%的人頻繁地使用手機。這樣的數據表示參與研究的民眾(對照組)使用手機的頻率比一般民眾還低。如此一來選擇性偏差會對研究結果有甚麼樣的影響？

若此樣本分布情形如下，



所以對照組 (參與研究的健康民眾) 使用手機的比率低於一般群眾 (上圖中的 c (未罹病且使用手機)降低，d (未罹病且未使用手機) 增加，因此在計算相對

勝算比(OR=ad/bc)時，會得到高估的估計值。

臨床實例 2-(ebook 03. p54 table 6.6)

研究背景：麻疹、德國麻疹、腮腺炎混合疫苗(measles-mumps-rubella vaccination, MMR)中的保存劑 (Thimerosal, 一種有機汞) 有可能會導致自閉症 (autism)或其他行為發展上的異常

研究目的：探討麻疹、德國麻疹、腮腺炎混合疫苗施打(X)與行為發展異常(Y)的相關性。

(A) 假設研究設計為世代研究，我們追蹤了 206,300 人，其資料如下，

	X	$\bar{X}$	total
Y	1,000	300	1,300
$\bar{Y}$	160,000	45,000	205,000
total	161,000	45,300	206,300

藉由此資料列聯表，我們可以計算相對風險比(Relative risk, RR)以及疾病風險勝算比(Odds ratio, OR)如下，

$$\text{RR}=\frac{1000/161000}{300/45300}=\frac{0.62\%}{0.66\%}=0.93\text{，}\text{OR}=\frac{1000/160000}{300/45000}=0.94$$

(B) 然而在研究資源有限的情況下，我們抽樣一部分的群眾進行世代研究的追蹤，其抽樣比例針對有接受疫苗施打與未接受疫苗施打兩組分別為 0.2(π_X, 暴露組抽樣比例)與 0.5($\pi_{\bar{X}}$,非暴露組的抽樣比例)，收集到的資料成為：

	X	$\bar{X}$	total
Y	200	150	350
$\bar{Y}$	32,000	22,500	54,500
total	32,200	22,650	54,850

在此資料分布下，我們可以計算相對風險比(Relative risk, RR)以及疾病風險勝算比(Odds ratio, OR)如下，

$$\text{RR}=\frac{200/32200}{150/22650}=\frac{0.62\%}{0.66\%}=0.93\text{，OR}=\frac{200/32000}{150/22500}=0.94$$

與(a)情境中追蹤所有群眾得到的估計值相同。

(C) 假設接受疫苗施打且小朋友行為發展異常的家庭，參與研究中的比例較高，反之小朋友未接受疫苗施打的家庭中，不管小朋友有無行為發展異常參與研究的比例相同：

假設$\pi_{XY}=0.9$ (有接受疫苗且有發展異常)，$\pi_{X\bar{Y}}=0.4$ (有接受疫苗且無發展異常)，$\pi_{\bar{X}Y}=0.2$(未接受疫苗且有發展異常),$\pi_{\bar{X}\bar{Y}}=0.2$ (未接受疫苗且無發展異常) ，則收集到的資料如下，

	X	$\bar{X}$	total
Y	900	60	960
$\bar{Y}$	64,000	9,000	73,000
total	64,900	9,060	73,960

$$\text{RR}=\frac{900/64900}{60/9060}=\frac{1.39\%}{0.66\%}=2.09\text{，OR}=\frac{900/64000}{60/9000}=2.11$$

由抽樣的結果所計算的估計值有高估的情形。

(D) 若在研究設計為病例對照研究(case-control study)，其收集的資料分布如下(ebook 03. p.54 table 6.7)，

	X	$\bar{X}$	total
Y	1,000	300	1,300
$\bar{Y}$	3,550	950	4,500
total	4550	1250	5800

注意:與最初以世代研究收集資料的方式((a)情境),在病例對照研究設計中，

所有罹病個案皆納入分析而對照組的個案僅有 4,500 人,約占原來的 2%。

抽樣比例為$\pi_{XY} = 1$ (有接受疫苗且有發展異常)，$\pi_{X\bar{Y}} = 0.022$ (有接受疫苗且無發展異常)，$\pi_{\bar{X}Y} = 1$(未接受疫苗且有發展異常),$\pi_{\bar{X}\bar{Y}} = 0.021$ (未接受疫苗且無發展異常)

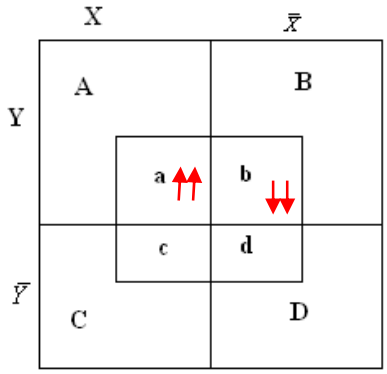
$$\text{OR}=\frac{1000/3550}{300/950}=0.89,$$

近似於原來以世代研究收集資料方式所估計出來的疾病風險勝算比。

3. 回憶性偏差(Recall bias)

在病例對照研究中，通常需要研究對象回憶其過去的情形。回憶性偏差通常會造成高估疾病與暴露的相關性。

利用 2X2 列聯表呈現回憶性偏差所造成的效果如下：

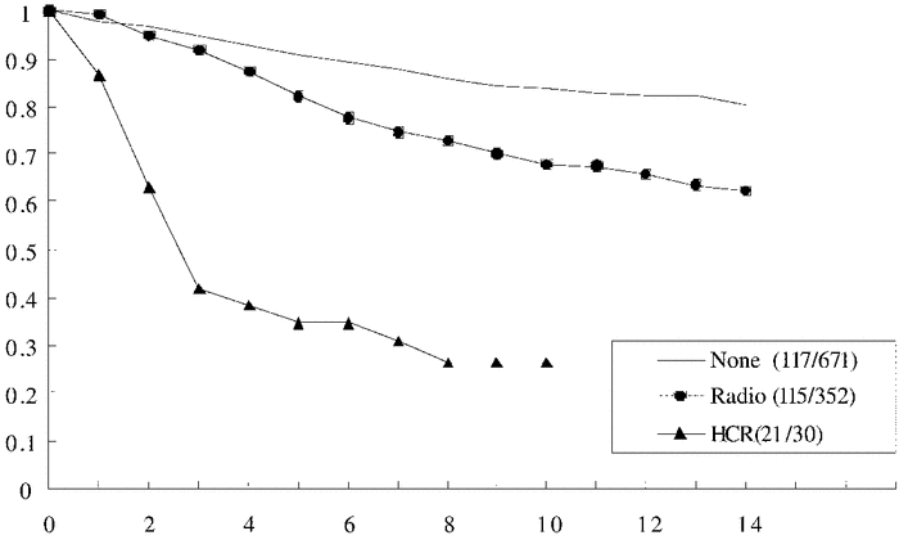


勝算比的計算為 OR=ad/bc; 在上述應用實例(1)的病例對照研究中罹病的人相對於對照組比較傾向於正確及努力回想起關於手機使用的經驗而造成回憶偏差(recall bias) 使得 a (罹病且有使用手機) 相對增加而 b (罹病且未使用手機)相對降低，因此估計得到的勝算比的值會高估。

4. 選擇性存活分析(Selected survival)

在臨床上許多治療存活曲線之比較會有選擇偏差之問題，下圖為單獨接受化學治療、同時接受化學治療與荷爾蒙藥物治療以及控制組的存活情形，由圖可以

看出同時接受化學治療與荷爾蒙藥物治療組的存活情形較差，其次為接受化學治療組，而未接受治療最好，這是因為同時接受化學治療與荷爾蒙藥物治療的病人其疾病嚴重度往往較僅接受化學治療組來得嚴重，因此若沒有考慮未治療前之疾病嚴重度比較上不公平，解決方式是將疾病嚴重度分層進行比較，而接受治療組比對照組疾病嚴重度來得差，亦即其存活情形，在未接受治療前本來就較差，這種情形就是選擇性偏差(Selection bias)。



HCR=any combination of hormone therapy and/or chemotherapy

※選擇性偏差的類型

由於造成低估或高估之偏差會因為內容不同而有差異，所以會有不同之名稱，在應用上至少有 35 種，若有興趣，請自行參考 Grimes DA & Schulz KF: Bias and casual association in observational research. Lancet 2002; 359:248-252.等人文章。

(1)「歸屬偏差」(Membership-bias)

(2)「柏克森偏差」(Berkson bias)又名「住院率偏差」(admission-rate bias)

(3)「奈曼偏差」(Neyman bias)又名「發生與盛行率偏差」(incidence-prevalence bias)

(4)「偵測偏差」(detection bias)又名「不反應偏差」(non-respondent bias)

(二) 錯誤分組 (Misclassification)

1. 錯誤分組概念

錯誤分組可分為兩種情形，無方向性(Non-differential misclassification)與有方向性偏差(Differential misclassification) (Y 為實際應變項、X 為實際研究自變項、Z 為觀察變項(可能有錯誤分組))，錯誤分組可能發生在 X，也可能在 Y。

(1) 無方向性偏差(Non-differential misclassification)

敏感度(Sensitivity) =Pr(Z=1|X=1)　註記為 SE

特異度(Specificity) =Pr(Z=0|X=0)　註記為 SP

(2) 有方向性偏差(Differential misclassification)

敏感度(Sensitivity) =Pr(Z=1|X=1, Y=y)　註記為 SE_y

特異度(Specificity) =Pr(Z=0|X=0, Y=y)　註記為 SP_y

2. 流行病學實例

(1) 世代追蹤研究之疾病錯誤分組—以無方向性偏差(Non-differential misclassification)為例

		暴露(Exposure)(X)	
		有(1)	無(0)
疾病(Disease) (Y)	有(1)	400	200
	無(0)	600	800
		1000	1000

得到相對風險比值 RR=(400/1000)/(200/1000)=2

假設透過校正性研究，我們可以得到實際暴露(Exposure)資料分布如下：

(A) 暴露(X=1)

		實際疾病分組(Disease)(Y)		
		有(1)	無(0)	
觀察疾病分組 (Classified status) (Z)	有(1)	320	60	380
	無(0)	80	540	620
		400	600	1000

(B) 未暴露(X=0)

		實際疾病分組(疾病 (Disease)(Y)		
		有(1)	無(0)	
觀察疾病分組 (Classified status) (Z)	有(1)	160	80	240
	無(0)	40	720	760
		200	800	1000

敏感度(SE_1)=Pr(Z=1|Y=1, X=1)=320/400=0.8

特異度(SP_1)=Pr(Z=0|Y=0, X=1)=540/600=0.9

敏感度(SE_0)=Pr(Z=1|Y=1, X=0)=160/200=0.8

特異度(SP_0)=Pr(Z=0|Y=0, X=0)=720/800=0.9

(C) 疾病(Disease)與暴露(Exposure)資料分布

		暴露(Exposure)(X)	
		有(1)	無(0)
觀察疾病分組 (Classified status) (Z)	有(1)	380	240
	無(0)	620	760
		1000	1000

若利用觀察資料計算相對風險比值 RR=(380/1000)/(240/1000)=1.58，但真實的相對疾病風險比值為 2；因此若使用觀察疾病分組求相對疾病風險比值會有低估(underestimation, toward the null) 之情形。

(2) 病例對照研究之暴露錯誤分組—以有方向性偏差(Differential misclassification)為例

		暴露(Exposure)(X)		
		有(1)	無(0)	
疾病(Disease)	有(1)	600	400	1000
(Y)	無(0)	300	700	1000

得到勝算比值 OR=(600x700)/(300x400)=3.5

假設透過校正性研究，我們可以得到疾病(Disease)資料分布如下：

(A) 得病(Y=1)

		真實暴露(Exposure)(X)		
		有(1)	無(0)	
觀察暴露分組	有(1)	540	120	660
(Classified status)(Z)	無(0)	60	280	340
		600	400	1000

(B) 無病(Y=0)

		真實暴露(Exposure)(X)		
		有(1)	無(0)	
觀察暴露分組	有(1)	180	70	250
(Classified status)(Z)	無(0)	120	630	750
		300	700	1000

敏感度(SE_1)=Pr(Z=1|X=1, Y=1)=540/600=0.9

特異度(SP_1)=Pr(Z=0|X=0, Y=1)=280/400=0.7

敏感度(SE_0)=Pr(Z=1|X=1, Y=0)=180/300=0.6

特異度(SP_0)=Pr(Z=0|X=0, Y=0)=630/700=0.9

(C) 疾病(Disease)與暴露(Exposure)資料分布

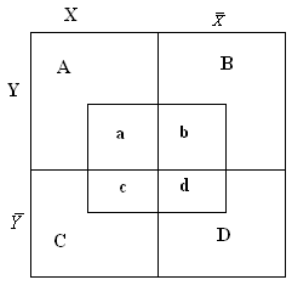
		觀察暴露分組(Classified status)(Z)	
		有(1)	無(0)
疾病(Disease)	有(1)	660	340
(Y)	無(0)	250	750
		1000	1000

觀察資料得到的勝算比值 OR=(660x750)/(340x250)=5.82，但實際的勝算比值為 3.5；因此使用觀察疾病分組求相對疾病風險比值會有高估 (overestimation, away from the null) 之情形.。

(3) 手機與腦部腫瘤錯誤分組之解釋：

在手機之實例中，頻繁地使用手機的定義也值得商確，如果高度頻繁地使用手機確實會造成腦部發生病變，若依照研究中的定義將非高度頻繁使用的民眾也當成暴露組，將會稀釋真正使用手機所造成的效果而影響評估結果。又假若參與研究的人的手機使用經驗是在數十年前，則當時的訊號可能為類比式而非數位式訊號，因此時兩者的手機暴露是不相同的，會造成的效果可能也不同，但這在研究中都被歸類在有手機使用那一組而無法區分。

使用 2X2 列聯表呈現錯誤分組(misclassification)可能造成的效果，若利用 a,b,c,d 占 A,B,C,D 的比例來呈現各組可能產生錯誤分組的機率，



在勝算比的計算中(OR=ad/bc)，錯誤分組可能造成高估(d 增加或者 a 增加) 也可能造成低估 (b 增加或者 c 增加)，端看錯誤分類的方向為何。此時可以利用兩階段抽樣研究設計 (two stage sampling design)，來推估錯誤分組的比率估計值。

(三) 運用實證醫學方法獲得正確的推論 (參考 e-book 01_ p25, p48-52)

為了使因果關係不受上述偏差的影響，西方醫學將解決這些偏差的研究設計稱做為"實證醫學 (Evidence-Based Medicine, EBM)"。

為達到實證醫學而進行的研究設計有以下四個大類，其效度強度依序減弱：

(1) 隨機分派對照試驗(randomized trial, RCT): 移除干擾因子及偏差(bias)

(2) 前瞻性研究(prospective study)-因在果之前，無進行隨機分派: 控制干擾因子及偏差。

(3) 病例對照研究(case-control study)-因為事件很少，需要追蹤的人數眾多故進行此研究: 控制干擾因子及偏差。

(4) 橫斷性研究(cross-sectional study)-因果都在同一時間點做測量，通常作為描述性流行病學之調查。

在推論因果關係時較不適用，通常作為描述性流行病學之調查。糖尿病與三酸甘油脂的因果，同一時間點測量兩個值都高，無法判別因果關係，但有時也有適用的例子。以色盲作為結果，去辨別與其他因子的關係，因為色盲不會因為時間改變而改變，故在因果關係的推論上有一致的結果， 但仍須考慮種種干擾因素以及偏差的影響。

以上四種研究若此時序性(判定因果關係最重要之性質)，前三種有時序性，因必定在果之前，第一種(隨機分派對照試驗，RCT)在因推果的過程中進行隨機分派，第二種(前瞻性研究，prospective study)為先收集因，再透過追蹤觀察得到果的資訊，第三種(病例對照研究，case-control study)先收集果(發病與否)的資訊再往前收集因(是否暴露於某些因子)的資訊；而第四種(橫斷性研究，cross-sectional study)無時序性，因果時間為同時。

流行病學方法論及實驗
(The Methods of Epidemiology and Practices)

偏好選擇流行病學故事之護理觀

授課教師：陳秀熙 教授/許辰陽 博士

台灣大學流行病學與預防醫學研究所

1

雌激素補充療法是否可以減少停經後女性心血管疾患 (CHD)的發生?

In Observational Studies：
A group: Take postmenopasual estrogen
B group: Not to take postmenopasual estrogen
CHD in A ＜＜ CHD in B
→ Estrogen use is effective in secondary prevention of CHD

- **Argument**：

Women who choose to take hormones are healthier and have a more favourable CHD profile than those who do not.

※ Selection bias: Absence of comparability between groups being studied

2

Argument

1) Participant and non-participants selection: Participants are frequently educated, healthier, lead better lifestyle, and have few complication
2) A hospital-based case-control study on myocardial infarction will underestimate relative risk (Neyman bias).
3) Knowledge of the exposure may lead to an increased rate of admission to hospital (ex. OC and thrombembolism)
4) Detection bias: An exposure may facilitate the unmasking of disease (ex. estrogen and endometrial cancer)

3

IARC Report to the Union for International Cancer Control (UICC) on the Interphone Study

International Agency for Research on Cancer
World Health Organization

Conclusions

Glioma and meningioma
Overall, no increase in risk of glioma or meningioma was observed with use of mobile phones. There were suggestions of an increased risk of glioma at the highest exposure levels, but biases and error prevent a causal interpretation. The possible effects of long-term heavy use of mobile phones require further investigation.

Acoustic neuroma
There was no increase in risk of acoustic neuroma with ever regular use of a mobile phone or for users who began regular use 10 years or more before the reference date. Elevated odds ratios observed at the highest level of cumulative call time could be due to chance, reporting bias or a causal effect. As acoustic neuroma is usually a slowly growing tumour, the interval between introduction of mobile phones and occurrence of the tumour might have been too short to observe an effect, if there is one.

The Interphone study
Mobile madness

The Economist

Please press "recall"

One problem was what statisticians call selection bias. Interphone began by gathering a group of people who had had the cancers of interest (glioma, meningioma, acoustic neurinoma and parotid gland tumour) and questioning them about their past use of mobile phones. The researchers then approached a number of healthy people in order to compare them with the cancer patients, and find out if there was a systematic difference in mobile-phone use between the two groups. Some of those approached agreed, and some declined. Of those who agreed to take part, 59% were regular mobile-phone users as defined by the study's protocol. Later on, those who had declined were recontacted and asked about their mobile use. Among this group, only 34% were regular users. That meant those in the control group were more likely than average to be regular users, and therefore were not representative of the population at large.

Moreover, the definition of "regular mobile-phone use" was itself questionable. Anyone who had used a phone just once a week for at least six months qualified. That is a pretty low rate of usage. If phones really do cause cancer, but only at high exposure, employing such a generous definition of regular use means that the effect might be diluted into undetectability.

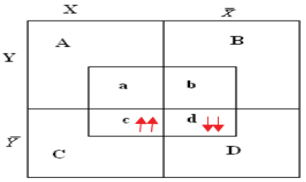
Another potentially serious flaw is that participants asked in 2001-02 about their mobile use a decade earlier will have been using analogue, not digital, handsets. That would lead to a different pattern of exposure and therefore of potential risk.

使用手機是否對於健康有影響

(1)

- Enrolled subjects: 59% were regular mobile-phone users
- Among refused group: 34% were regular users
  - **control group were more likely than average to be regular users**
- What happen to the result for selection bias ?

Underestimate



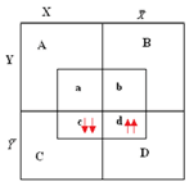
6

Another scenario

(2)

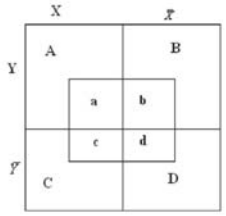
- Enrolled subjects: 34% were regular mobile-phone users
- Among refused group: 59% were regular users
  - **control group were less likely than average to be regular users**
- What happen to the result for selection bias ?

Overestimate



$$OR = \frac{A/C}{B/D}$$

$$\hat{OR} = \frac{a/c}{b/d}$$



α =a/A
β =b/B
γ=c/C
δ=d/D

Target population

Source population

X: exposure Y: disease status

•For cohort study:

$$P(Y \mid X) / P(Y \mid \bar{X}) = RR$$

$$\frac{P(Y \mid X) / P(\bar{Y} \mid X)}{P(Y \mid \bar{X}) / P(\bar{Y} \mid \bar{X})} = OR$$

Sampling fraction

$$P(S=1 \mid X, Y) = \pi_{XY}$$
$$P(S=1 \mid X, \bar{Y}) = \pi_{X\bar{Y}}$$
$$P(S=1 \mid \bar{X}, Y) = \pi_{\bar{X}Y}$$
$$P(S=1 \mid \bar{X}, \bar{Y}) = \pi_{\bar{X}\bar{Y}}$$

施打疫苗是否與兒童發展異常有關

- Background
  - One of content in MMR vaccination may results in the occurrence of autism or other developmental disorder.
- Study Goal
  - the relation between MMR vaccination (X) and developmental disorder (Y)

	X	$\bar{X}$	total
Y	1,000	300	1,300
$\bar{Y}$	160,000	45,000	205,000
total	161,000	45,300	206,300

$$RR = \frac{1000/161000}{300/45300} = \frac{0.62\%}{0.66\%} = 0.93$$

$$OR = \frac{1000/160000}{300/45000} = 0.94\ (0.9375)$$

	X	$\bar{X}$	total
Y	1,000	300	1,300
$\bar{Y}$	160,000	45,000	205,000
total	161,000	45,300	206,300

π_X=0.2 $\pi_{\bar{X}}$=0.5

	X	$\bar{X}$	total
Y	200	150	350
$\bar{Y}$	32,000	22,500	54,500
total	32,200	22,650	54,850

$$RR = \frac{1000/161000}{300/45300} = \frac{0.62\%}{0.66\%} = 0.93$$

$$OR = \frac{1000/160000}{300/45000} = 0.94\ (0.9375)$$

$$RR = \frac{200/32200}{150/22650} = \frac{0.62\%}{0.66\%} = 0.93$$

$$OR = \frac{200/32000}{150/22500} = 0.94\ (0.9375)$$

If the study includes a higher proportion of subjects with vaccination history and have developmental disorder :

	X	$\bar{X}$	total
Y	1,000	300	1,300
$\bar{Y}$	160,000	45,000	205,000
total	161,000	45,300	206,300

π_{XY} = 0.9

	X	$\bar{X}$	total
Y	900	60	960
$\bar{Y}$	64,000	9,000	73,000
total	64,900	9,060	73,960

$$\pi_{XY} = 0.9,\ \pi_{X\bar{Y}} = 0.4$$
$$\pi_{\bar{X}Y} = \pi_{\bar{X}\bar{Y}} = 0.2$$

$$RR = \frac{1000/161000}{300/45300} = \frac{0.62\%}{0.66\%} = 0.93$$

$$OR = \frac{1000/160000}{300/45000} = 0.94\ (0.9375)$$

overestimate

$$RR = \frac{900/64900}{60/9060} = \frac{1.39\%}{0.66\%} = 2.09$$

$$OR = \frac{900/64000}{60/9000} = 2.11$$

Case-control study

	X	$\bar{X}$	total
Y	1,000	300	1,300
$\bar{Y}$	160,000	45,000	205,000
total	161,000	45,300	206,300

$\pi_{XY} = \pi_{\bar{X}Y} = 1$

	X	$\bar{X}$	total
Y	1,000	300	1,300
$\bar{Y}$	3,550	950	4,500
total	4550	1250	5800

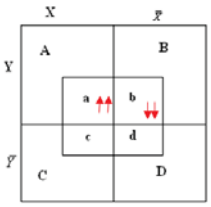
$\pi_{X\bar{Y}} = \pi_{\bar{X}\bar{Y}} = 0.02$

$$OR = \frac{1000/3550}{300/950} = 0.89$$

回憶性偏差(Recall bias)

- Retrospective study

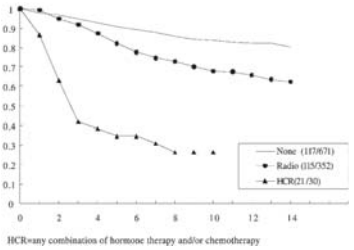
"Someone subsequently diagnosed with a brain tumour might easily be biased…
to exaggerate the former…"

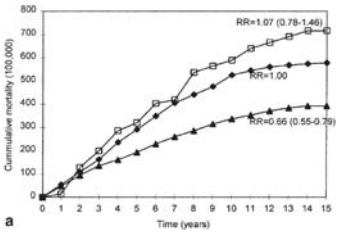


- Does recall-bias lead to under-estimation or over-estimation of association between the use of mobile phone and brain tumour?

→they may lead to over-estimation

選擇性存活分析(Selected survival)





※選擇性偏差的類型

- "Comparability" problem
- Bias due to selection
  - Types of selection bias
    - Membership-bias
    - Berkson bias (admission-rate bias)
    - Neyman bias (incidence-prevalence bias)
    - detection bias (non-respondent bias)

Self-selection and bias in a large prospective pregnancy cohort in Norway

Paediatric and Perinatal Epidemiology, **23**, 597–608.

Summary

Nilsen RM, Vollset SE, Gjessing HK, Skjærven R, Melve KK, Schreuder P, Alsaker ER, Haug K, Daltveit AK and Magnus P. Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatric and Perinatal Epidemiology* 2009; **23:** 597–608.

Self-selection in epidemiological studies may introduce selection bias and influence the validity of study results. To evaluate potential bias due to self-selection in a large prospective pregnancy cohort in Norway, the authors studied differences in prevalence estimates and association measures between study participants and all women giving birth in Norway. Women who agreed to participate in the Norwegian Mother and Child Cohort Study (43.5% of invited; $n = 73\,579$) were compared with all women giving birth in Norway ($n = 398\,849$) using data from the population-based Medical Birth Registry of Norway in 2000–2006. Bias in the prevalence of 23 exposure and outcome variables was measured as the ratio of relative frequencies, whereas bias in exposure-outcome associations of eight relationships was measured as the ratio of odds ratios.

Statistically significant relative differences in prevalence estimates between the cohort participants and the total population were found for all variables, except for maternal epilepsy, chronic hypertension and pre-eclampsia. There was a strong under-representation of the youngest women (<25 years), those living alone, mothers with more than two previous births and with previous stillbirths (relative deviation 30–45%). In addition, smokers, women with stillbirths and neonatal death were markedly under-represented in the cohort (relative deviation 22–43%), while multivitamin and folic acid supplement users were over-represented (relative deviation 31–43%). Despite this, no statistically relative differences in association measures were found between participants and the total population regarding the eight exposure-outcome associations.

Using data from the Medical Birth Registry of Norway, this study suggests that prevalence estimates of exposures and outcomes, but not estimates of exposure-outcome associations are biased due to self-selection in the Norwegian Mother and Child Cohort Study.

Background characteristics	Total population No. (%)	MoBa participants No. (%)	Ratio of relative frequencies [95% CI][a]
Maternal age at delivery[b]			
<25 years	67 707 (17.0)	8734 (11.9)	0.70 [0.69, 0.71]
25–34 years	266 171 (66.7)	52 638 (71.5)	1.07 [1.07, 1.08]
>34 years	64 961 (16.3)	12 207 (16.6)	1.02 [1.00, 1.03]
Marital status[c]			
Single	25 547 (6.4)	2599 (3.5)	0.55 [0.53, 0.57]
Cohabiting	172 287 (43.2)	33 930 (46.1)	1.07 [1.06, 1.07]
Married	195 944 (49.1)	36 732 (49.9)	1.02 [1.01, 1.02]
Parity			
0	162 983 (40.9)	31 763 (43.2)	1.06 [1.05, 1.06]
1	142, 211 (35.7)	26 486 (36.0)	1.01 [1.00, 1.02]
2	65 770 (16.5)	11 840 (16.1)	0.98 [0.96, 0.99]
>2	27 885 (7.0)	3490 (4.7)	0.68 [0.66, 0.70]
Previous stillbirths[d]			
0	232 408 (98.5)	41 402 (99.0)	1.00 [1.00, 1.01]
1	3220 (1.37)	387 (0.93)	0.68 [0.62, 0.74]
>1	238 (0.10)	27 (0.06)	0.64 [0.42, 0.86]
Maternal asthma			
Yes	16 468 (4.1)	3155 (4.3)	1.04 [1.01, 1.07]
Maternal epilepsy			
Yes	3066 (0.77)	554 (0.75)	0.98 [0.91, 1.06]
Pregestational diabetes			
Yes	2701 (0.68)	415 (0.56)	0.83 [0.76, 0.91]
Chronic hypertension			
Yes	2232 (0.56)	376 (0.51)	0.91 [0.83, 1.00]
All deliveries	398 849 (100)	73 579 (100)	

Exposures and pregnancy complications	Total population No. (%)	MoBa participants No. (%)	Ratio of relative frequencies [95% CI][a]
In vitro fertilization			
Yes	7353 (1.8)	1542 (2.1)	1.14 [1.09, 1.19]
Smoking[b]			
Unknown	60 254 (15.1)	9512 (12.9)	0.86 [0.84, 0.87]
No	295 618 (74.1)	59 548 (80.9)	1.09 [1.09, 1.10]
Yes	42 977 (10.8)	4519 (6.1)	0.57 [0.55, 0.58]
Multivitamin use[c]			
Unknown	66 443 (16.7)	10 310 (14.0)	0.84 [0.83, 0.85]
No	225 628 (56.6)	37 391 (50.8)	0.90 [0.89, 0.90]
Yes	106 778 (26.8)	25 878 (35.2)	1.31 [1.30, 1.33]
Folic acid use[c]			
Unknown	66 443 (16.7)	10 310 (14.0)	0.84 [0.83, 0.84]
No	185 974 (46.6)	24 551 (33.4)	0.72 [0.71, 0.72]
Yes	146 432 (36.7)	38 718 (52.6)	1.43 [1.42, 1.44]
Medication use[d]			
Yes	85 300 (21.4)	17 531 (23.8)	1.11 [1.10, 1.13]
Gestational diabetes			
Yes	3444 (0.86)	587 (0.80)	0.92 [0.86, 0.99]
Pre-eclampsia			
Yes	15 879 (4.0)	2861 (3.9)	0.98 [0.94, 1.01]
Placental abruption[e]			
Yes	1726 (0.43)	282 (0.38)	0.88 [0.79, 0.98]
All deliveries	398 849 (100)	73 579 (100)	

Exposure-outcome associations	Total population (n = 391 011) UOR	Total population (n = 391 011) AOR [95% CI][b]	MoBa participants (n = 72 159) UOR	MoBa participants (n = 72 159) AOR [95% CI][b]	Ratio of AORs [95% CI][c]
Smoking and low birthweight(<2500 g)					
Smoker[d]	1.89	1.85 [1.76, 1.94]	1.88	1.77 [1.53, 2.05]	0.95 [0.83, 1.10]
Smoking and placental abruption					
Smoker[d]	1.72	1.73 [1.50, 1.98]	1.95	1.85 [1.24, 2.74]	1.07 [0.70, 1.50]
Smoking and stillbirth					
Smoker[d]	1.32	1.19 [1.06, 1.34]	1.45	1.31 [0.87, 1.97]	1.08 [0.69, 1.55]
Chronic hypertension and gestational diabetes					
Hypertension	2.55	2.31 [1.72, 3.09]	2.81	2.59 [1.28, 5.25]	1.12 [0.45, 1.95]
Birthweight and neonatal death[e]					
Birthweight <2500 g	30.2	30.2 [25.9, 35.3]	37.7	43.1 [28.7, 64.8]	1.43 [0.98, 2.07]
Vitamin use and placental abruption					
User of vitamins[f]	0.72	0.74 [0.66, 0.83]	0.72	0.75 [0.57, 0.99]	1.01 [0.79, 1.31]
Parity and pre-eclampsia					
Parity 0	1.00	1.00 Reference	1.00	1.00 Reference	1.00 Reference
Parity 1	0.46	0.47 [0.45, 0.49]	0.45	0.46 [0.42, 0.50]	0.98 [0.90, 1.07]
Parity 2	0.43	0.43 [0.40, 0.45]	0.44	0.44 [0.38, 0.50]	1.03 [0.90, 1.15]
Parity >2	0.44	0.43 [0.39, 0.46]	0.47	0.46 [0.37, 0.58]	1.07 [0.87, 1.32]
Marital status and preterm birth (<37 weeks)					
Single	1.00 Reference	1.00 Reference	1.00 Reference	1.00 Reference	1.00 Reference
Cohabiting	0.71	0.77 [0.73, 0.81]	0.64	0.70 [0.60, 0.81]	0.90 [0.79, 1.04]
Married	0.67	0.75 [0.72, 0.79]	0.58	0.67 [0.57, 0.78]	0.89 [0.78, 1.03]

Misclassification

Size Effect due to Misclassification

Table I. Data from case-control study of sudden infant death syndrome (SIDS) and reported maternal antibiotic use.*

	Self-reported use of antibiotics (Z)	
Case-control status (Y)	1	0
1	122	442
0	101	479

*Reference: Drews *et al.* [16].

To illustrate sensitivity analysis for outcome misclassification, let us now pretend that sampling for the SIDS study had been done in a prospective or cross-sectional manner and that the outcome (Y) rather than the exposure (X) was error-prone. We then view the data in Table I as depicting Y^* vs. X. Numerically maximizing the likelihood based on equation (8) assuming non-differential error with SE=0.6 and SP=0.9 yields $\hat{\beta}=\ln(\hat{OR})=0.982$ (0.727), so that $\hat{OR}=2.67$. Assuming instead a differential scenario with $SE_1=0.6$, $SP_1=0.8$, $SE_0=0.7$, and $SP_0=0.9$ gives $\hat{\beta}=\ln(\hat{OR})=1.33$ (0.735), so that $\hat{OR}=3.78$. In both of these cases, the naïve OR estimate of 1.31 would be markedly attenuated.

21

錯誤分組 (Misclassification)

- 無方向性(Non-differential error)
  - 敏感度(Sensitivity) =Pr(Z=1|X=1) 註記為SE
  - 特異度(Specificity) =Pr(Z=0|X=0) 註記為SP
- 有方向性偏差(Differential error)
  - 敏感度(Sensitivity) =Pr(Z=1|X=1, Y=y) 註記為SE_y
  - 特異度(Specificity) =Pr(Z=0|X=0, Y=y) 註記為SP_y

Y：實際應變項
X：實際自變項
Z：觀察變項
(錯誤分組可能發生在X，也可能發生在Y)

22

流行病學範例
世代追蹤研究之疾病錯誤分組

- 以無方向性偏差(Non-differential error)為例

		暴露(Exposure)(X)	
		有(1)	無(0)
疾病(Disease) (Y)	有(1)	400	200
	無(0)	600	800
		1000	1000

相對風險比值RR=(400/1000)/(200/1000)=2

23

(A) 實際有暴露(X=1)

		疾病(Disease)(Y)		
		有(1)	無(0)	
觀察疾病分組 (Classified status) (Z)	有(1)	320	60	380
	無(0)	80	540	620
		400	600	1000

(B) 實際未暴露(X=0)

		疾病(Disease)(Y)		
		有(1)	無(0)	
觀察疾病分組 (Classified status) (Z)	有(1)	160	80	240
	無(0)	40	720	760
		200	800	1000

無方向性偏差(Non-differential error)
敏感度(SE_1)=Pr(Z=1|Y=1, X=1)=320/400=0.8
特異度(SP_1)=Pr(Z=0|Y=0, X=1)=540/600=0.9
敏感度(SE_0)=Pr(Z=1|Y=1, X=0)=160/200=0.8
特異度(SP_0)=Pr(Z=0|Y=0, X=0)=720/800=0.9

陽性預測值 (PPV)
$$=\frac{\text{盛行率}\times\text{敏感度}}{\text{盛行率}\times\text{敏感度}+(1-\text{盛行率})\times(1-\text{特異度})}$$

陰性預測值 (NPV)
$$=\frac{(1-\text{盛行率})\times\text{特異度}}{\text{盛行率}\times(1-\text{敏感度})+(1-\text{盛行率})\times(\text{特異度})}$$

(C)調整錯誤分組下，疾病(Disease)與暴露(Exposure)資料分布

		暴露(Exposure)(X)	
		有(1)	無(0)
觀察疾病分組 (Classified status) (Z)	有(1)	380	240
	無(0)	620	760
		1000	1000

•相對風險比值RR=(380/1000)/(240/1000)=1.58
•調整錯誤分組後，相對疾病風險比值從2降低為1.58
•若使用觀察疾病分組求相對疾病風險比值會有低估(underestimation, toward the null) 之情形

流行病學範例

病例對照研究之暴露錯誤分組

– 以有方向性偏差(Differential error)為例

		暴露 (Exposure)(X)		
		有(1)	無(0)	
疾病 (Disease) (Y)	有(1)	600	400	1000
	無(0)	300	700	1000

得到勝算比值OR=(600x700)/(300x400)=3.5

假設透過校正性研究，得到實際疾病(Disease)資料分布如下

(A) 實際得病(Y=1)

		真實暴露(Exposure)(X)		
		有(1)	無(0)	
觀察暴露分組 (Classified status)(Z)	有(1)	540	120	660
	無(0)	60	280	340
		600	400	1000

(B) 實際無病(Y=0)

		真實暴露(Exposure)(X)		
		有(1)	無(0)	
觀察暴露分組 (Classified status)(Z)	有(1)	180	70	250
	無(0)	120	630	750
		300	700	1000

有方向性偏差(Differential error)
敏感度(SE_1)=Pr(Z=1|X=1, Y=1)=540/600=0.9
特異度(SP_1)=Pr(Z=0|X=0, Y=1)=280/400=0.7
敏感度(SE_0)=Pr(Z=1|X=1, Y=0)=180/300=0.6
特異度(SP_0)=Pr(Z=0|X=0, Y=0)=630/700=0.9

陽性預測值 (PPV) $= \frac{\text{盛行率} \times \text{敏感度}}{\text{盛行率} \times \text{敏感度} + (1-\text{盛行率}) \times (1-\text{特異度})}$

陰性預測值 (NPV) $= \frac{(1-\text{盛行率}) \times \text{特異度}}{\text{盛行率} \times (1-\text{敏感度}) + (1-\text{盛行率}) \times (\text{特異度})}$

(C)調整錯誤分組下，疾病(Disease)與暴露(Exposure)資料分布

		觀察暴露分組 (Classified status)(Z)	
		有(1)	無(0)
疾病(Disease) (Y)	有(1)	660	340
	無(0)	250	750
		1000	1000

•勝算比值OR=(660x750)/(340x250)=5.82
•在調整錯誤分組後，勝算比值從3.5上升為5.82
•若顯示使用觀察疾病分組求相對疾病風險比值有高估(overestimation, away from the null)

運用實證醫學方法獲得正確的推論

- To study X (cause) → Y (Consequence), we had better have good study design (such as randomized controlled trial, RCT) to demonstrate a specific causal relationship

To achieve EBM, we have the following design.

1. Randomized trial-remove confounding and bias
2. Prospective study-control confounding and bias
3. Case-control study-control confounding and bias
4. Cross-sectional study

六、 護理人如何校正偏好選擇之實例

1. Qin et al.: Estimating effects of nursing intervention via propensity score analysis. Nurs Res. 2008; 57(6): 444-452.

2. Stukel et al.: Analysis of observational studies in the presence of treatment selection bias: effects of invasive cardiac management on AMI survival using propensity score and instrumental variable methods. JAMA 2007; 297(3):278-285.

一、利用 Propensity score 校正偏好選擇並正確評估介入之效益

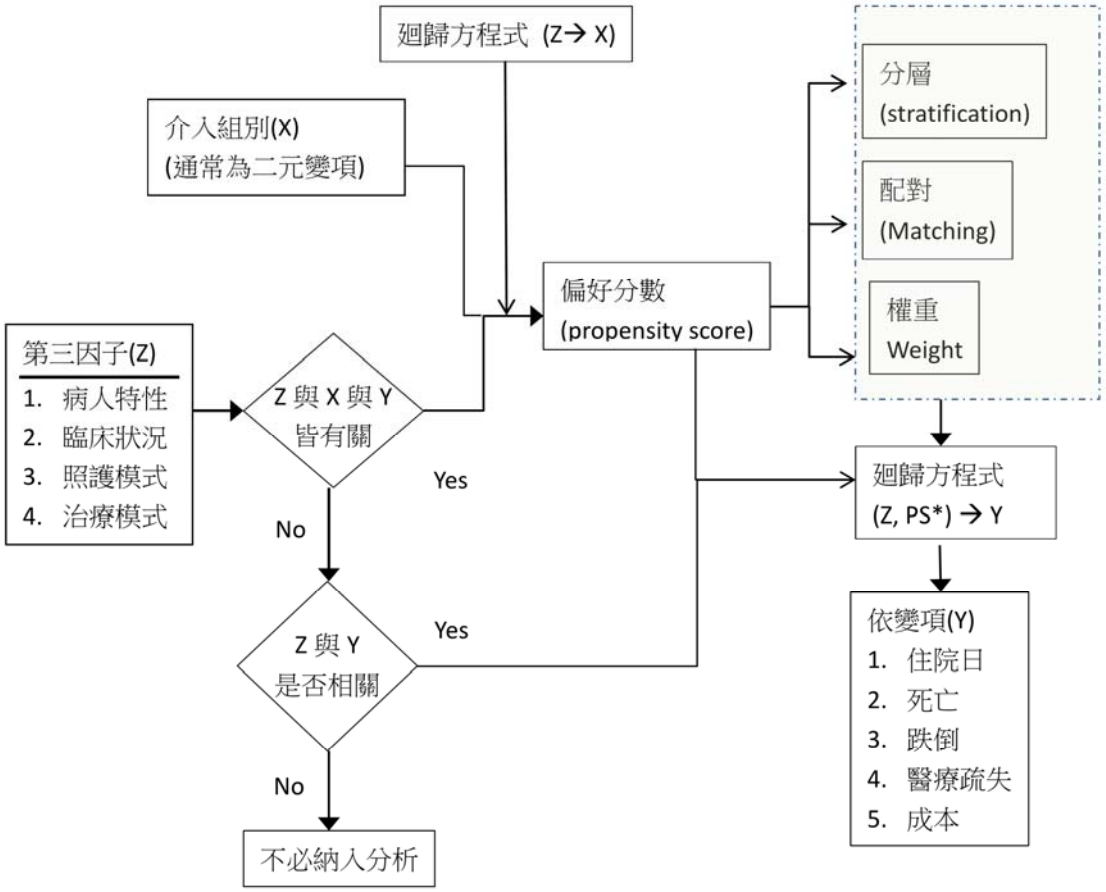
一般而言，評估效益研究若無法以隨機分派試驗去除比較時可能發生的種種偏誤與干擾，則可以利用前述的分層分析方法以及迴歸方法將可能的干擾因子所造成的影響納入考慮。另外可以利用 propensity score 處理干擾因子對於效益評估所帶來的影響。

我們可以利用研究參與者影響接受治療的機率所得到的偏好分數，建立接近若以隨機分派研究設計檢視治療及介入所得到的效益之不偏估計。亦即，若有一個接受治療參與者與另一個未接受治療者有相近的偏好分數，如此就如同以隨機分派之機轉將一位參與者分派到治療組，而另一位分派到對照組一般。因此 propensity score 就成為一個分析非隨機分派試驗的觀察性研究分析效益不偏估計的利器。

以下先介紹分析架構，並以護理中疼痛管理與住院日相關為例。

I 偏好分數 (propensity score) 分析模式架構:以護理效益研究(effectiveness research) 為例 (Qin et al. 2008)

圖一為該研究利用偏好分數之架構。在以下的研究中針對 523 位接受 hip procedure 之病患(其中 41 位住院 2 次，兩位住院 3 次)，總共 568 個住院次數。在 568 次住院中，有 214 次接受疼痛管理，利用偏好分數分析 (propensity score analysis) 回答 ”疼痛照護是否會影響住院日?”



圖一、偏好分數分析架構圖 (*PS: propensity score)

1. 選擇相關第三因子(extraneous variable)

在疼痛管理之例子中，可能的影響因子包括病人特徵、臨床狀況、照護內容、治療以及護理介入等。在偏好分數分析中，若納入太多因子則會造成統計檢定力不足之問題。

2. Propensity score 分析

(1) 產生 propensity score

(2) 利用得到之 propensity 校正偏誤

一般在求偏好分數若介入組別是二元變項，可以利用 logistic regression model，有關 Z 與 X 是否有相關可以利用不同的統計方法及模式。在此例中，前項是利用 GEE ，因為有些是同樣病人會有相關之問題，再利用 Spearman's 相關分析 Z 與 X 與 Y 之相關得到 9 個干擾因子(Z)和疼痛管理(X)與住院天數(Y)皆有關。13 個和 Y 有相關但和 X 呈弱相關及 5 個和疼痛管理(X)有強相關但和住院天數(LOS,Y)呈弱相關。

3. 如何評估所產生的 propensity score 是否可有效校正偏誤

(1) 以接受者作業曲線評估 propensity score 是否可有效區別介入組與未介入組

(2) 以分佈圖或盒鬚圖(box plot)檢視介入組與未介入組之 propensity score 是否有重

疊

(3) 以 propensity score 將資料區分為 5 組並比較介入組與未介入組之基礎分佈 (baseline distribution)

4. 如何利用 propensity score 校正偏誤

(1) 多變項分析 (multivariable analysis)

將偏好分數及其他影響住院日(Y)之變項(Z)放入模式中當成控制變項。

結果見原表二。將疼痛管理(X)放入模式中評估在控制偏好分數及其他變項(Z')之後，X 是否呈現有統計顯著意義。在此例中，偏好分數並沒有呈現統計上之相關(p=0.152)，而 X 呈現有意義之正相關(p=0.002)。

(2) 分層分析 (stratified analysis)

結果見原表三。將住院日(LOS, Y)依偏好分數分成 5 群。由偏好分數最低至最高再進行迴歸分析，發現只有在最低層(p=0.001)及最高層(p=0.031)中，疼痛管理與住院天數(Y)具有統計相關。

(3) 配對分析 (matching)

利用標準化後 logit 值 0.2 內進行配對分析，故只留下 308 對資料進行迴歸分析。發現疼痛管理與住院日在統計上仍呈正相關(p=0.006)。

(4) 以 propensity score 對觀察資料作加權 (weight) (見 McNiel and Binder, 2007)

5. Instrumental variable (IV)

(1) 定義: 與選擇組別(X)有關但與結果(Y)無關之變項。IV 在意義上較接近中介變項 (intermediate endpoint)。

(2) 目的: 藉由加入 IV 調整未知的干擾因子使得分析結果能夠與隨機分派試驗所得到的結果相近。

(3)假設:

(a) 影響到組別分派但與結果無關

(b) 與組別分派之相關性越強則越能達到 IV 之目的。若利用 chi-square 值來評估，在模式中 IV 除以所增加的自由度應該至少達到 20(即利用 likelihood ratio test 衡量 IV 的影響)。 若在線性迴歸中中以 F 值評估，加入 IV 至少要達到 10 的差異。

(c) IV 必須與模式中的其他變項以及結果無關

(d) IV 與介入程度需有單調(monotonic)之關係

七、 貝氏柯南護理流行病學

I 臨床推理—檢驗及診斷工具的正確性(Accuracy)

(參考文獻：

1. Berg WA, et al.: Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. JAMA 2008;299(18):2151-63.
2. Braden and Bergstrom: Predictive validity of the Braden scale for pressure sore risk in a nursing home population)

在臨床決策中，檢驗工具的正確性對於治療的決策扮演著重要的角色，在個案實際上有病的情況之下，若工具無法正確檢驗出來，則個案可能因為診斷被忽略而疏於治療，使得疾病持續進展，而錯失治癒的時機；反之而言，若個案實際上未罹患疾病，卻被誤診為有病，則可能導致過度診斷(Over-diagnosis)或過度治療(Over-treated)的情況產生，進而影響個案之生活品質。因此，具有正確性高的診斷工具在醫學臨床及護理照護上是非常重要的。本次課程重點之一就是從流行病學及統計學的觀點來看檢驗工具正確性的測量。

1. 敏感度(Sensitivity)、特異度(Specificity)、陽性預測值(Positive Predictive Value, PPV)與陰性預測值(Negative Predictive Value, NPV) (e-Book 01: Chap 12, p988-1000; e-Book 03: Chap 12, p124-129, Chap 13, p139-143)

臨床實例 1. 乳癌篩檢工具比較

(1) 2X2 Table

		乳癌 罹病	乳癌 未罹病	
乳房攝影術 + 乳房超音波	陽性	a (31)	b (275)	a+b (306)
	陰性	c (9)	d (2322)	c+d (2331)
		a+c (40)	b+d (2597)	

利用不同的條件機率，我們可以將診斷特性中的敏感度，特異度，陽性預測值以及因性預測值作如下的定義:

敏感度： $Sen\ (\text{Sensitivity}) = P(X=1 \mid Y=1) = \frac{a}{a+c} = \frac{31}{31+9} = 77.5\%$

特異度： $Spe\ (\text{Specificity}) = P(X=0 \mid Y=0) = \frac{d}{b+d} = \frac{2322}{275+2322} = 89.4\%$

陽性預測值：

$PPV(\text{PositivePredictiveValue}) = P(Y=1 \mid X=1) = \frac{a}{a+b} = \frac{31}{31+275} = 10.1\%$

陰性預測值：

$NPV(Negative\ \text{Predictive Value}) = P(Y=0 \mid X=0) = \frac{d}{c+d} = \frac{2322}{2322+9} = 99.6\%$

偽陰性： $FN(False\ negative) = P(X=0 \mid Y=1) = \frac{c}{a+c} = \frac{9}{31+9} = 22.5\%$

偽陽性： $FP(False\ posotive) = P(X=1 \mid Y=0) = \frac{b}{b+d} = \frac{275}{275+2322} = 10.9\%$

在上列的條件機率關係中，偽陰性為敏感度的互補 (1-sen);而偽陽性為特異度的互補 (1-spe).

下表為利用乳房攝影術合併乳房超音波相對於僅使用乳房攝影術進行篩檢之工具表現：

指標 (95%信賴區間)	**乳房攝影術合併乳房超音波**	**僅使用乳房攝影術**
Sensitivity (%)	77.5 (61.6-89.2)	50 (33.8-66.2)
Specificity (%)	89.41 (88.16-90.57)	95.53 (94.67-96.30)
Area under ROC curve (%)	90 (83-95)	68 (53-80)
Positive predictive value (%)	10.1 (7.0-14.1)	14.7 (9.2-21.8)
Negative predictive value (%)	99.61 (99.27-99.82)	99.20 (98.77-99.51)

(2) 盛行率、陽性預測值與陰性預測值之關係

(A) 高盛行率：盛行率=25%

敏感度=80%, 特異度=80%, 陽性預測值=57.1%

		罹病狀態 (Y)		
		罹病 (Y=1)	未罹病 (Y=0)	小計
檢驗工具 (X)	陽性 (X=1)	400	300	700
	陰性 (X=0)	100	1200	1300
	小計	500	1500	2000

(B) 低盛行率：盛行率=5%

敏感度=80%, 特異度=80%, 陽性預測值=17.4%

		罹病狀態 (Y)		
		罹病 (Y=1)	未罹病 (Y=0)	小計
檢驗工具 (X)	陽性 (X=1)	80	380	460
	陰性 (X=0)	20	1520	1540
	小計	100	1900	2000

- 在某個特定族群中，高特異度的檢驗工具會得到較高的陽性預測值，若疾病盛行率增加，則其陽性預測值亦會隨之增加。
- 在某個特定族群中，高敏感度的檢驗工具會得到較高的陰性預測值，若疾病盛行率減少，其陰性預測值會增加。

歸納整理

Prior probability (Prevalence) ↑ ➔ PPV ↑, NPV ↓

Prior probability (Prevalence) ↓ ➔ PPV ↓, NPV ↑

Specificity ↑ ➔ PPV ↑

Sensitivity ↑ ➔ NPV ↑

臨床實例 2. 褥瘡風險預測分數 (Braden scale)

在病患照護上，如何評估患者處於產生褥瘡的高風險是一項重要議題。Braden scale 是褥瘡風險評估工具之一，其利用六項問題進行風險評估。該研究利用敏感度、特異度、陽性預測值陰性預測值評量該工具的效用。原文表 5 列出在不同的分數切點值下的敏感度、特異度、陽性預測值以及陰性預測值。表中可以看出分數切點不同其敏感度和特異度及陽性預測值、陰性預測值呈現消長之關係。

2. 貝氏理論之臨床推理 (e-Book 01: Chap 12, p1000-1013; e-Book 02: Chap 8, p.187-194; e-Book 03: Chap 12, p123-134, Chap 13, p139-144)

令 Y 為真正的疾病狀態(Y=1: 有病, Y=0: 沒病)，隨機變數 Y 服從二項分佈，分佈參數π即為群體中處於疾病狀態個體的比率 (盛行率)。令 X 為檢驗結果 (X=1: 陽性, X=0: 陰性)，隨機變數 X 亦服從二項分佈。則陽性預測值 (PPV=Pr(Y=1|X=1)) 可以視為事後機率 (進行檢驗後所得到的處於疾病狀態的比率)，前述的盛行率即為事前機率。運用貝氏定理可得到陽性預測值, 盛行率以及敏感度三者的關係：

$$\text{陽性預測值}(PPV) = P(Y=1 \mid X=1) = \frac{P(Y=1, X=1)}{P(X=1)}$$

$$= \frac{P(Y=1) \times P(X=1, Y=1) / P(Y=1)}{P(X=1)} \qquad (6)$$

$$= \underbrace{P(Y=1)}_{\text{事前機率, 盛行率}} \times \frac{P(X=1 \mid Y=1)}{\underbrace{P(X=1)}_{\text{正規化常數}}} \leftarrow \text{敏感度}$$

觀念: 若欲得到 P(Y=1|X=1) 以及 P(X=1|Y=1) 之間得關係時, 必定會運用到貝氏定理.

我們可以從流行病學觀點以及統計學觀點理解上式:

流行病學觀點: PPV ∝ 盛行率 × 敏感度

統計學觀點: 事後機率 ∝ 事前機率 × 概似度

其中, P(X=1)為邊際分佈，與疾病真正狀態是無關的，其來源可分為真正有病中陽性個案及沒病中陽性的個案，故可利用總和機率法則 (total law of probability) 分解為

$$P(X=1)=P(Y=1) \times P(X=1|Y=1)+ P(Y=0) \times P(X=1|Y=0)$$

$$= P(Y=1)\ P(X=1|Y=1)+[1-P(Y=1)]\ P(X=1|Y=0)$$

$$=\text{盛行率}\times\text{敏感度}+(1-\text{盛行率})\times(1-\text{特異度})$$

以統計學觀點來看，P(X=1) 也稱做 "正規化常數"，使事後機率介於 0 和 1 之間。

在此注意因為概似度並非為機率值，其可能大於 1。

若用流行病學的角度可以寫成:

$$PPV = P(Y=1) \times \frac{P(X=1|Y=1)}{P(X=1)}$$

$$= \frac{\text{盛行率}\times\text{敏感度}}{\text{盛行率}\times\text{敏感度}+(1-\text{盛行率})\times(1-\text{特異度})}$$

另以勝算(odds)的角度來看，若機率為 p，勝算則為 p/1-p。

事後機率亦可以事後勝算來表示：

$$\underbrace{P(Y=1|X=1)/(P(Y=0|X=1)}_{\text{事後勝算比}} = \underbrace{[P(Y=1)/P(Y=0)]}_{\text{事前勝算比}} \times \underbrace{[P(X=1|Y=1)/(P(X=1|Y=0)]}_{\frac{\text{敏感度}}{1-\text{特異度}}\ (\text{即偽陽性})}$$

{即概似度比}

移項可得：

$$\frac{\text{事後勝算比}}{\text{事前勝算比}} = \text{概似度比}\left(=\frac{\text{敏感度}}{1-\text{特異度}}\right)$$

範例：HIV 陽性盛行率與檢驗

在某個 HIV 盛行率為 0.001 的族群，利用檢驗工具得下表：

		HIV +	HIV –	
Test	+	95	1998	2093
Test	-	5	97902	97909

100 99900 100,000

⇨ Sen=0.95 Spe=0.98, LR=47.5, PPV=0.0454, Posterior odds =0.048.

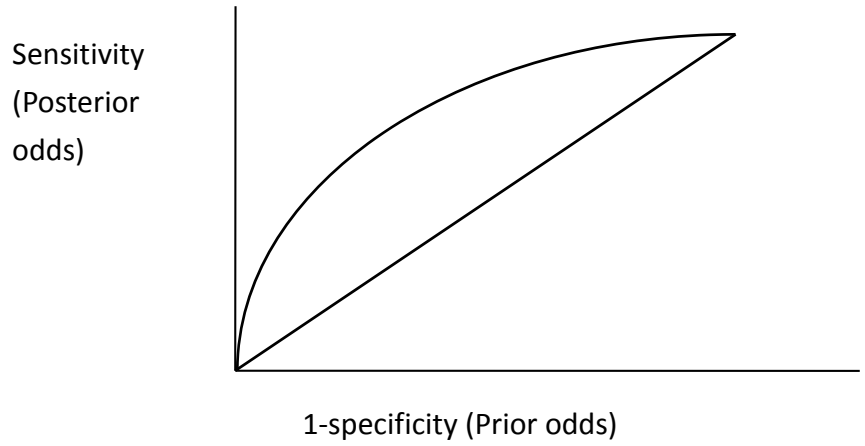
3. 接收者操作特徵曲線(Receiver operating characteristic, ROC curve)

(e-Book 03: Chap 12, p129-132)

對於評估臨床診斷的準確性，操作特徵曲線(ROC curve)是一個十分有用的指標，臨床運用上常可利用操作特徵曲線來表示檢測工具的預測效果。在曲線下的面積(area under curve，AUC)越大，表示檢測工具的正確性越高。

Y 軸：事後勝算比(Posterior odds)，在臨床檢驗中為敏感度

X 軸：事前勝算比(Prior odds)，在臨床檢驗中為 1-特異度



⇨ 事後機率(P(Y=1|X=1)) 與事前機率(P(Y=1))有高度相關，P(Y=1)越大則 P(Y=1|X=1)越大。

若事前機率與事後機率相等 (P(Y=1|X=1) = P(Y=1)，即 ROC curve 中的對角線,其斜率為 1)，表示檢測工具並不會影響診斷機率，則事後機率與事前機率為相等，表示檢測是無用的。

ROC curve 即為敏感度(Y 軸)對 1-特異度(X 軸)作圖, 因此, 圖中的曲線斜率(Y 軸比上 X 軸, 或敏感度比上 1-特異度) 即上式中的概似度比，所以概似度比為 1 即是 ROC curve 的斜率為 1，表示事後勝算比與事前勝算比相等，因此檢查前與檢查後對於判斷患者是否罹病並無不同，因此此種檢查是無效的。概似比

越大則 ROC 曲線下的面積會越大。

對於檢查工具而言,我們希望 ROC 曲線下的面積越大越好，最理想的情況是敏感度及特異度皆為 100%，但在現實情況下是不可能的，一般來說敏感度高,特異度就會低,兩者呈現消長的關係。基於這樣的消長關係,我們可以利用 ROC curve 尋找呈現連續數值檢測工具(如膽固醇、血壓值)的最佳切點值，使得利用該工具判斷罹病與否(事後機率)時得到最高的敏感度以及特異度。

臨床實例 1：乳癌篩檢工具比較

文中以 ROC curve 評估使用乳房攝影加上乳房超音波比上乳房攝影，節•果顯示乳房攝影加上乳房超音波對於乳癌診斷的 ROC 曲線下面積較大，亦即乳房攝影術合併乳房超音波的表現較僅使用乳房攝影術為佳。

運用實例: 免疫法糞便潛血檢查對於大腸直腸癌病變的診斷力

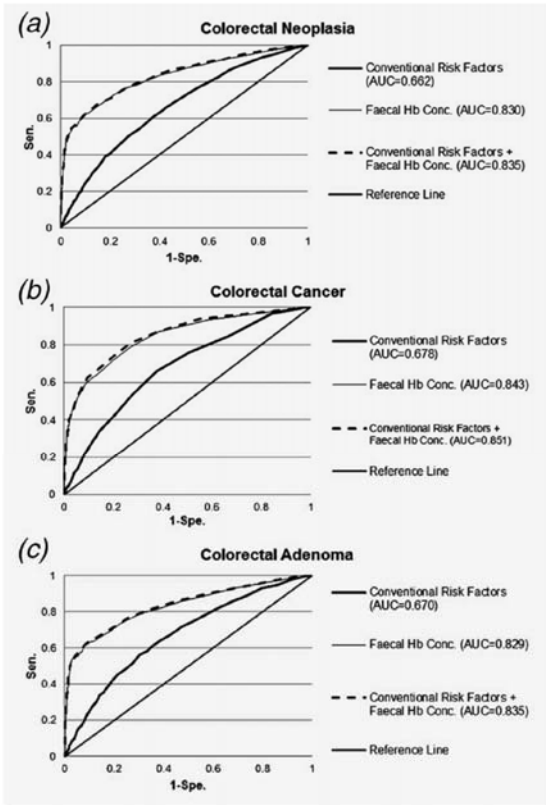


Figure 2. The receiver operating characteristic curves for the predictive model based on conventional risk factors only, faecal Hb concentration only, and the combination of conventional risk factors and FHbC.

參考文獻: Yen et al.: A new insight into fecal hemoglobin concentration-dependent predictor for colorectal neoplasia.

臨床實例 2.　褥瘡風險預測分數 (Braden scale)

文中利用不同切點值所算的敏感度以及特異度，繪製 ROC 曲線，並顯示以 18 分作為切點可以達到最佳的敏感度與特異度。(見原文圖 1)

II 臨床決策分析

(參考文獻

1. Kolmstrom et al.: Prostate specific antigen for early detection of prostate cancer: longitudinal study. BMJ 2009;339:b3537.

3. Kalkman et al.: Preoperative prediction of severe postoperative pain. Pain 2003 (105) 451-23)

1. 決策閾值模式 (Threshold model) (參考 e-book 03. Chap 13, p.139-150，Pauker SG, Kassirer JP. The threshold approach to clinical decision making. N Engl J Med. 1980 May 15;302(20):1109-17.)

A 在臨床上病人選擇治療可能模式，包括不必接受治療且不檢查、直接接受治療且不檢查及先經過檢查再進行治療 (兩階段)，因為通常檢查含有不準確部分，即上述所學偽陽性及偽陰性，因此會有如文中之決策分析之模式圖。

原文圖中，Tt 之臨界點表示，選取檢查與不治療是沒有差別，而 Trx 之臨界點是表示選擇檢驗和治療病人但不接受檢驗是沒有差別，這些臨界點值決定因素，包括檢查之危險性 (the risk of diagnostic test)、治療有病之病人之益處 (the benefit of treatment to patients who have the disease)、治療有病及無病病人之危險性 (the risk of the treatment to patients with and without disease)，以及檢查之準確性 (accuracy of test)。

Pauker 等人文章詳細計算在考慮這些不同因素下之決策值閾模式 (Threshold model)，以下我們先針對檢查之準確性示範如何利用前述貝氏臨床推理進行治療決策分析，以及以大腸直腸癌篩檢計畫為例，利用成本效益決策分析，呈現，除了篩檢檢測精確度外，考慮篩檢與未篩檢之利益與危險，在各方面考量下，是否要選擇進行大腸直腸癌篩檢。

臨床實例 1. 鏈球菌感染咽峽炎

一個 8 歲左右男孩向他的醫師主訴他三天來的症狀包括喉嚨痛、流鼻水、及發燒。他目扁桃腺有紅班及淋巴結腫大。他的父親十分擔心他為咽喉炎。主要的診斷應為鏈球菌性咽峽炎，如果延誤治療，疾病會更趨嚴重，潛在的併發症包含咽喉膿腫、風濕熱、腎絲球腎炎。鏈球菌性咽峽炎快速檢測法相當準確，敏感度及特異度可分別達 90%及 95%。

抗生素雖有好的療效，但有少部份副作用，例如過度使用會對社區有不良的嚴重後果。但不用抗生素治療，將會增加副作用及死亡的風險。醫師必需在抗生素治療對病人效益與可能的社區損害間取得平衡。

鏈球菌性咽峽炎發生的事前機率來自小兒臨床的表現。鏈球菌感染常伴隨喉嚨痛、發燒及扁桃腺腫。然而鏈球菌性咽峽炎就上呼道感染，較其他常見病毒性感染而言是較為罕見，但較少出現流鼻水，也常見有淋巴結滲出液，但這些症狀並未在這個男孩上看見。綜合這些情形，此男孩罹患鏈球菌性咽峽炎的事前機率約為 10%。

(1) 2X2 聯列表

事前機率：10%

敏感度=0.9，特異度=0.95

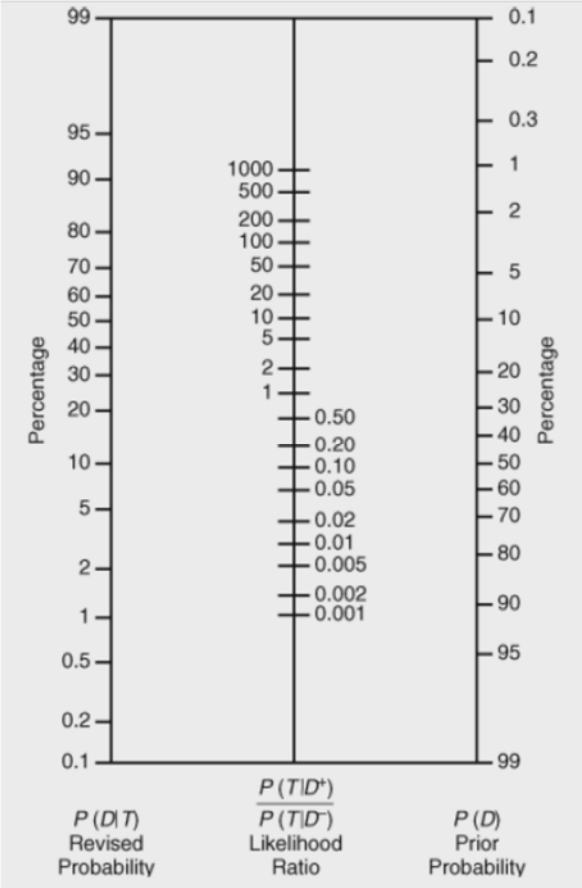
		鏈球菌感染咽峽炎		
		是	否	合計
快速檢驗	陽性	90	45	135
	陰性	10	855	865
	合計	100	900	1,000

PPV=67% →以抗生素治療

NPV=99% →不治療繼續追蹤

(2) 利用貝氏諾謨圖 (Bayesian nomogram)

為能夠方便使用貝氏臨床推理於臨床治療決策，流行病學家與統計學家，將事前得病機率與資料所得概似對比值 (Likelihood ratio)，所計算出事後機率做成一個臨床可便利查詢之圖，稱為 Bayesian nomogram。



貝氏諾謨圖(Nomogram for using Bayes' theorem)
(參考文獻

1. Fagan TJ. Nomogram for Bayes' theorem. (Letter.) N Engl J Med 1975;293:257.

2. Kolmstrom et al.: Prostate specific antigen for early detection of prostate cancer: longitudinal study. BMJ 2009;339:b3537)

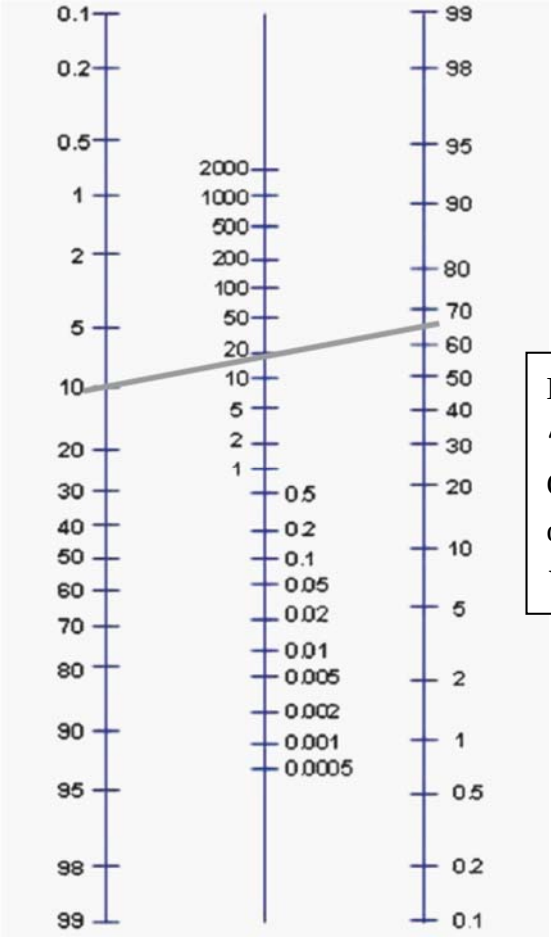
(A) 檢查為陽性

Likelihood ratio positive= $\frac{\text{Sen}}{1-\text{Sp}} = \frac{0.90}{0.05} = 18$

陽性事後勝算比=$\frac{\text{P(D|+)}}{\text{P}(\bar{\text{D}}\,|+)} = \frac{\text{P(D)}}{\text{P}(\bar{\text{D}})} \times \frac{\text{Sen}}{1-\text{Spe}} = \frac{1}{9}$ X18= 2 (P(D)=0.1)

事後機率= $\frac{2}{1+2} = \frac{2}{3}$ = 0.67 →選擇使用抗生素

其他情況可以用下面諾謨圖直接求得，如此可避免每次計算 2X2 聯列表。



Likelihood ratio nomogram for positive rapid strep test.
快速鏈球菌檢查的概似度比值(以諾謨圖方式呈現)
Given a positive rapid strep test, the post-test probability of disease is about 67% 在快速鏈球菌檢查為陽性結果下，檢查後得病機率約67%

Pre-Test probability 檢查前得病機率(%)　Likelihood ratio 概似度比值　Post-Test probability 檢查後得病機率(%)

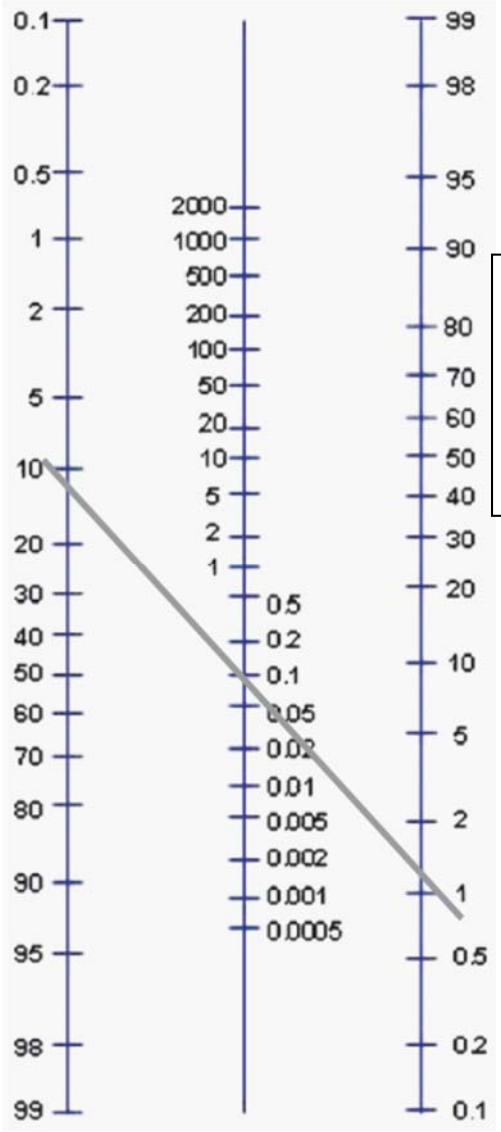
(B) 檢查為陰性

Likelihood ratio negative = $\frac{1-\text{Sen}}{\text{Sp}}$

陰性事後勝算比= $\frac{P(D|-)}{P(\bar{D}|-)}=\frac{P(D)}{P(\bar{D})}$ X $\frac{P(-|D)}{P(-|\bar{D})}$ ($=\frac{1-\text{Sen}}{\text{Sp}}$) = $\frac{1}{9}\times\frac{0.1}{0.95}$=0.012

P(D|-)=$\frac{0.012}{1+0.012}$=0.012

如果檢驗為陰性，則得病之機率為 1.2%。 →選擇觀察，不選擇使用抗生素。其他情況也可以用下面諾謨圖直接求得。



Likelihood ratio nomogram for positive rapid strep test.
快速鏈球菌檢查的概似度比值(以諾謨圖方式呈現)
Given a negative test result, the post-test probability of disease is about 1% 在快速鏈球菌檢查為陰性結果下，檢查後得病機率約1%

Pre-Test probability 檢查前得病機率(%)

Likelihood ratio 概似度比值

Post-Test probability 檢查後得病機率(%)

臨床實例 2. 術後疼痛之預測

研究目的: 發展利用術前資料預測病患術後嚴重疼痛之工具

統計方法: 多變相羅吉斯迴歸

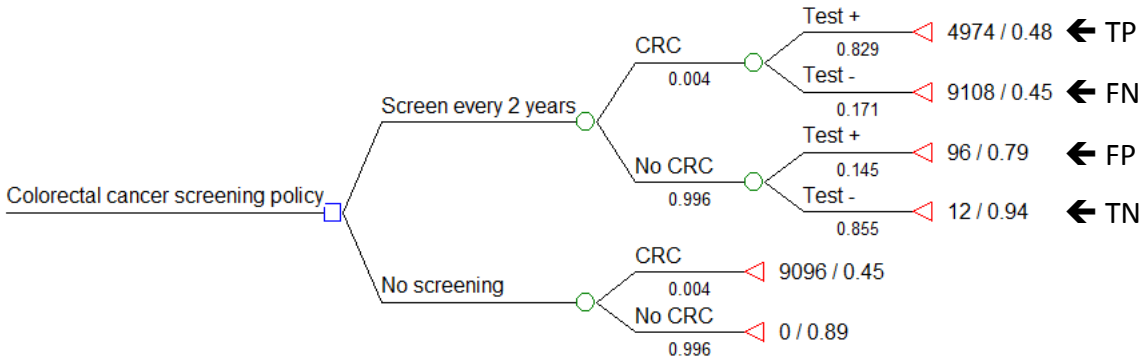
諾謨圖運用:迴歸分析結果可以表示為諾謨圖，並用一預測術後疼痛發生的可能性。

例如: 有一名女性病患(3)，55 歲(8)，術前疼痛分數 8 (18)，將接受骨科手術 (14)，手術切口屬於中到大(medium to large)(3)，術前評估關於受術訊息需求程度為第二級(9)、焦慮分數 12 (6)。相對應的術前評估總分為 61 分，因此術後嚴重疼痛可能性約為 65%。

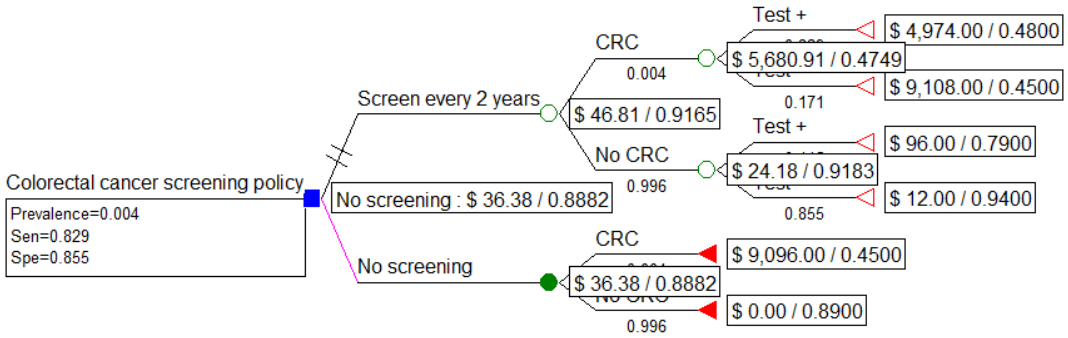
2. 成本效益決策分析(Cost-effectiveness descision analysis) (參考 e-book 01. Chap 12, p.1022-1048)

下圖為大腸直腸癌篩檢決策圖，在圖中□稱為決策節，其後續銜接可能考慮的決策，圖中○稱為機率節，表示後續分枝進展由機率決定，在本範例中，不篩檢的決策之後，會將族群分為有病(CRC)與無病(No CRC)兩類，有病決策節所佔的比例則由盛行率(Prevalence)決定。兩年一次篩檢(Screen every 2 years)的決策之下亦將族群分為有病(CRC)與無病(No CRC)兩類，不論哪一類人，均會接受篩檢，可能會有陽性及陰性的結果，在有病的族群之中，篩檢結果陽性的機率即由敏感度(Sen)決定，亦即真陽性(TP)，另一群人則為偽陰性(FN)；無病的族群之中，篩檢結果陰性的機率即由特異度(Spe)決定，亦即真陰性(TP)，另一群人則為偽陽性(FP)。由於結果於此底定，因此最後以終結節(Payoff)(◁)表示。在本範例中，我們示範成本效益分析(Cost-effectiveness analysis)的結果，因此，在終結節之後則顯示此分枝至此所對應之成本及效益。結果可以看到真陽性(TP)的個案因為早期偵測出癌症，因此其相對應的成本較未篩檢策略下的癌症個案低，且效益(此例為生命品質)較高；偽陰性(FN)個案因為篩檢之時並未檢測出，因此其效益同不篩檢的癌症個案，但因為曾接受篩檢，因此成本會略高。在篩檢決策之下沒有罹病的

個案仍需考慮其篩檢成本，或者偽陽性(FP)的個案會再增加其確診成本，在效益方面，偽陽性(FP)會因為錯誤的診斷，使個案憂慮罹病後果而造成效益較沒有篩檢策略之下的無病個案為低，相對地，真陰性(TN)個案則因確定沒有罹病而有較佳的效益。因為兩個決策之下各有權衡之處(Trade-off)，因此需要決策分析加以釐清。



決策分析架構建立完畢之後，可利用 Rollback 方式回推各個決策之下總結之所需成本及對應之效益，本範例之 Rollbak 結果如下圖，並整理如下表。



結果顯示，未篩檢及兩年一次篩檢所對應之平均成本及效益分別為$36.38、$46.81 及 0.8882、0.9165，結果顯示篩檢較未篩檢平均每人需多花$10.43，此為增加成本(Incremental cost)，但可多得 0.0283 品質調整人年(Quality-adjusted life-year, QALY)，此為增加效益(Incremental Effect)，兩者相除所得到的值為增加

成本效益比值(Incremental cost-effectiveness ratio, ICER)，代表每欲多獲得一個QALY，所需花費的成本。

策略	平均成本 ($)	增加成本 ($)	平均效益 (品質調整人年, QALY)	增加效益 (品質調整人年, QALY)	增加成本效益比值
未篩檢	36.38		0.8882		
兩年一次篩檢	46.81	10.43	0.9165	0.0283	369

流行病學方法論及實驗
(The Methods of Epidemiology and Practices)
貝氏柯南護理流行病學

授課教師：陳秀熙 教授/許辰陽 博士

台灣大學流行病學與預防醫學研究所

I 臨床推理— 檢驗及診斷工具的正確性

Accuracy of Test and Diagnosis

臨床實例(1) 乳癌篩檢工具比較:
乳房攝影合併超音波 VS 乳房攝影

ORIGINAL CONTRIBUTION

Combined Screening With Ultrasound and Mammography vs Mammography Alone in Women at Elevated Risk of Breast Cancer

Wendie A. Berg, MD, PhD
Jeffrey D. Blume, PhD
Jean B. Cormack, PhD
Ellen B. Mendelson, MD
Daniel Lehrer, MD
Marcela Böhm-Vélez, MD
Etta D. Pisano, MD
Roberta A. Jong, MD
W. Phil Evans, MD
Marilyn J. Morton, DO
Mary C. Mahoney, MD
Linda Hovanessian Larsen, MD
Richard G. Barr, MD, PhD
Dione M. Farria, MD, MPH
Helga S. Marques, MS
Karan Boparai, RT
for the ACRIN 6666 Investigators

JAMA. 2008;299(18):2151-2163

Using Sonography to Screen Women with Mammographically Dense Breasts
以乳房超音波篩檢緻密性乳房

- 67-year-old woman with dense Breast Imaging Reporting and Data System (BI-RADS) [14] category 3 breast tissue.
  - **A** and **B**, Mediolateral oblique (**A**) and craniocaudal (**B**) screening mammograms reveal no abnormalities.
  - **C**, Screening sonogram shows solid hypoechoic mass that measures 9 mm wide by 5 mm high. Angular margins (*arrows*) between mass and surrounding tissues are suspicious for malignancy. Sonographically guided biopsy (not shown) revealed invasive ductal carcinoma.
  - **D** and **E**, Right mediolateral (**D**) and right craniocaudal (**E**) mammograms obtained after sonographically guided wire localization show uniformly dense breast tissue with no evidence of mass in hookwire area.

1. Sensitivity (敏感度), specificity (特異度), positive predictive value (PPV) (陽性預測值) and negative predictive value (NPV) (陰性預測值)

(e-Book 01: Chap 12, p988-1000; e-Book 03: Chap 12, p124-129, Chap 13, p139-143)

2x2 Table (列聯表)

• 關鍵字：**Sensitivity (敏感度)**, **specificity (特異度)**, **positive predictive value(PPV) (陽性預測值)** and **negative predictive value (NPV) (陰性預測值)**

		乳癌個案(Y) 是 (Y=1)	否 (Y=0)	
乳房攝影加超音波檢查(X)	陽性 (X=1)	a (31)	b (275)	a+b (306)
	陰性(X=0)	c (9)	d (2322)	c+d (2331)
		a+c (40)	b+d (2597)	

Indicators

2x2 Table

		乳癌個案(Y) 是 (Y=1)	否 (Y=0)	
乳房攝影加超音波檢查(X)	陽性 (X=1)	a (31)	b (275)	a+b (306)
	陰性(X=0)	c (9)	d (2322)	c+d (2331)
		a+c (40)	b+d (2597)	

$$Sen\ (\text{Sensitivity}) = P(X=1 \mid Y=1) = \frac{a}{a+c} = \frac{31}{31+9} = 77.5\%$$

$$Spe\ (\text{Specificity}) = P(X=0 \mid Y=0) = \frac{d}{b+d} = \frac{2322}{275+2322} = 89.4\%$$

$$PPV\,(\text{PositivePredictiveValue}) = P(Y=1 \mid X=1) = \frac{a}{a+b} = \frac{31}{31+275} = 10.1\%$$

$$NPV\,(\textit{Negative}\ \text{Predictive Value}) = P(Y=0 \mid X=0) = \frac{d}{c+d} = \frac{2322}{2322+9} = 99.6\%$$

$$FN(\textit{False negative}) = P(X=0 \mid Y=1) = \frac{c}{a+c} = \frac{9}{31+9} = 22.5\%$$

$$FP(\textit{False posotive}) = P(X=1 \mid Y=0) = \frac{b}{b+d} = \frac{275}{275+2322} = 10.9\%$$

Clinical Example of breast cancer

Estimate (95% CI)	Mammography + Ultrasound	Mammography alone
Sensitivity (%)	77.5 (61.6-89.2)	50 (33.8-66.2)
Specificity (%)	89.41 (88.16-90.57)	95.53 (94.67-96.30)
Area under ROC curve (%)	90 (83-95)	68 (53-80)
Positive predictive value (%)	10.1 (7.0-14.1)	14.7 (9.2-21.8)
Negative predictive value (%)	99.61 (99.27-99.82)	99.20 (98.77-99.51)

7

臨床實例(2) 褥瘡風險預測分數

Predictive Validity of the Braden Scale for Pressure Sore Risk in a Nursing Home Population

Barbara J. Braden and Nancy Bergstrom

8

The relationship between prevalence and PPV and NPV

- High prevalence: Prevalence=25%
- Sensitivity=80%, Specificity=80%, PPV=57.1%

		Disease status (Y)		
		Disease (Y=1)	Non-disease (Y=0)	Total
Test (X)	Positive (X=1)	400	300	700
	Negative (X=0)	100	1200	1300
	Total	500	1500	2000

- Low prevalence: Prevalence=5%
- Sensitivity=80%, Specificity=80%, PPV=17.4%

		Disease status (Y)		
		Disease (Y=1)	Non-disease (Y=0)	Total
Test (X)	Positive (X=1)	80	380	460
	Negative (X=0)	20	1520	1540
	Total	100	1900	2000

- A test with high specificity will yield a high PPV in a given population. When disease prevalence in the population increases the PPV will increase.
- A test with high sensitivity will yield a high NPV in a given population. As disease prevalence in the population decreases the NPV will also increase.

9

The relationship between prevalence and PPV and NPV

- Prior probability (Prevalence) ↑ ➔ PPV ↑, NPV ↓
- Prior probability (Prevalence) ↓ ➔ PPV ↓, NPV ↑

- Specificity ↑ ➔ PPV ↑
- Sensitivity ↑ ➔ NPV ↑

10

2. Bayesian Theorem for clinical reasoning

e-Book 01: Chap 12, p1000-1013;
e-Book 02: Chap 8, p.187-194;
e-Book 03: Chap 12, p123-134, Chap 13, p139-144)

Y ：true disease status (Y=1: Disease Y=0: non-disease)

X ：result of test (X=1: positive; X=0:negative)

The **positive predictive value (PPV)** is

$$P(Y=1|X=1) = \frac{P(Y=1) \times P(X=1|Y=1)}{P(X=1)}$$

Normalizing constant.
A marginal distribution

$$= P(Y=1) \times P(X=1|Y=1) + P(Y=0) \times P(X=1|Y=0)$$
$$= P(Y=1) \times P(X=1|Y=1) + [1 - P(Y=1)] \times P(X=1|Y=$$

Epidemiological viewpoint: PPV α Prevalence × Sensitivity

Statistical viewpoint: Posterior α Prior × Likelihood

11

PPV: Bayesian perspective

Remind that PPV

$$P(Y=1|X=1) = \frac{P(Y=1) \times P(X=1|Y=1)}{P(X=1)}$$

➔

$$\frac{P(Y=1|X=1)}{P(Y=0|X=1)} = \frac{P(Y=1)}{P(Y=0)} \times \frac{P(X=1|Y=1)}{P(X=1|Y=0)}$$

$$\{\text{Posterior Odds}\} = \{\text{Prior Odds}\} \times \{Likelihood\ Ratio\ (=$$

12

Bayesian Theorem for clinical reasoning: HIV example

Prevalence and HIV test

In a population with an HIV prevalence of 0.001 (prior odds: 0.001)

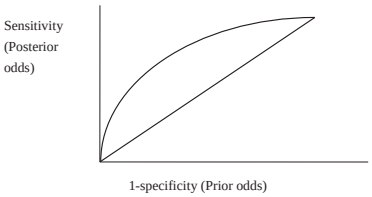
		HIV +	HIV –	
Test	+	95	1998	2093
Test	-	5	97902	97909
		100	99900	100,000

Sen=0.95
Spe=0.98
LR=47.5
PPV=0.0454
Posterior odds =0.048

3. Receiver operating characteristics (ROC)

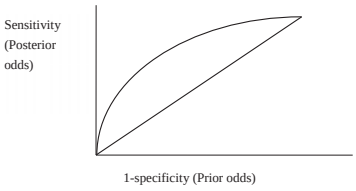
(e-Book 03: Chap 12, p129-132)

- A very useful indicator for assessing the accuracy of clinical diagnosis.
- The larger the area under curve (AUC), the more accurate the test is.



Receiver operating characteristics : Interpretation

P(Y=1)↑ => P(Y=1|X=1)↑.



- When P(Y=1|X=1) = P(Y=1), the posterior probability is equivalent to the prior probability (the angle of the linear line is 45^0).
- Optimal cutoff given the test is based on a continuous variable such as cholesterol, hypertension etc.

II 臨床決策分析

1. 決策閾值模式 (Threshold model)

(e-Book 01: Chap 12, p985-987;
e-book 03. Chap 13, p.139-150)

臨界點
T_t: 選取**檢查**與**不治療**沒有差別

T_{trx}: 選擇**檢驗**和**直接治療病人但不接受檢驗**沒有差別

決定因素：

- 檢查之危險性 (the risk of diagnostic test)
- 治療有病之病人之益處 (the benefit of treatment to patients who have the disease)
- 治療有病及無病病人之危險性 (the risk of the treatment to patients with and without disease)
- 檢查之準確性 (accuracy of test)

臨床實例
鏈球菌感染咽峽炎
(streptococcal pharyngitis)

- 臨床案例

一個8歲左右男孩向他的醫師主訴他三天來的症狀包括喉嚨痛、流鼻水、及發燒。他目扁桃腺有紅班及淋巴結腫大。他的父親十分擔心他為咽峽炎。主要的診斷應為鏈球菌性咽峽炎，如果延誤治療，疾病會更趨嚴重，潛在的併發症包含咽喉膿腫、風濕熱、腎絲球腎炎。

臨床實例
鏈球菌感染咽峽炎
(streptococcal pharyngitis)

- 檢驗及治療方案
  - 鏈球菌性咽峽炎快速檢測法 (Rapid streptococcal test)
    - 敏感度 (Sensitivity) 90%
    - 特異度 (Specificity) 95%
  - 抗生素治療
    - 療效好，但有少部份副作用，例如過度使用會對社區有不良的嚴重後果。
    - 不使用抗生素治療，將會增加副作用及死亡的風險。
    - 醫師必需在抗生素治療對病人效益與可能的社區損害間取得平衡。

19

臨床實例
鏈球菌感染咽峽炎
(streptococcal pharyngitis)

- 事前機率
  - 來自小兒科醫師的臨床經驗。
  - 鏈球菌性咽峽炎較其他常見病毒性感染而言是較為罕見，較少出現流鼻水，也可能有淋巴結滲出液。
  - 鏈球菌感染常伴隨喉嚨痛、發燒及扁桃腺腫，但這些症狀並未在這個男孩上看見。
  - 綜合這些情形，此男孩罹患鏈球菌性咽峽炎的事前機率約為10%。

20

2x2聯列表(Contingency table)

		鏈球菌感染咽峽炎 (streptococcal pharyngitis)		
		是	否	合計
快速檢驗 (Rapid streptococcal test)	陽性	90	45	135
	陰性	10	855	865
	合計	100	900	1,000

PPV=67% →以抗生素治療
NPV=99% →不治療繼續追蹤

21

快速鏈球菌檢查為陽性結果下
(Given a positive rapid strep test)

- 陽性概似度比值 (Likelihood ratio positive)

$$= \frac{\text{Sen}}{1-\text{Sp}} = \frac{0.90}{0.05} = 18$$

- 陽性事後勝算比 (Posterior odds (given test is positive)) (當P(D)=0.1)

$$= \frac{P(D|+)}{P(\bar{D}|+)} = \frac{P(D)}{P(\bar{D})} \times \frac{\text{Sen}}{1-\text{Spe}} = \frac{1}{9} \text{ X18= 2}$$

- 事後機率

$$= \frac{2}{1+2} = \frac{2}{3} = 0.67$$ → 選擇使用抗生素!!

22

參考 Fagan T.J. Nomogram for Bayes' theorem. (Letter.) N Engl J Med 1975;293:257.

快速鏈球菌檢查為陽性結果下
(Given a positive rapid strep test)

- 見文中貝氏諾莫圖

Given a positive rapid strep test, the post-test probability of disease is about 67%
在快速鏈球菌檢查為陽性結果下，檢查後得病機率約67%

23

快速鏈球菌檢查為陰性結果下
(Given a negative rapid strep test)

- 陰性概似度比值 (Likelihood ratio negative)
- $\frac{1-\text{Sen}}{\text{Sp}}$
- 陰性事後勝算比 (Posterior odds (given test is negative)) (當P(D)=0.1)

$$= \frac{P(D|-)}{P(\bar{D}|-)} = \frac{P(D)}{P(\bar{D})} \text{ X } \frac{P(-|D)}{P(-|\bar{D})} \left(= \frac{1-\text{Sen}}{\text{Sp}} \right) = \frac{1}{9} \times \frac{0.1}{0.95} = 0.012$$

- 事後機率

$$P(D|-) = \frac{0.012}{1+0.012} = 0.012$$ → 繼續觀察

24

參考 Fagan TJ. Nomogram for Bayes' theorem. (Letter.) N Engl J Med 1975;293:257.

快速鏈球菌檢查為陰性結果下 (Given a negative rapid strep test)

貝氏諾謨圖 (Bayesian nomogram)

Given a negative test result, the post-test probability of disease is about 1%
在快速鏈球菌檢查為陰性結果下，檢查後得病機率約1%

臨床實例(2)術後疼痛之預測

- 臨床案例

有一名女性病患，55歲，術前疼痛分數8，將接受骨科手術，手術切口屬於中到大(medium to large)，術前評估關於受術訊息需求程度為第二級、焦慮分數12。

術後嚴重疼痛可能性為何?

Preoperative prediction of severe postoperative pain

C.J. Kalkman[a,*], K. Visser[b], J. Moen[a], G.J. Bonsel[c], D.E. Grobbee[d], K.G.M. Moons[a,d]

C.J. Kalkman et al. / Pain 105 (2003) 415–423

- 有一名女性病患(3)，55歲(8)，術前疼痛分數8 (18)，將接受骨科手術 (14)，手術切口屬於中到大(medium to large)(3)，術前評估關於受術訊息需求程度為第二級(9)、焦慮分數12 (6)

- 相對應的術前評估總分為61分，因此術後嚴重疼痛可能性約為 65%

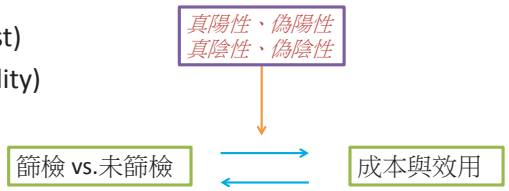
2. 成本效益決策分析 (Cost-effectiveness decision analysis)

- Strategies:

篩檢 vs.未篩檢 (Screening compared with no screening)

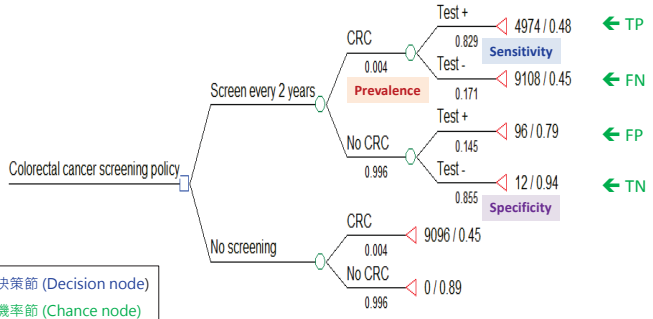
– 成本 (Cost)

– 效用 (Utility)



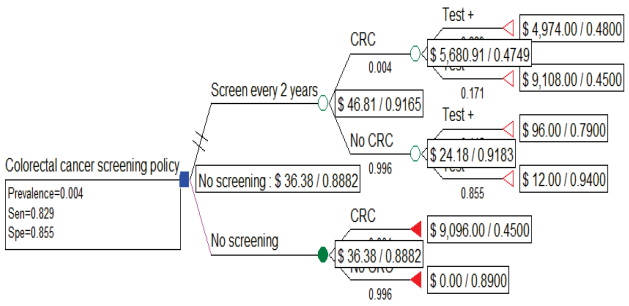
(e-book 01. Chap 12, p.1022-1048)

實例: 大腸直腸癌篩檢 (Colorectal cancer screening)



(e-book 01. Chap 12, p.1022-1048)

成本分析決策模式 (Cost-effectiveness decision analysis)



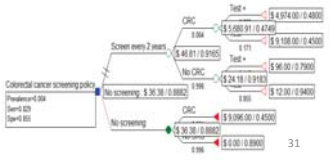
(e-book 01. Chap 12, p.1022-1048)

成本分析決策模式
(Cost-effectiveness decision analysis)

策略(Strategy)	平均成本 (Mean cost) ($)	增加成本 (Incremental cost)($)	平均效益(Mean Effect) (品質調整人年, QALY)	增加效益 (Incremental Effect) (品質調整人年, QALY)	增加成本效益比值 (ICER)
未篩檢	36.38		0.8882		
兩年一次篩檢	46.81	10.43	0.9165	0.0283	369

ICER (Incremental cost-effectiveness ratio)

$$= \frac{\text{Incremental cost}}{\text{Incremental effectiveness}}$$



31

八、 賽先生護理流行病學

一、隨機分派對照研究之觀念 (參考 e-book 01_ p25, p48-52、e-book 02_Ch2_ p.28-31、consort checklist)

在因果關係之研究中，為了研究 X(因)導致 Y(果) 的產生，我們須要利用實證醫學之原則中最具公信力的研究設計 (隨機分派對照研究設計，Randomized controlled trial， RCT，或以簡寫"R"表示) 證明兩者的因果關係。關於隨機分派對照研究設計，有兩種主要的設計方式: 平行設計以及交叉設計，最常見為平行設計，也是本次之重點，交叉設計較複雜在下次說明。平行設計圖示如下：

R
- X → Y_X
- $\bar{X}$ → $Y_{\bar{X}}$

"R" 代表 randomization 或 random allocation。上圖表示在試驗進行前，以一種隨機方式或過程(random process)將試驗參與者隨機分派到介入組 (X) 或非介入組 ($\bar{X}$)，分別產生所關心的結果，(介入組的結果:Y_X，以及非介入組的結果:$Y_{\bar{X}}$); 隨機的方式分配方式有很多種方式，例如抽取亂數，奇數者分派到介入組而偶數者分派到非介入組。總之，這隨機的分派方式使得試驗參與者是否接受介入不是經由個人意志而是隨機過程所決定。

隨機分派對照試驗利用"R"，也就是 randomization 或 random allocation，使參與試驗的受試者除了介入的不同外，其他干擾及影響偏差之因素的特性的比例都相近，因此消除了除了介入外其他可能造成結果不同的因素，這是為何隨機分派對照試驗是實證醫學中最具有公信力之科學證據，因此非採用隨機分派對照試驗(例如觀察性研究) ，就會產生謬誤(fallacy)。

***安慰劑效應以及對照組之功用**

安慰劑效應(Placebo effect)

在芝加哥， Hawthorne 工廠的老闆發現工廠產量不好，員工反應是因為照

明設施不好，因此他改進了照明設施，因而增加了產量。過了一段時間照明設備又變不好，但產量並沒有減少。

論點：產量的增加可能是因為老闆聽取員工意見進而鼓勵員工，員工心情變好而更認真工作，進而產量增加，所以即使後來照明設備不好但產量仍是穩定的。此情形稱做“Hawthorne 效應”。下列之例子亦有相同之現象。

(1) 學生考試時老師站在旁邊，學生感覺到壓力而導致考試考不好。

(2) 營養學研究無法親自到家中測量營養，因為當受試者被研究人員監測時， 受試者對食物的攝取會變得較拘謹，平常大吃大喝可能會變得少放醬油、鹽的使用減少等等。

(3) 臨床試驗中的安慰劑(placebo)，試驗分成兩組，一組吃藥，另一組沒吃藥的也需要吃跟藥做的一模一樣的安慰劑，安慰劑造成的效應即是 Hawthorne effect，因為有吃藥跟沒吃藥對受試者會有不同的影響，而安慰劑可以使試驗組與對照組都有吃藥的動作。

(4) 中醫研究中如針灸、穴道按壓很容易受到 Hawthorn effect 的影響，因為實驗組有受到治療的動作，而對照組沒有；故中醫研究的對照組須進行"Sham control"：例如想研究針灸對某疾病是否有影響時，在對照組找一個沒有副作用但也不會產生療效的地方也刺一下 (如同試驗組針灸的動作)，以辨別究竟是在穴位上針灸有用還是刺一下這個心理行為有用(如同 Hawthorn effect) 。欲解決 Hawthorne 效應的方法即是進行 RCT 並且在對照組使用安慰劑，故後來稱做 placebo effect。

二、隨機分派基礎研究設計與分析方法 (參考 e-book 03_ p68-73、consort checklist)

上述之平行設計又因有無前測分為下列兩種:

1. 有前測的完全隨機化設計 (Complete randomized design with pre-test)

$$R\begin{cases} O & X & O \\ O & \bar{X} & O \end{cases}$$

某試驗欲研究減重計畫 (weight reduction program) 是否有效， 採用有前測的完全隨機化設計(CRD with pre-test)如下:

X	$\bar{X}$
(介入組，參加減重計畫)	(對照組，未參加減重計畫)
O_1 O_2	O_3 O_4

註: O_1，O_3: 前測

O_2，O_4: 後側

這樣設計其爭議點在於是否需要前測？使用前測的原因有三:

(A) 因隨機分派的機制可保證兩組在前測的分布應是相同的，因此若兩組前測有所差異則表示隨機分派的機制有問題， 此時前測可以用來評估隨機分派的機制是否成功(通常隨機分配試驗之文章結果的第一個表)。

(B) 因為進行 CRD 會有介入組及非介入組兩組，而又各有前後測(O_1、O_3 為前測，O2、O4 為後測)，若隨機分派機制成功則不需要在乎 O_1 以及 O_3，只需比較 O_2 與 O_4。但在臨床運用上也有需要看介入前後變化的情形，這時必須要做前測。

(C) 分析具有前測的資料時，可以將前測(O_1、O_3) 如同對干擾因子的調整一般，作為變項(X)放入迴歸中。這樣通常在分析介入得較果時會更容易得到顯著的結果，在統計上會得到好處: 也就是將 O_1、O_3 當作基準值(Baseline)放入迴歸時通常會改善對 O_2 以及 O_4 分析時的精確度(precision)。

2. 沒有前測的完全隨機化設計 (Complete randomized design without pre-test):

$$R\begin{cases} X & O_2 \\ \bar{X} & O_4 \end{cases}$$

如此只要比較後測之結果(O_2 及 O_4)

3. 統計分析之概念

	結果變項(Y)性質	**樣本統計值平均值**	**非迴歸分析方法**	**迴歸分析方法**
二組 (X=1/0)	(1)連續變項	$\bar{Y}$(平均值)	獨立 t-test $(\bar{Y}_1 - \bar{Y}_0)$	線性迴歸 Y=a+bx
	(2)二元變項 (Y=1/0)	$\hat{p}$(比例)=d/n d: 事件 n: 總數	X^2 test(2×2 列聯表)及比例檢定 $\widehat{p_1} - \widehat{p_0}$	羅吉斯迴歸 Logit(P(Y=1|X))=a+bx
	(3)存活時間 (Y=time to event)	存活函數 S(t) 風險函數 h(t)	Hazard Rate $=h_1(t)/\ h_0(t)$	Cox 迴歸模式(Cox proportional hazard regression model) $h(t)=h_0(t)\exp(a+bx)$
多組 (k=3) X=1/2/3	(1)連續變項	$\bar{Y}_1 - \bar{Y}_2)$/ $\bar{Y}_2 - \bar{Y}_3)$/ $\bar{Y}_1 - \bar{Y}_3)$	One-way Analysis of Variance(ANOVA)	$Y=a+b_1X_1+b_2X_2$ X=1 X=2 X=3 X_1 0 1 0 X_2 0 0 1

迴歸模式之基本想法

$$Y=a+bx+e \quad e\sim N(0,\sigma^2)$$

$X=1 \quad \bar{y_1} = a + b$

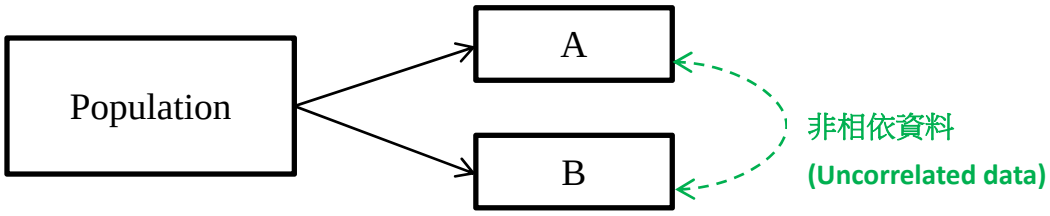
$X=0 \quad \bar{y_0} = a$

$\bar{y}_1 - \bar{y}_0 = b$ (迴歸係數也就是斜率), b 越大, 則兩組平均值差異越大

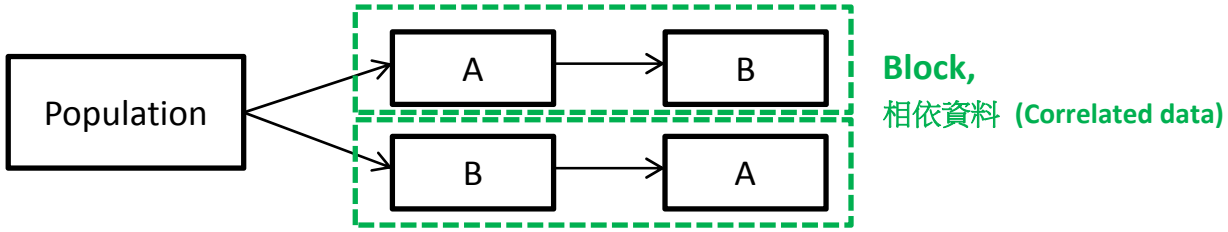
三、隨機分派研究設計應用實例

一般研究設計分為平行設計以及組內設計如下圖。平行設計以隨機分配(randomization)精神稱為完全隨機化設計(completely randomized design)；而組內設計一般稱為隨機區集設計(randomized block design)。

平行設計(Parallel design)



組內設計(Within group design)

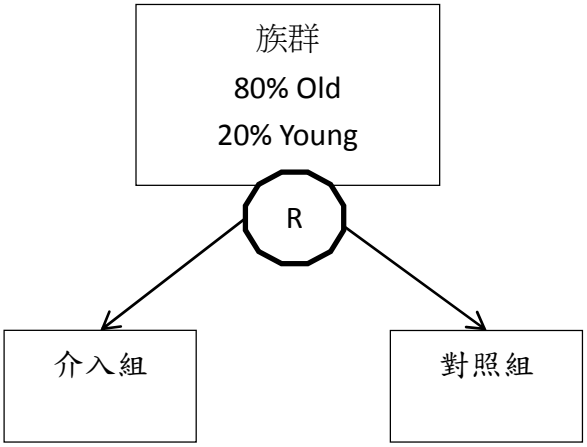


在 4/10 的課程中介紹了平行設計(parallel design),今天的課程將介紹隨機區集設計(randomized block design,RBD) 特別是交叉設計 (crossover design)) 及其他複雜設計 (other complex designs)。

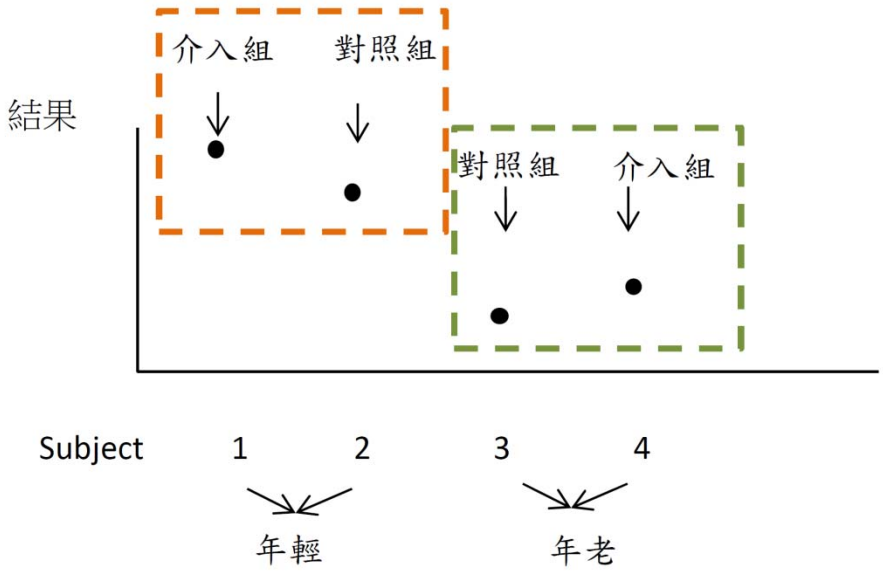
1. 隨機區集設計 (Randomized Block Design, RBD)

使用隨機區集設計(RBD)的緣由通常是在進行平行設計時兩組可能因樣本數太少或某些變項分佈比例差異太大，導致在隨機分配(randomization)過程其兩組有差異。以下舉一例來示範。

如下圖，當欲分派之族群中年輕個案占的比例很低(20%)，而且年齡與研究感興趣之結果有關時:例如年輕者有較佳的反應,而年老者的反應較差,圖示如下：



由於族群中年輕個案占的比例很低，所以在兩組間的年齡分佈即使經由隨機分配過程兩組也會有不平衡（age imbalance)之狀態，如此年齡之干擾也可能就無法藉由隨機分配過程來消除。故此時可採用隨機區集試驗。將年輕的個案 1、2 視為一個區集(block)，隨機分派一位至介入組，一位至對照組；相同的也將年老的個案 3、4 視為一個區集(block)，隨機分派一位至介入組，一位至對照組。



上述例子中可拓展將隨機分配單位化成不同單位如：鄉鎮、城市、學校、班級或者是人(測量多次)。原則上區集內(within block)之個人其有關結果要保持同質性(homogeneous)，而區集間(between block)要保持異質性(heterogeneous)，如此可以使得原來若用完全隨機化設計(complete

randomized design, CRD)之誤差變異(error variance)減小，並可以利用較少樣本數達到相同效益，提高效率。有關使 block 保持均質性之方法有下列幾種：

(1) **配對(Matching)：**

可利用 matching 方式將相同特質 k 個個案加以配對，而每一個配對就是一個 block，每一個 block 中的每個個案僅接受一種介入(intervention)，而須接受那種介入則由隨機分派來決定。

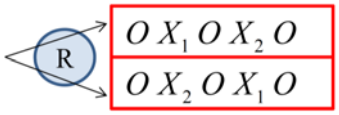
有些變項(如 sex 及 age)在實驗設計時並無法由研究者來操控(例如上述的 age imbalance 的情形)，為了降低研受試者間的變異，研究者只好使用 block design 降低原本若使用完全隨機化設計(CRD)所得到之誤差變異量。

例子 1.

某研究欲探討體能訓練計畫(physical training program)是否可以加強體力。因為男女體力先天上可能有別，研究者以性別做為配對(block)，一個男性同時配對另一個男性;其中一個接受 physical training program，另一個則未接受。女性 block 也依照相同的方式分派。依此原則完成隨機區集設計。

(2) **重複測量(Repeated Measurement)：**

也就是每個 subject 接受 m 個介入(intervention)而且其介入之順序不同，通常可以使用 counter balancing design 來決定順序，例如 m=2 之設計如下：



例子 2.

在臨床試驗中所使用之交叉設計(crossover design)就是一種用同一個人作為區集(block)接受不同治療方法之隨機區集設計(RBD)，以兩種治療方法為例：

	Run-in period	階段一	洗刷期(Washout)	階段二
結果	Z_{ij}	O_{ij1}		O_{ij2}
第一組（n_1）	---	A	---	B
第二組（n_2）	---	B	---	A

附註說明 i: group, j: subject

A: 治療組 A, B: 治療組 B

n_1:第一組樣本數, n_2:第二組樣本數

Z_{ij}: run-in period 觀察值

O_{ij1}: group i, subject j, 階段一之觀察結果

O_{ij2}: group i, subject j, 階段二之觀察結果

此研究設計中，為什麼需要“run-in period”，其原因在於因為採用交叉設計通常樣本數都比較少，如果找到一群合作意願較不佳之個案，試驗參與者在接受階段一的治療後就失聯 (loss follow-up)，將會造成第二階段皆收集不到結果，如此一來，原來的交叉設計就變成平行設計; 但原先樣本數的計算是基於交叉設計進行計算，所以對於平行設計而言就會有樣本數不足的問題發生，而在分析上產生統計檢定力不足的情形。因此可以透過run-in period去檢視且尋找一群合作意願較高(compliance rate)的個案，來達到收集完整兩階段之目的。

臨床實例 1

有一研究以隨機分派試驗設計探討 Nurse navigator 對於癌症病患之就醫過程與生活品質是否有所助益。罹患大腸直腸癌、乳癌或肺癌病患以及診治醫師為單位隨機分派到試驗組(Nurse navigator)介入或對照組(加強護理照護)。(Wagner EH et al.: Nurse navigator in early cancer: a randomized, controlled trial. Journal of clinical oncology. 2013; 25:1-8)。研究流程見原文圖 1。

分析方法: 以線性迴歸調整病患生活品質之基礎值後比較生活品質以及就醫過程經驗分數之差異。

研究結果: 相較於加強護理照護，Nurse navigator 對於生活品質並沒有影響，但對於病患在就醫過程之疑問以及社會心理相關問題則有顯著的幫助。

臨床實例 2

一研究欲探討口服嗎啡對於是否對於病患呼吸困難的感覺有所影響(AP Albernethy et al.: Randomised, double blind, placebo controlled crossover trial of sustained release morphine for the management of refractory dyspnoea. BMJ 2003; 327: 523-8). 研究流程見原文：

上述研究是採用交叉研究設計(crossover)，其主要結果為利用視覺等效量尺評分法(visual analogue scale,VAS)量測呼吸困難程度，研究結果見原文表 2。

分析方法則採用配對 t 檢定(Paired *t*-test)來比較兩組對於呼吸困難的程度(以 VAS 測量)是否不同:

$$t = \frac{\bar{X}_d - \mu_d}{s/\sqrt{n}} = \frac{6.6}{15/\sqrt{38}}$$

結論: 嗎啡對於呼吸困難之程度(以視覺等校量表量測,VAS)具有顯著降低

之效益。

該研究另採用 McNemar Test 來比較兩組對於呼吸困難改善的比率(以改善與否測量)是否不同，結果見原文表 3。

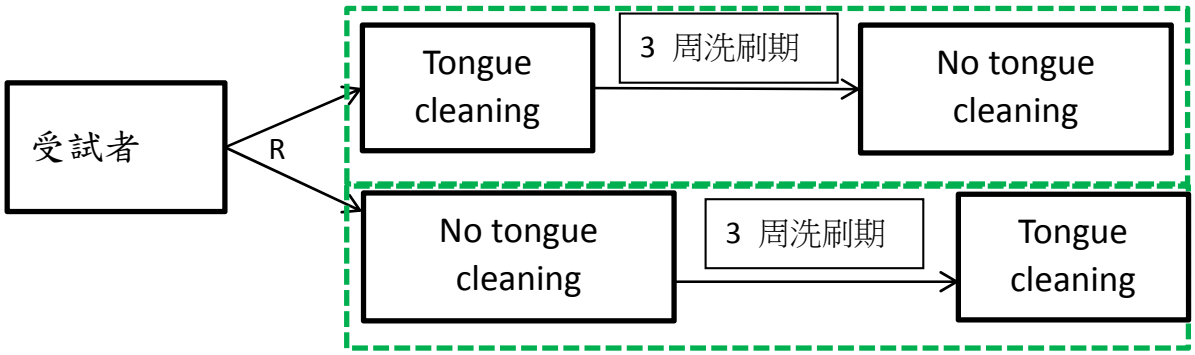
$$\chi^2 = \frac{(|8-1|-0.5)^2}{8+1} = 4.69 > \chi_1^2 = 3.84$$

結論: 嗎啡治療對於呼吸困難改善之比率顯著較高。

臨床實例 3

有一研究欲探討舌部清潔(tongue cleaning)對於口腔細菌含量以及牙菌斑是否有所影響。研究採用單盲隨機分派對照設計以及交叉試驗設計 (crossover design)。(Matsui et al.: Effect of tongue cleaning on bacterial flora in tongue coating and dental plaque: a crossover study. BMC Oral Health 2014(**14**):4)

研究流程如下圖：



結果量測:

舌部菌量(Winkel tongue coating index, WTCI)

牙菌斑內菌量

研究結果:

舌部菌量: 見原文圖 1

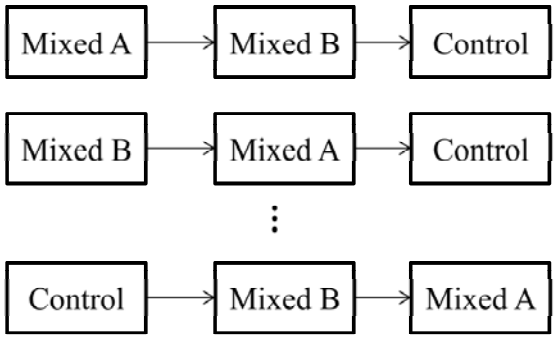
牙菌斑內菌量: 見原文圖 2

結論:

a. 舌部清潔對於舌部菌量在第 3 天時具有減少之效益
b. 舌部清潔對於牙菌斑菌量減少並無效益

臨床實例 4

有一研究欲探討色素添加劑對於兒童產生過動症狀(hyperactive behavior)是否有所影響。該研究使用雙盲隨機分派對照試驗及交叉(crossover)設計。受試者分別接受含有色素添加劑以及安慰劑之混合物(mix A 以及 mix B)，結果評估為發生過動症狀之評分指標 (global hyperactive aggregate (GHA)) (McCann et al.: Food additives and hyperactive behavior in 3-year-old and 8/9-year-old children in the community: a randomized, double blinded. Placebo-controlled trial. Lancet 2007 61306-3.)



分析結果: 見原文表 3

結論: 因為迴歸係數為 0.2>0 (H_0:β=0)，且 95%信賴區間不包含 0，顯示色素添加劑 A 對於 3 歲兒童之過動狀況(以 GHA 分數評估)具有加劇之效果。

2. 多因子實驗設計(Factorial Design)

當研究者在同一實驗想要操控兩個以上之變項，就必須使用多因子實驗設計(Factorial Design)。而在多因子實驗設計中研究者仍然依照隨機分配原則將所有可能組合之介入分派至各個受試者。

(1) 完全隨機分派之多因子研究設計

例如有 2 個介入變項 A 及 B，其中 A 有 2 個 level 而 B 有 3 個 level，則所有可能的組合共 6 種，若依完全隨機化設計(CRD)為：

$$R \begin{cases} X_{A1} \; X_{B1} \; O_1 \\ X_{A1} \; X_{B2} \; O_2 \\ X_{A1} \; X_{B3} \; O_3 \\ X_{A2} \; X_{B1} \; O_4 \\ X_{A2} \; X_{B2} \; O_5 \\ X_{A2} \; X_{B3} \; O_6 \end{cases}$$

以上可以用列聯表的方式表示如下：

X_A		X_B		
		X_{B1}	X_{B2}	X_{B3}
	X_{A1}	O_1	O_2	O_3
	X_{A2}	O_4	O_5	O_6

利用多因子研究設計，研究者不但可以測試A及B之主效應(main effect)，而且可以檢視 A 因子與 B 因此之間是否有交互作用(interaction)存在。例如我們可以比較:採用 A1 者，其結果是否在 B1、B2 與 B3 有不同? 如有不同則代表 A1 因子會因為 B 因子是哪一種而有不同的反應;故 B 因子為 A 因子之修飾因子(effect modifier) ，及兩者有交互作用存在。

臨床實例 5

有一研究欲探討抗病毒藥物(acyclovir)以及類固醇對於顏面神經麻痺之治療效果是否有所助益。該研究採隨機分派對照試驗及多因子研究設計(抗病毒藥物(acyclovir)以及類固醇(prednisolone)二因子)方式進行。

研究流程見原文圖 1:

研究結果: 見原文表 2

結論:

a. 抗病毒藥物與類固醇治療對於顏面神經麻痺之改善效果無交互作用。

b. 早期使用類固醇治療對於顏面神經麻痺改善有所助益。

c. 早期使用抗病毒藥物對於顏面神經麻痺改善並無助益。

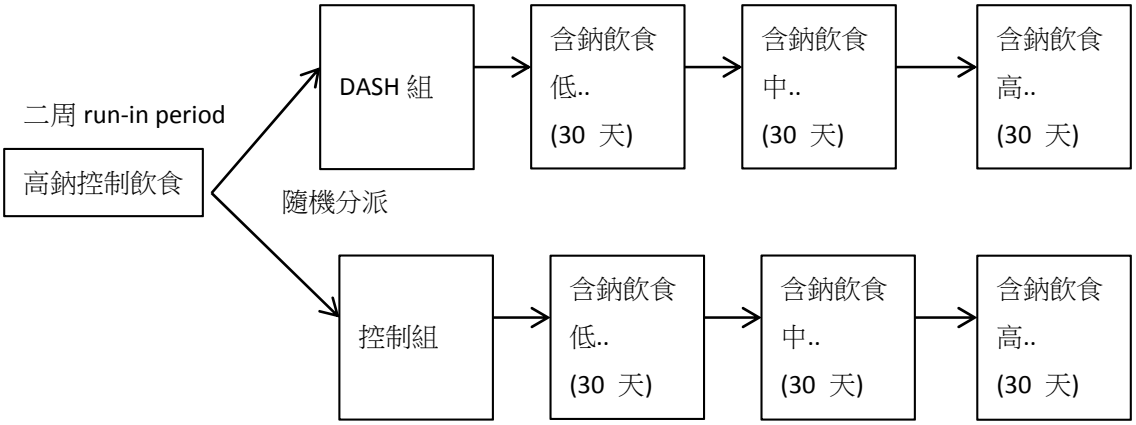
(3) 複雜試驗設計 (Complex experimental design)

1. 平行設計與交叉設計結合多因子設計 (the combined parallel and cross-over factorial design)

臨床實例 6

一研究欲評估飲食中含鈉量與預防高血壓之高蔬果低脂飲食(Dietary Approaches to Stop Hypertension (DASH) diet)對降血壓之效果。介入有兩種飲食(DASH diet 與一般飲食)以及三種不同的含鈉量(高、中與低)。經 2 周的 run-in period 後，實驗將通過 run-in period 的研究個案利用平行設計概念，隨機分派到兩種飲食的其中一種。隨後再利用交叉設計之概念，每個個案接受三種不同含鈉量飲食各 30 天，而三種含鈉量飲食之順序為隨機指派。研究最主要之結果為飲食介入後之 30 天時所測得之收縮壓。(FM Sacks et al.: Effects on blood pressure of reduced dietary sodium and the dietary approaches to stop hypertension (DASH) diet. NEJM 2001;344:3-10).

上述研究方法採用之**研究設計**，圖示說明如下:



*三個階段之飲食為高、中或低鈉之順序為隨機決定。

DASH 組與一般飲食組之隨機分配是運用平行設計但鈉濃度的分配則採用組內設計(crossover)；因為有兩個因子(飲食以及鈉濃度)，所以該研究設計為多因子研究設計(factorial design)。

研究結果:見原文圖 1。

結論:

a. 低鈉飲食、DASH 飲食分別對於血壓之降低皆具有顯著效果

b. 低鈉飲食加上 DASH 飲食對於血壓降低之效果具有交互作用，DASH 飲食在高鈉飲食組對於血壓降低之效果較為顯著，但在低鈉飲食組其效果較不顯著。

(2) Split-unit design

在某些實驗性研究設計研究者欲探討 2 個因子之多因子試驗設計，而其中一個因子是另一個因子之次單位(subunit)。研究者首先將大單位因子視為 ”Main plots” 在每個 main plot 再將其分為一系列 ”Subplots” ，而隨機分派也是先執行 k 個和 main plot 相關之介入，在每個 main plot 中再執行和 subplot 相關之隨機分派。這種研究設計稱為"Split-plot design"，最先應用於農業之實驗設計。例如: Main plot：土壤條件、Subplot：不同的肥料。

此概念可應用於醫學之相關例子如下:

Main plot	Subplot
受試者	相同人(動物)不同時間
同一窩	飼養於同一窩之動物
天	每天之不同週期

以研究藥物在每天不同時間對於受試者之反應為例:

受試個案	A 因子 “ 藥物” (main plot)	B 因子 “時間” (subplot)		
		1	2	3
1	A3	B1	B3	B2
2	A1	B1	B2	B3
3	A1	B3	B1	B2
4	A2	B1	B3	B2
5	A3	B1	B2	B3
6	A2	B3	B1	B2
7	A1	B1	B3	B2
8	A3	B1	B2	B3
9	A3	B3	B1	B2
10	A2	B1	B3	B2
11	A1	B1	B2	B3
12	A2	B3	B1	B2

有關上述複雜的研究設計之統計方法如有興趣可參考隨附文章之方法描述以及研究所參考書(P Armitage, G Berry, JNS Matthews: Statistic Method in Medical Research. Ch 9. pp.266, Ch 12.5 pp.418-424)

(4) 相依資料結構與統計分析方法整理對照表

結果變項(Y)性質	樣本統計值平均值	結果評估	非迴歸分析方法	迴歸分析方法
(1)連續變項	$\bar{Y}$(平均值)	$\overline{Y_1} - \overline{Y_0}$ 或斜率 β	配對 t-test	混合線性迴歸模式 (linear mixed model) $Y_i = \alpha_i + \beta X_i + \varepsilon_i$ *i:* i_{th} block
(2)二元變項 (Y=1/0)	$\hat{p}$(比例)=d/n d: 事件 n: 總數	勝算比(OR) 或斜率 β	2×2 列聯表 McNemar test	條件式羅吉斯迴歸 (conditional logistic regression) **於後續課程說明*

流行病學方法論及實驗
(The Methods of Epidemiology and Practices)

賽先生護理流行病學

授課教師：陳秀熙 教授/許辰陽 博士

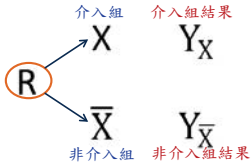
台灣大學流行病學與預防醫學研究所

e-book 01: p25, p48-52
e-book 02:Ch2_ p.28-31)

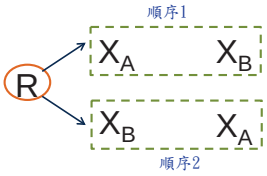
一、隨機分派對照研究之觀念

X(因) → Y(果)

- (1) 平行設計



- (2)交叉設計



"R" 代表randomization 或 random allocation

非採用隨機分派對照試驗

- 如觀察性研究, 謬誤(fallacy)產生
  - (1)干擾因子(辛普森矛盾**Simpson's paradox**)
  - (2)選擇偏差(Selection Bias)
  - (3)安慰劑(Placebo effect)

安慰劑(Placebo effect)

- 在芝加哥, Hawthorne工廠的老闆發現工廠產量不好，員工反應是因為照明設施不好，因此他改進了照明設施，因而增加了產量。過了一段時間照明設備又變不好，但產量並沒有減少。
  - 論點：產量的增加可能是因為老闆聽取員工意見進而鼓勵員工，員工心情變好而更認真工作，進而產量增加，所以即使後來照明設備不好但產量仍是穩定的。
  - **霍桑效應(Hawthorne Effect)**
  - 欲解決Hawthorne效應的方法即是進行RCT，並且在對照組使用安慰劑，故後來稱做placebo effect

e-book 01: p25, p48-52
e-book 02:Ch2_ p.28-31)

分析方法

- 平行設計又因有無前測分為下列兩種:
  - (1) 有前測的完全隨機化設計

$$R\begin{cases} O\ X\ O \\ O\ \bar{X}\ O \end{cases}$$

X(介入組，參加減重計畫)	$\bar{X}$(對照組，未參加減重計畫)	註: O_1、O_3: 前測 O_2、O_4: 後測
O_1 O_2	O_3 O_4	

  - (2)沒有前測的完全隨機化設計

$$R\begin{cases} X & O_2 \\ \bar{X} & O_4 \end{cases}$$

•只要比較後測之結果(O_2及O_4)

統計分析之概念

	結果變項(Y)性質	樣本統計值平均值	非迴歸分析方法	迴歸分析方法
二組 (X=1/0)	(1)連續變項	$\bar{Y}$(平均值)	獨立 t-test $(\bar{Y}_1 - \bar{Y}_0)$	線性迴歸 Y=a+bx
	(2)二元變項 (Y=1/0)	$\hat{p}$(比例)=d/n d: 事件 n: 總數	X^2 test(2×2 列聯表)及比例檢定 $\hat{p}_1 - \hat{p}_0$	羅吉斯迴歸 Logit(P(Y=1|X))=a+bx
	(3)存活時間 (Y=time to event)	存活函數 S(t) 風險函數 h(t)	Hazard Rate $=h_1(t)/\ h_0(t)$	Cox 迴歸模式(Cox proportional hazards regression model) $h(t)=h_0(t)\exp(a+bx)$
多組 (k=3) X=1/2/3	(1)連續變項	$\bar{Y}_1 - \bar{Y}_2$)/ $\bar{Y}_2 - \bar{Y}_3$)/ $\bar{Y}_1 - \bar{Y}_3$)	One-way Analysis of Variance(ANOVA)	$Y=a+b_1X_1+b_2X_2$ X=1 X=2 X=3 X_1 0 1 0 X_2 0 0 1

迴歸模式之想法

$Y=a+bx+e \quad e\sim N(0,\sigma^2)$

$X=1 \quad \bar{y}_1 = a+b$

$X=0 \quad \bar{y}_0 = a$

$\bar{y}_1 - \bar{y}_0 = b$ (迴歸係數也就是斜率), b 越大, 則兩組平均值差異越大

e-book 01: p25, p48-52)

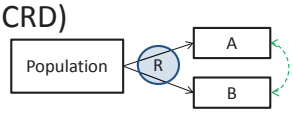
實證醫學 (Evidence-Based Medicine, EBM)

- 為了使因果關係不受偏差的影響，西方醫學將解決偏差的研究設計稱為"實證醫學 (Evidence-Based Medicine, EBM)
  - **(1) 隨機分派對照試驗(randomized trial, RCT)**
    - 移除干擾因子及偏差(bias)
  - **(2) 前瞻性研究(prospective study)**
    - 因在果之前,無進行隨機分派
    - 控制干擾因子及偏差
  - **(3) 病例對照研究(case-control study)**
    - 因為事件很少，需要追蹤的人數眾多故進行此研究
    - 控制干擾因子及偏差
  - **(4) 橫斷性研究(cross-sectional study)**
    - 因果都在同一時間點做測量
  - **整合分析(Meta-analysis)**
  - **成本效益分析(Cost-Effectiveness Analysis)**

7

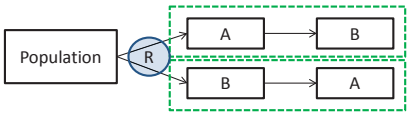
二、試驗設計 (Study design)

- 平行試驗設計 (completely randomized design, CRD)



獨立樣本, uncorrelated data

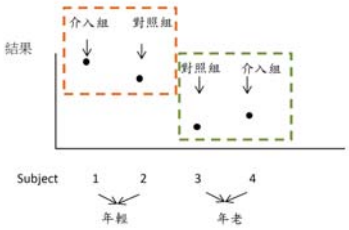
- 隨機區集試驗設計(randomized block design, RBD)

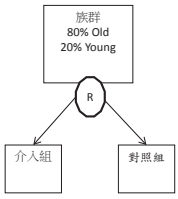


區集(Block)相依樣本 (Correlated data)

8

1. 使用隨機區集設計的緣由





9

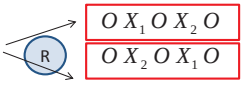
隨機區集設計 (Randomized Block Design, RBD)

- 隨機分派之單位為區集 (block)
  - 一個家庭 (family)
  - 配對之受試者 (matched subjects)
  - 同一個體 (individual)
- 區集內具有同質性 (homogenous)
- 區集間維持異質性 (heterogeneous)

10

維持區集內同質(homogenous)的方法

1. 配對 (Matching)
   - 將特質相近的受試者進行配對(性別、年齡…)
   - 每一個配對組形成一個區集
   - 同一區集內的受試者隨機分派到治療組或對照組
2. 重複測量 (Repeat measurement)
   - 每一個受試者接受多種介入其順序為隨機指派, 若介入有2種:



- 優點: 更有效的偵測介入的效益(efficient)
- 缺點: Carry-over effect

11

例子 1: 體適能訓練計畫

- 某試驗欲探討訓練計劃對於體適能之增進是否有所助益
  - 已知年齡對於體適能之優劣有相當程度之影響
  - 利用隨機區集設計，以年齡作為配對條件(matching)建立區集 (block)
    - 同一年齡組的受試者在隨機分派到介入組或控制組

12

例子 2: 交叉試驗設計(crossover)

- 交叉試驗設計 (crossover): 一個受試者接受多種介入，一個受試者即為一個區集 (block)，以兩種介入為例

	Run-in	Period 1	Washout	Period 2
Outcome	Z_{ij}	O_{ij1}		O_{ij2}
Group Ⅰ(n_1)	---	A	---	B
Group Ⅱ(n_2)	---	B	---	A

附註說明 i: group, j: subject
A: 治療組 A, B: 治療組 B, n_1:第一組樣本數, n_2:第二組樣本數
Z_{ij}: run-in period 觀察值
O_{ij1}: group i, subject j, 階段一之觀察結果
O_{ij2}: group i, subject j, 階段二之觀察結果

臨床實例 1

J Clin Oncol 31. © 2013 by American Society of Clinical Oncology

Nurse Navigators in Early Cancer Care: A Randomize Controlled Trial

Edward H. Wagner, Evette J. Ludman, Erin J. Aiello Bowles, Robert Penfold, Robert J. Reid, Carolyn M. Rutter, Jessica Chubak, and Ruth McCorkle

臨床實例 2
嗎啡是否對於呼吸困難之緩解有所助益?

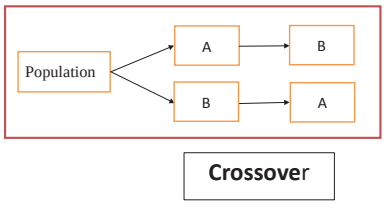
呼吸困難

- 呼吸困難的治療在緩和醫療照顧上是重要的問題
- 生理層面與心理層面的因素
- 憂鬱、焦慮、失眠多重因素交互影響

應用鴉片類藥物治療呼吸困難

- 缺乏一致的共識
  - 對於呼吸困難之症狀緩解可能有所助益
  - 對於慢性阻塞性肺疾可能加劇病情
  - 具有呼吸抑制之效果並造成二氧化碳滯留
- 缺少高品質的臨床試驗

以雙盲性之隨機分配對照試驗評估嗎啡對於呼吸困難之治療效益:
交叉試驗設計
(Randomized, double blind, placebo controlled crossover trial of sustained release morphine for the management of refractory dyspnea)



AP Albernethy et al.: Randomised, double blind, placebo controlled crossover trial of sustained release morphine for the management of refractory dyspnoea. BMJ 2003; 327: 523-8

臨床實例 2: 研究目的與試驗設計

- 研究目的 (Objective)
  - 評估嗎啡對於基本病因 (underlying aetiology)已進行最佳治療之患者其呼吸困難之程度 (sensation of breathlessness)是否具有緩解之效益
- 試驗設計 (Design)
  - 隨機分派、雙盲、對照組使用安慰劑
  - **交叉試驗設計 (crossover design)**

AP Albernethy et al.: Randomised, double blind, placebo controlled crossover trial of sustained release morphine for the management of refractory dyspnoea. BMJ 2003; 327: 523-8

臨床實例 2: 受試者與結果量測

- 受試者參與條件 (Participant)
  - 收案48人,未曾接受過嗎啡治療, 38 人完成
  - 多為慢性阻塞性肺疾患者 (CODP, 88%)
  - 介入: 4天 20mg 長效型嗎啡緩釋劑
  - 對照: 4天安慰劑
- 結果量測 (Outcome measurement)
  - 主要結果: 呼吸困難程度 (100mm 視覺等校量尺, visual analogue scale, VAS)
  - 其他結果: 睡眠品質、自覺症狀改善、體能狀況、治療副作用

AP Albernethy et al.: Randomised, double blind, placebo controlled crossover trial of sustained release morphine for the management of refractory dyspnoea. BMJ 2003; 327: 523-8

臨床實例 2: 樣本數估計

Sample size

We calculate si
reports of im ly
and local expectations for data variability in this cross-over study; the predicted standard deviation of the visual analogue scale was 16 mm.[16] We estimated that 48 participants would provide 80% power to detect a 10 mm difference in the scale, with an α of 0.05, allowing for a 20% dropout rate.

樣本數公式:

$$n = \frac{(Z_{1-\beta} + Z_{1-\alpha})^2}{\left(\frac{\mu_0 - \mu_1}{\sigma}\right)^2}$$

樣本數決定於四個成分:
(A) 統計檢定力, 1-β
(B)顯著水準, α
(C) H_0與 H_1之真實差異, μ_0-μ_1
(D) 母群體變異數, σ^2

- **虛無假說**: morphine與安慰劑的效果相同。
- **治療效益**: VAS 相差 10 mm
- **Significant level(型一錯誤)**: 設定最大可容忍型一錯誤為5%。
- **統計檢定力(1-型二錯誤)**:80%

20

臨床實例 2
結果: 呼吸困難改善程度 (VAS量測)

- **Paired t-test**
- H_0: mean improvement =0

(同一個人使用morphine與使用placebo相比,其VAS改善為0)

$$\left(S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_0^2}{n_1+n_2-2}\right)$$

$$t = \frac{\bar{X}_d - \mu_d}{s_p/\sqrt{n}} = \frac{6.6}{15/\sqrt{38}}$$

$$95\%CI = \left(6.6 - t_{37,0.975} \times \frac{15}{\sqrt{38}}, 6.6 + t_{37,0.975} \times \frac{15}{\sqrt{38}}\right)$$

自由度(degree of freedom)

21

臨床實例 2 結論

1. **本研究使用何種試驗設計?**

- 隨機區集設計 (randomized block design, RBD), 交叉試驗設計 (crossover design)

2. **主要結果量測為何?使用何種統計方法檢定?**

(1) 呼吸困難視覺等效量尺 (VAS)

配對t檢定 (paired t-test)進行兩組比較

$$t = \frac{\bar{X}_d - \mu_d}{s/\sqrt{n}}$$

(2) McNemar's test (二元變項)

$$\chi^2 = \frac{(|b-c|-0.5)^2}{b+c}$$

3. **結論:** 嗎啡對於呼吸困難之程度(以視覺等效量表量測,VAS)具有顯著降低之效益。

22

臨床實例 3
舌部清潔(tongue cleaning)對於口腔細菌含量以及牙菌斑是否有所影響?

研究設計:

結果量測:

- 舌部菌量(Winkel tongue coating index, WTCI)
- 牙菌斑內菌量

Matsui et al.: Effect of tongue cleaning on bacterial flora in tongue coating and dental plaque: a crossover study. BMC Oral Health 2014

24

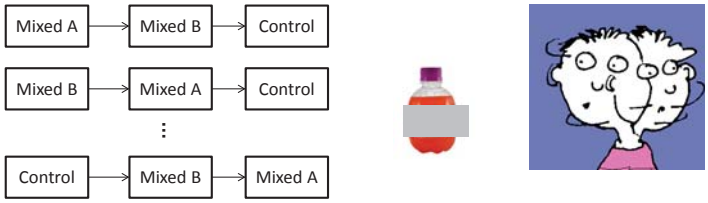
臨床實例 3 分析方法與結論

結論:

a. 舌部清潔對於舌部菌量在第3天時具有減少之效益

b. 舌部清潔對於牙菌斑菌量減少並無效益

Food additives and hyperactive behaviour in 3-year-old and 8/9-year-old children in the community: a randomised, double-blinded, placebo-controlled trial



Linear Mixed Regression Model

結論:
因為迴歸係數為0.2>0 (H_0:β=0)，且95%信賴區間不包含0，顯示色素添加劑A對於3歲兒童之過動狀況(以GHA分數評估)具有加劇之效果。

Linear mixed re

$$Y_{ij} = \alpha_i + \beta X_{ij} + \varepsilon_{ij}$$

$$\alpha_i = \alpha_0 + \xi_i$$

$$\xi_i \sim N(0,\sigma_\alpha^2), \varepsilon_{ij} \sim N(0,\sigma^2)$$

Mixed model

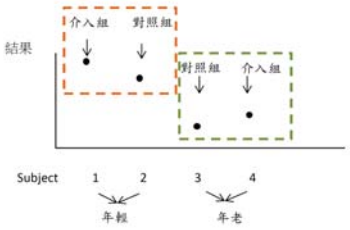
Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	17.5000	1.4532	9	12.04	<.0001
X	-2.7000	1.1552	9	-2.34	0.0442

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
X	1	9	5.46	0.0442

$$t = -2.34 = \frac{\bar{X}_d - \mu_d}{s/\sqrt{n}} = \frac{-2.7}{3.6530/\sqrt{10}}$$




```
run;
```

Paired t-test

The TTEST Procedure

Difference: drug - placebo

N	Mean	Std Dev	Std Err	Minimum	Maximum
10	-2.7000	3.6530	1.1552	-8.0000	4.0000

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-2.7000	-5.3132	-0.0868	3.6530	2.5127	6.6890

DF	t Value	Pr > |t|
9	-2.34	0.0442

2. 多因子試驗設計 (Factorial Design)

- 當研究者在同一實驗想要操控兩個以上之變項，就必須使用多因子實驗設計(Factorial Design)
- 依照隨機分配原則將所有可能組合之介入分派至各個受試者。

2.多因子試驗設計 (Factorial Design)

1.<u>完全隨機分派之多因子研究設計</u>

– 例如有2個介入變項A及B，其中A有2個level而B有3個level，則所有可能的組合共6種，若依完全隨機化設計(CRD)為：

X_A		X_B		
		X_{B1}	X_{B2}	X_{B3}
X_{A1}		O_1	O_2	O_3
X_{A2}		O_4	O_5	O_6

R
$X_{A1}\ X_{B1}\ O_1$
$X_{A1}\ X_{B2}\ O_2$
$X_{A1}\ X_{B3}\ O_3$
$X_{A2}\ X_{B1}\ O_4$
$X_{A2}\ X_{B2}\ O_5$
$X_{A2}\ X_{B3}\ O_6$

– 研究者不但可以測試A及B之主效應(main effect)，而且可以檢視A因子與B因此之間是否有**交互作用**(interaction)存在

臨床實例 5.

抗病毒藥物與類固醇對於治療顏面神經麻痺之角色

- 抗病毒藥物(acyclovir)以及類固醇對於顏面神經麻痺之治療效果是否有所助益。
- 隨機分派對照試驗及多因子研究設計

臨床實例 5: 結論

研究結果

1. 抗病毒藥物與類固醇治療對於顏面神經麻痺之改善效果無交互作用。
2. 早期使用類固醇治療對於顏面神經麻痺改善有所助益。
3. 早期使用抗病毒藥物對於顏面神經麻痺改善並無助益。

There was no significant interaction between prednisolone and acyclovir at either 3 months or 9 months (P=0.32 and P=0.72, respectively). Ta-

3. 複雜試驗設計 (complex experimental design)

1. 結合平行設計與交叉設計之多因子試驗 (the combined parallel and cross-over factorial design)

2. Split unit design

臨床實例 6.

- DASH study:評估飲食中含鈉量與預防高血壓之高蔬果低脂飲食(**Dietary Approaches to Stop Hypertension (DASH**) diet)對降血壓之效果

介入:

- 兩種飲食 (DASH 以及一般飲食)
- 三種不同的含鈉量 (高、中、低)

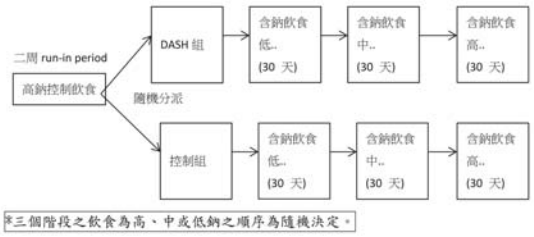
- 2周的run-in period後, 受試者利用平行設計 (parallel study design) ，隨機分派到兩種飲食的其中一種.
- 隨後再利用交叉設計 (crossover design)，每個個案接受三種不同含鈉量飲食各30天.

主要測量結果:飲食介入後之30天時所測得之收縮壓

FM Sacks et al.: Effects on blood pressure of reduced dietary sodium and the dietary approaches to stop hypertension (DASH) die NEJM 2001;344:3-10

臨床實例 6: 研究設計

- 多因子設計結合隨機區集化試驗設計方法 (Factorial design with RBD (crossover))



臨床實例 6: 研究結果

- 低鈉飲食、DASH飲食分別對於血壓之降低皆具有顯著效果
- 低鈉飲食加上DASH飲食對於血壓降低之效果具有交互作用，DASH飲食在高鈉飲食組對於血壓降低之效果較為顯著，但在低鈉飲食組其效果較不顯著。

Split-unit design

- 在某些實驗性研究設計研究者欲探討2個因子之多因子試驗設計，而其中一個因子是另一個因子之次單位(subunit)。
- 研究者首先將大單位因子視為 ”Main plots” 在每個main plot再將其分為一系列 ”Subplots” 。
- 而隨機分派也是先執行*k*個和 main plot 相關之介入，在每個 main plot 中再執行和 subplot 相關之隨機分派。

Split-unit design: 運用

- Split-unit design 始於農業研究:
  - Main plots: 不同條件之土壤 (日照、水源…)
  - Subplots: 不同的肥料
- 運用於生物醫學研究中:

Main plot	Subplot
受試者(Subject)	同一受試者不同的介入時期(Period of subjects (crossover))
同一窩 (litter)	飼養於同一窩之動物 (Study animal within litter)
試驗天 (Day)	每天中不同的週期 (Period of the day)

Split-unit design: 生物醫學運用

- 以研究藥物在每天不同時間對於受試者之反應為例:

Subject	A factor " Drug" **(main plot)**	B factor "Drug during period" **(subplot)**		
		1	2	3
1	A3	B1	B3	B2
2	A1	B1	B2	B3
3	A1	B3	B1	B2
4	A2	B1	B3	B2
5	A3	B1	B2	B3
6	A2	B3	B1	B2
7	A1	B1	B3	B2
8	A3	B1	B2	B3
9	A3	B3	B1	B2
10	A2	B1	B3	B2
11	A1	B1	B2	B3
12	A2	B3	B1	B2

三、 相依資料結構與統計分析方法整理對照表

結果變項(Y)性質	樣本統計值平均值	結果評估	非迴歸分析方法	迴歸分析方法
(1)連續變項	$\bar{Y}$(平均值)	$\bar{Y}_1 - \bar{Y}_0$ 或斜率 β	配對 t-test	混合線性迴歸模式 (linear mixed model) $Y_i = \alpha_i + \beta X_i + \varepsilon_i$ i: i_{th} block
(2)二元變項 (Y=1/0)	$\hat{p}$(比例)=d/n d: 事件 n: 總數	勝算比(OR) 或斜率 β	2×2 列聯表 McNemar test	條件式羅吉斯迴歸 (conditional logistic regression) *於後續課程說明

九、 信效度護理流行病學 (1)

效度(Validity)及信度(Reliability)

一、效度(Validity)之概念

Cook 及 Campell 兩人認為效度在研究法之概念為“the best available approximation to the truth or falsity of propositions， including propositions about cause”，受到 Popper 否證論之影響，他們認為永遠不可能知道真實狀態為何，最多只知何者無法否證。因此儘量使用‘approximately’或‘tentatively’。

Campell 和 Stanely 於 1963 年提出兩種效度：

1、 Internal Validity:

It is defined as the approximate validity with which we infer that a relationship between two variables is causal or that the absence of a relationship implies the absence of causes。

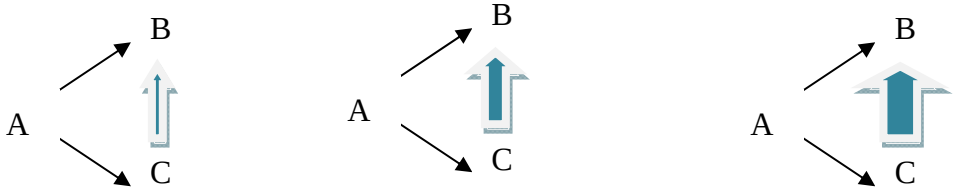
2、 External Validity:

It refers to the approximate validity with which we can infer that the presumed causal relationship can be generalized to and across alternate measures of the cause and effect and across different types of persons, selling, and items.

二、Internal Validity

1、 內在效度之簡介

基本上就內在效度在因果關係之扮演上最重要的是在因(A)→果(B)過程中如何受到第三者(C)影響評估相當重要且不容易，而且其干擾方向性也相當困難，一般而言 C 通常接 A→B 使其減弱或消失，然而也可能 C 出現 A→B 才出現，以下整理其可能出現



2、 影響內在效度之因素

Campell 將影響內在效度歸為幾類：

(1). History

History 通常是發生在前測及後測中間受到某些其他因素之影響， 一般而言這種效度在 Laboratory 比較不可能發生，因為實驗室研究可以隔絕外在之影響，而在一般社會及行為科學研究比較不可能是 closed system，因此受到 History 之影響可能性相當高。

例：某衛生教育設計(E)提高子宮頸抹片篩檢率，針對某社區一群進行 E 衛生教育介入，然後比較介入前及介入後其子宮頸抹片篩檢率，結果在介入前後之間受到“電視廣告六分鐘護一生”廣告影響(History)，因此到底是 History 之效應或真正有效。

(2). Maturation

任何一種前測及後測其結果改變可能會因為受訪者年長，變聰明，有經驗而不是因為介入之影響，稱為 Maturation。

例：某研究欲探討是否 IGT(Impaired Glucose Tolerance)可以經由某種降低血糖之飲食控制達成，研究者於是選擇一群 IGT 病人在前測先測 fasting blood sugar 再施予飲食控制，發現血糖恢復正常→可能是 maturation。

(3). Testing

在前測及後測有些結果是以某些 test(如問卷)來得到，若結果有變化有些時候並非是 intervention 影響，而是受測者對於 test 已有經驗，有所謂 test effect。

(4). Instrumentation

前測及後測之測量工具可能改變，因此結果改變並非是因為 Intervention，這種工具改變可能是因為觀察者變得有經驗或可能是尺度改變，後者一般稱為 ceiling effect，也就是末端 scale 之改變遠比中點來得不容易改變。

(5). Statistical Regression toward the mean

一般而言在測量不是很穩定情況下，容易產生 pretest 分數很高或很低，則在經過後測之後很容易產生較低或較高後測分數，這種現象稱為 regression toward the mean，這種改變很容易干擾介入之效應。一般而言 regression 是朝向 population mean，而這種誤差大小視 test-retest reliability of a measure，以及 subgroup mean 和 population mean 之差異。一般而言 statistical regression 可以用貝氏定理來解釋，如果 population mean＝μ，表示 prior＝μ，而 pre-test 所得到的 score X 視為樣本，而 post-test 所得到的 score 為 X_2，如何能夠調整 regression toward the mean 再和 X_2 比較，其做法如下：

假設 score 為 normal distribution，則透過貝氏定理

Posterior α Prior $\times$ likelihood

如此 Posterior 之 distribution 為(μ_p，σ_p^2).

$\mu_p = (\mu * 1/\sigma^2 + x * 1/\sigma_{x1}^2)/(1/\sigma^2 + 1/\sigma_{x1}^2)$

$\sigma_p^2 = \sigma^2 + \sigma_{x1}^2$

如此比較μ_p與後測之結果，可以得知在調整了 regression toward the mean 是否 intervention 有效。

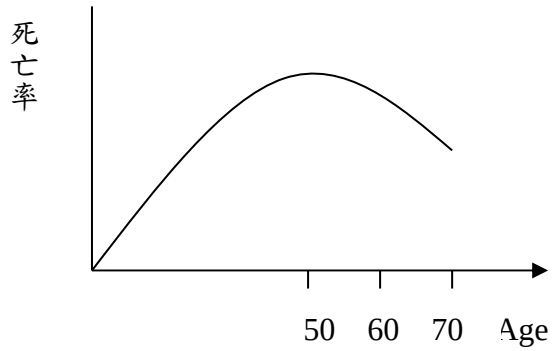
(6). Selection

如果結果之改變是因為兩組之間某些特性不同 (例如：年齡，關心自己健康程度)，而這些特性不同在兩組所造成結果不同，會產生所謂 selection bias。

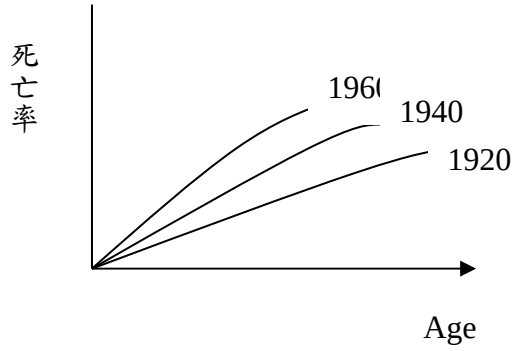
例：Type 2 DM screening 中吾人發現 Attender 與 Non-attender 經過十年後累積死亡率分別為 $50 / 10^3$ 及 $50 / 10^3$，其 ratio Non-attender 是 Attender 的 3 倍，因此下結論說

screening 有益，此差異很可能是 Non-attender 比較不關心自己健康，生活型態也不好，造成死亡機率高於 Attender，而非實驗 screening benefit 造成，此可以比較兩組間 total mortality 而得到佐證。

下列是某大腸直腸癌症年齡別死亡率當代曲線(1990)



吾人發現隨著 Age 增加，死亡率呈現下降不太合理，若改成出生世代曲線則隨著年齡上升而上升相當合理



這是流行病學標準世代效應(Cohort effect)，而這種 Cohort effect 在方法學上相當於 Maturation，也就是早出生世代死亡率較低而晚出生世代死亡率較高。

(7). Mortality

如果任何結果改變是因為兩組之間 drop-out 不同，而這些 drop-out 影響結果在兩組也不同，這是因為 loss to follow-up。

例：在探討降血壓新藥是否能夠降低 borderline hypertension，研究者分別尋找兩群人

測量用藥前血壓，隨意分一組接受新藥 A，另一組接受舊藥 B，經過一段時間後，再測量其血壓,如果兩組皆有 40% drop-out,而且在 A 組 drop-out 之對象其 social class 較 B 組高，如果 social class 高之對象血壓皆較高，則可能造成 A 組血壓降低程度較 B 組多，這可能不是新藥的效果，而是 loss to follow-up 在兩組所造成之效應不同。

(8). Interaction with selection

上述這些影響 validity 之因素皆會和 selection 共同對結果產生影響，包括：

a. Selection-Maturation

某衛生教育介入計畫欲改善婦女對於乳癌認知之影響，研究人員找尋兩群人，一組是 high social-class (A) ，另一組是 low social-class (B) ，經過一段時間後發現 A 組較 B 組影響為大，這可能有兩個因素，第一是其 social class 之選擇偏差，第二則為隨著時間增加可能認知能力就會增加，而其增加在 high social-class 組增加更大，也就是 social class 和 maturation 產生 interaction。

b. Selection-history

相同道理如果類似上述以子宮頸癌篩檢為例,如果選擇 SES 較高及 SES 較低者,SES 較高則受到〝廣告〞干擾可能比 SES 較低者來得大，這是 SES 之 selection 和 history 產生 interaction 結果。

c. Selection-instrumentation

假如所選擇的兩組，其中一組 scale 和另一組有選擇性偏差，也就是其中一組是在較高〝ceiling〞或較低〝floor〞位置，則有 ceiling effect 及 floor effect，這是經過 selection 之後，再和 instrumentation 產生 interaction 造成。

(9). 因果關係時序性

這是一般 cross-sectional study 中常見的 fallacy，因為在 cross-sectional study 因與果同時測量，因此無法得知因果關係的時序性。

例：探討 Triglyceride (TG)和 Type 2 DM 之關係，若使用 cross-sectional study 時，因

為 Type 2 DM 是以 fasting blood sugar ≧126 來定義，如此到底是 TG 高 → fasting blood sugar 高，還是 fasting blood sugar →TG 高，無法釐清。一般而言，在這種情況下研究者會以 correlation 來表示，不須強調方向性。

(10). 控制組受到干擾

一般的因果關係研究，其控制組會因為受到某種原因影響，一樣受到實驗組之介入，如此在比較介入之影響時恐怕會有低估的現象。

例：在乳癌大規模篩檢中，研究者以隨機分派方法將被研究者依序分派至篩檢組及對照組時，對照組可能會從篩檢組婦女得知資訊而要求或設法得到更多 screening service，如此在比較篩檢組及對照組的 BC 比率降低時，就可能低估，這是明顯 control contamination 例子。

(11). 對照組無介入之補償

這在一般行為科學好的介入或營養學預防性介入可以看到，因為執行研究之人員同情對照組沒有接受到好的介入，可能會用另外一種方法來補償，而這種補償又剛好和介入有關，因此會干擾介入對於結果的影響。

例：某研究欲探討維他命 C 是否可以預防 cancer 之發生，研究者為了補償對照組給予另一種營養劑 β-carotene (可能是預防 cancer 之另一種營養劑)，如此降低維他命 C 之效果。

(12).〝輸人不輸陣〞之干擾

在某些智力助益性計畫評估中，對照組由於覺得被忽視，於是尋求 extra effort 去改進 performance，於是讓助益性計畫之效果減低。

例：假設教育部為促進學童電腦能力，於是設計助益性計畫，並將學童分成兩組，其一接受助益性計畫(A)，另一組不接受(B)，比較其助益性計畫之效果，結果 B 組學童家長不甘勢弱尋找〝博士兒〞電腦補習，最後評估結果並無差異這是〝輸人不輸陣〞干擾。

(13).〝自暴自棄〞干擾

和上述相反，對照組覺得受到忽視便開始自暴自棄，如此造成介入效果之高估。

3、 測量效度（Validity in Measurement）

上述提及之內在效度大部份和研究設計相關之效度概念，也就是因研究取樣偏差、時間效應、環境變化、樣本行為造成研究因果關係之假相關或不正確推論，而若因測量變項是否可真正測到其真正含意及概念之效度則通常和測量有相關，以定義而言，測量效度可定義為 ”how well it measures what it sets out to measure”。例如，測量 pain 之 item 就應該測量 pain，而不是測量不相關之變項如 anxiety；測量和情緒相關的生活品質不要測量為 depression。以下介紹常見幾種和測量相關效度：

(1) 表面效度（Face validity）

表面效度基本上是經由一群未受專業訓練人員來判定是否所測量為『正確』。例如，測量”Aggressive behavior”，如果某問題是「請問您是否經常生氣」，如果將此問題給予您的朋友或家人判定是否此問題可代表”Aggressive behavior”，這樣之效度評估為 face validity，通常這是最不 scientific 的方法。

(2) 內容效度(Content validity)

內容效度通常是由一群對此問題了解的人員(專家)去評估是否此變項內容可以真正測量到所代表概念及含意。一般而言，內容效度是一群專家綜合的意見，嚴格來說它並非是一個 scientific measure，但是至少提供一個在方法學比 face validity 更嚴謹之效度評估。值得注意的是，因為內容效度通常是由專家判斷，常會 miss 某些層面，這也是 quantitative study(尤其是 social science)最易被 qualitative study 批評為完全是 PI or Expert viewpoint 而沒有從 information 觀點來看，因此近年來內容效度也開始相邀 client 或其 family 進入評估 content validity，其實也就是將部份 face validity 帶入一起評估。

例：某研究欲以婚姻互動來測量 health-related quality of life 之某一層面，研究者發展約 16 個問題，包括夫妻溝通、彼此信賴，對於婚姻之探討。「您是否經常下班後和您太太(先生)探討帶小孩的問題」。

研究者欲使用此 new scale 評估是否這樣的 social support 可以幫助接受化療之已婚癌症病人，為了評估內容效度他聘請 315 位 Reviewers（包括 3 oncologists、3 Psychologists、2 social workers、1 oncology nurse practitioners、4 cancer patients，及 2 spouses of cancer patients）去 review 所有相關之變項，並要求所有 reviewer 以 1-5 scale 或 1-10 scale 來量化各個變項和婚姻互動相關的適當性(Appropriateness)及相關性(Relevance)，並請 15 位列出沒有包括在這 16 個項目之相關項目，一旦這個 reviewer 完成後，研究者可以根據這些評估決定是否新的工具具有 content validity，如果研究者須評估表面效度可以請其朋友、父母親看看這些變項是否代表婚姻互動。

(3) 指標效度（Criterion validity）

指標效度是測量某一測量項和其他工具或預測相一致程度，這種效度是提供比較量化之證據，其測量相當多種，視目前文獻或過去研究所做有關工具種類及程度多寡而定，一般而言，指標效度可以分為兩種：

(A) Current validity

通常指是用目前工具與其他"gold standard"工具評估其效度為何，如果以統計觀點而言，可以使用 correlation coefficient 來表示兩者相關，相關度愈高則 concurrent validity 愈好。

例：某研究利用一個新的項目評估一群準備開刀病人之 pain tolerance，評估項目是利用病人過去對於痛的經驗，而這 4 個項目是從過去 pain tolerance index score 摘取出來，所花費時間為 1 分鐘。為了評估 concurrent validity 研究者將此四個問卷項目總合與原來 45 個項目之總合做相關得到 correlation coefficient 為 0.92，因此認為有不錯 concurrent validity。Concurrent validity 最大缺點是必須選擇適當的 gold standard。

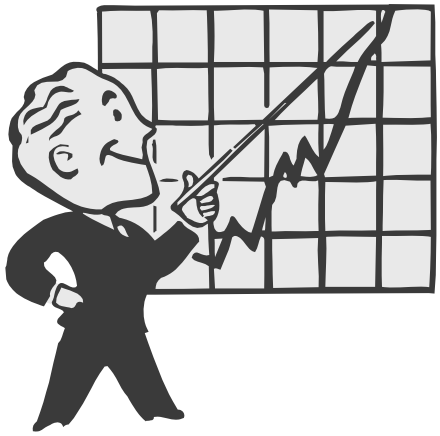
(B) 預防效度(Predictive validity)

通常是指工具預測將來事件、行為、態度及結果之能力。

九、 信效度護理流行病學 (2)

測量效度與信度

(Validity and Reliability of Measurement)



授課老師：陳秀熙 教授

測量效度與信度(Validity and Reliability of Measurement)

I Validity

上述提及之內在效度大部份和研究設計相關之效度概念，也就是因研究取樣偏差、時間效應、環境變化、樣本行為造成研究因果關係之假相關或不正確推論，而若因測量變項是否可真正測到其真正含意及概念之效度則通常和測量有相關，以定義而言，測量效度可定義為 "how well it measures what it sets out to measure"。例如，測量 pain 之 item 就應該測量 pain，而不是測量不相關之變項如 anxiety；測量和情緒相關的生活品質不要測量為 depression。以下介紹常見幾種和測量相關效度：

1. 表面效度(Face validity)

表面效度基本上是經由一群未受專業訓練人員來判定是否所測量為『正確』。例如，測量"Aggressive behavior"，如果某問題是「請問您是否經常生氣」，如果將此問題給予您的朋友或家人判定是否此問題可代表"Aggressive behavior"，這樣之效度評估為 face validity，通常這是最不 scientific 的方法。

2. 內容效度(Content validity)

內容效度通常是由一群對此問題了解的人員(專家)去評估是否此變項內容可以真正測量到所代表概念及含意。一般而言，內容效度是一群專家綜合的意見，嚴格來說它並非是一個 scientific measure，但是至少提供一個在方法學比 face validity 更嚴謹之效度評估。值得注意的是，因為內容效度通常是由專家判斷，常會 miss 某些層面，這也是 quantitative study(尤其是 social science)最易被 qualitative study 批評為完全是 PI or Expert viewpoint 而沒有從 informant 觀點來看，因此近年來內容效度也開始相邀 client 或其 family 進入評估 content validity，其實也就是將部份 face validity 帶入一起評估。

例：某研究欲以婚姻互動來測量 health-related quality of life 之某一層面，研究者發展約 16 個問題，包括夫妻溝通、彼此信賴，對於婚姻之探討。「您是否經常下班後和您太太(先生)探討帶小孩的問題」。

研究者欲使用此 new scale 評估是否這樣的 social support 可以幫助接受化療之已婚癌症病人，為了評估內容效度他聘請 15 位 Reviewers(包括 3 oncologists、3 Psychologists、2 social workers、1 oncology nurse practitioners、4 cancer patients，及 2 spouses of cancer patients)去 review 所有相關之變項，並要求所有 reviewer 以 1-5 scale 或 1-10 scale 來量化各個變項和婚姻互動相關的適當性(Appropriateness)及相關性(Relevance)，並請 15 位列出沒有包括在這 16 個項目之相關項目，一旦這個 reviewer 完成後，研究者可以根據這些評估決定

是否新的工具具有 content validity，如果研究者須評估表面效度可以請其朋友、父母親看看這些變項是否代表婚姻互動。

3. 指標效度(Criterion validity)

指標效度是測量某一測量項和其他工具或預測相一致程度，這種效度是提供比較量化之證據，其測量相當多種，視目前文獻或過去研究所做有關工具種類及程度多寡而定，一般而言，指標效度可以分為兩種：

(1) Concurrent validity

通常指是用目前工具與其他"gold standard"工具評估其效度為何，如果以統計觀點而言，可以使用 correlation coefficient 來表示兩者相關，相關度愈高則 concurrent validity 愈好。

例：某研究利用一個新的項目評估一群準備開刀病人之 pain tolerance，評估項目是利用病人過去對於痛的經驗，而這 4 個項目是從過去 pain tolerance index score 摘取出來，所花費時間為 1 分鐘。為了評估 concurrent validity 研究者將此四個問卷項目總合與原來 45 個項目之總合做相關得到 correlation coefficient 為 0.92，因此認為有不錯 concurrent validity。Concurrent validity 最大缺點是必須選擇適當的 gold standard。

(2) 預測效度(Predictive validity)

通常是指工具預測將來事件、行為、態度及結果之能力。

它可以用來預測 intervention 之效益(成功與否)，time to a clinical endpoint 等，相較於 concurrent validity，predictive validity 較常用於評估如果 time frame 很長。如何利用所得之結果預測未來資料所呈現之結果，一般而言可以用下列方法來評估：

(a)Correlation coefficient：

例：上述 Pain score 之例子，如果研究者利用新的 scale 經過 10 年之後，研究者欲利用所發展指標預測接受 hysterectomy 病人其 narcotic 之需要量，研究者已測試過 concurrent validity，如今欲測試 predictive validity 研究，目前有 24 個開完刀病人，因此每一個可以計算 index score，score 越高對於 pain 之 tolerance 越高；研究者也收集了 number of doses of narcotic，因此研究將 number of narcotic 和 pain index score 做相關，產生一 0.84，顯示 predictor validity 相當好。

(b)測量之敏感度及精確度：

Predictive validity 是測量 test 工具其敏感度及精確度重要指標，而工具可能

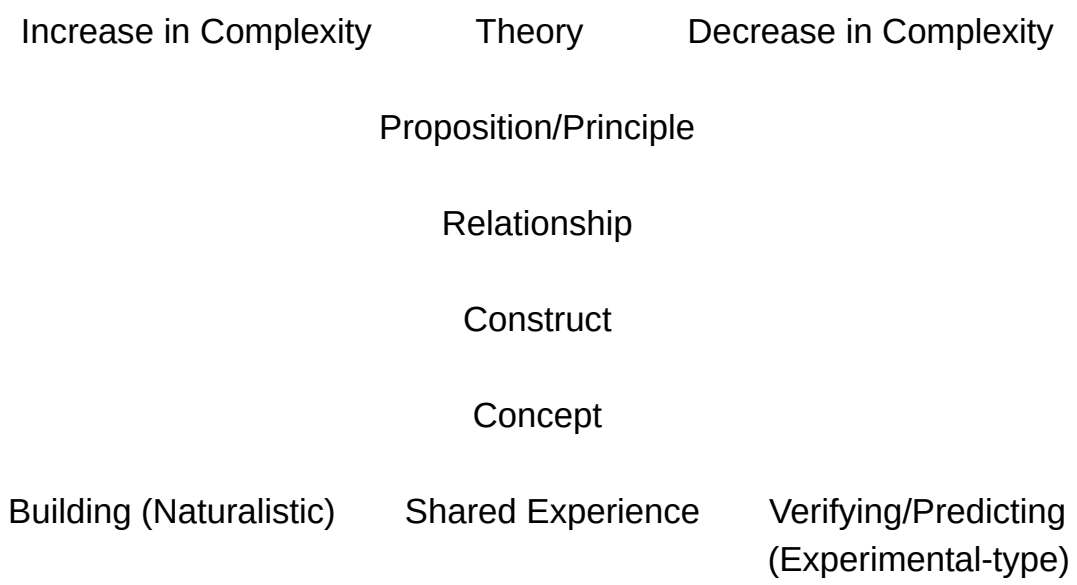
是問卷得到，也可以是血液生化變項。

4. Construct validity：

(1) 抽象概念層次（Level of Abstraction）

因為建構效度牽涉到建構(construct)因此有必要針對 construct 加以說明。Construct 是來自理論(Theory)而 Theory 是解釋或預測現象，若根據 Kerlinger 定義為"a set of interrelated constructs, definitions, and propositions that present a systematic view of phenomena by specifying relations among variables, with the purpose of explaining or predicting phenomena"根據此定義，理論提供了一個 structural map，在此 map 內利用所觀察及經歷過來促進或提供預測能力，而理論預測能力則視實證資料來支持，根據上述定義是一組相互有關成份互相架構而成，因此牽涉到抽象概念(Abstraction)層次不同，而抽象概念通常是代表 shared experience 之象徵(symbol)例：沒有寬度之圖形為線，有毛有尾巴且有四隻腳 Dog，這些例子中我們用感官，經驗創造了 Dog(symbol)例：許多生物學家發現人類遺傳結構都是雙股物(shared experience from experiment) DNA，也就是我們利用了 symbols 來描述共同經驗而人們生活中 words 也是一重抽象概念不同 words 代

表不同抽象概念下圖顯示四個不同層次抽象概念：



在這個架構下，所有抽樣概念是架構在 Shared experience，這是大家共同交換經驗及感覺之基礎點而 Shared experience 可以用 deduction 方法得到也可以用 induction 方法獲得，這個抽樣概念層次有四：

第一層是 Concept，Concept 基本上是觀察或經歷之象徵(Symbol)，Concept 可以幫助我們互相溝通經驗及概念。

沒有 concept 就沒有 language。例如在上述 dog 之例子中，有毛之生物，"dog"是一個 concept，它描述了大家所共同觀察到，furry、tail、legs、bark 也是 concept，因為這些是直接感覺到，以 deductive 層面來看 concepts 是在研究開始之前就 select 且定義好。因此可以直接

觀察或測量，在 induction 層面，concepts 是來自直接觀察。

第二層次抽樣概念是 construct，它是和 shared experience 有關但無法直接觀察及經驗到的層次。例如(1)以上述 dog 為例，對於沒有經歷過怕 dog， fear 是一個 construct 但 fear 可以用一組可觀察之 concept 組成如 sweating, shaking, turn pale 及 get away from the dog。

(2)Categories：分類本身是一種 constructs 每種分類是無法有直接觀察，但是每種分類下皆包含一組可以觀察之 concept 例如就與 CVD 相關社會分類而言，可以將人分為 Type A 及 Type B, Type A 及 Type B, 是 construct 而 Type A 及 Type B 下可以觀察到一組 concept 來代表這個 construct。

(3)Health, wellness, life roles, rehabilitation, healthy city 等是 construct 但每個 construct 下皆包含許多可以觀察之 concept。

第三層次抽象概念是 relationship，是用來描述單一 construct 或 concept 之間的關係。

例：The size of dog and the level fear that one experience

在這個 relationship 中有兩個 construct, size 及 fear 及一個 concept 就

是 dog。

第四層抽象概念是 proposition/ principles 它是用來描述不同 relationship 之間的聯結。

例：有人提出一個 proposition

Negative child hood experience → fear of large dog

這個 proposition 說明兩個 relationships:

Size of dog and fear

Childhood experience and size

不但如此 proposition 也描述了不同關係之間的方向性。

抽象概念最後一個層次是 Theory 以上述 dog 為例，以 theory 來說明如下：

Fear of large dog 是來自 Childhood experience，如此可以用心理治療法來恢復 negative effects of childhood experience 而獲得治癒。

從上述抽象概念層次我們可以將抽象概念視為 symbolic naming 及 presentation of shared experience 越遠離 shared experience 抽象概念層次越高越複雜，而層次越高所需要研究設計越複雜越謹慎。

例：

Construct 相關研究：

某研究欲探討 head injury 病人其 cognitive recovery 之狀況。這是一個有關 construct: cognitive recovery 之描述性研究，只要找到適當的 Target population 即可。

Relationship 或 Proposition 相關研究：

Age at insult and recovery rate 在這個研究必須 define" recovery " 這個 construct 及描述 age 這個 concept 和 recovery 之關係。這是一個因果相關之研究，如果將更多 construct 如 level of severity of injury 及 pre-morbidity 加入則形成 proposition，研究設計更加複雜。

與抽象概念相關之研究型態

一般而言，為了檢視這個抽象概念架構，通常有兩種 approach 方式；Experimental-type 及 naturalistic-type，簡述如下：

（1）Experimental-type：

其抽象概念架構如下

Theory

Hypothesis

Operational Definition of Concepts

Findings

Observation

例：

"Control theory"可以用來解釋 deviant behavior of adolescents

Strong "attachments" or "bonds"□deviant behavior

此例中 Strong "attachments"及 deviant behavior 代表兩種 construct:

根據此 theory 吾人可以提出下列 hypothesis" There is a negative association between attachment to parents and adolescent smoking regardless of the smoking behavior or attitudes of the adolescent's parents 在此 hypothesis 之下吾人可以將 concept 賦予 operational definition 吾人透過觀察得到資料進而決定是否 verify; refute 或 modify. 這個 theory, 這是一種 Scientific process 主要是將抽象概念層次具體化至可以檢視此理論中之關係，進而判定這理論是否正確，進而加以預測及控制

(1) Experimental-type research。主要有六個特點：

(a) Logical-deductive process

(b) Primary theory testing

(c) More from theory to less level of abstraction

(d) Assumptions of unitary reality can be measured

(e) Knowing through existing conceptions

(f) Focus on measurable parts of phenomena

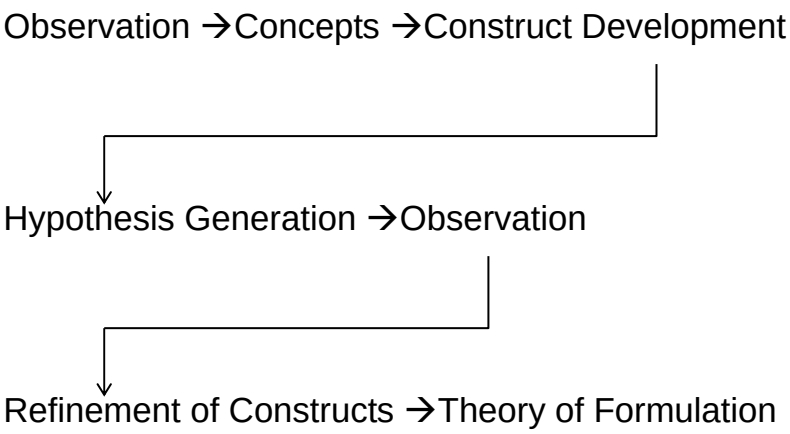
這個過程基本上是一種 logical deduction。

(2). Naturalistic-type 和 Experimental-type 相反之 approach 主要不是 test theory 而是透過 Observation 去得到 concept 並非在實驗設計前就已經定義好，而是經由整個 data collection on shared experience 來定義 concept 及 construct 通常這樣方法所形成之 theory 稱為"Ground theory"，這樣的想法通常屬於 Qualitative study，而其推論過程也比較像 inductive process。在某些 naturalistic inquiries 過程中，研究者會加入引用 well-established theories 來解釋這些觀察值，使這些觀察值在解釋上更合理。

例：再觀察 gift exchange 過程中，民族 ethnographer 可能會在收集資料後使用已知理論來解釋這種 gift exchange 之現象。

一般 Experimental-type research 研究者會 criticize Naturalistic-types research 使用 theory 來解釋這些 observation 有違背 naturalistic principle，其實並沒有因為利用 theory 來解釋 observation 是在 data collection 之後，而非之前。

對於 Naturalistic-type researcher 而言，其過程如下：



Naturalistic-type design 之主要特徵包括：

(a)Primary inductive

(b)Primary theory generating

(c)More from shared experience to higher level

(d)Assumption of discovering meaning through multiple subjective understandings

(e)Informant as knower

(f)Focus on understanding complexity

建構效度在抽象概念之角色

因為 Construct 是來自 theory 而且是不可觀察的，如果是使用 Experimental-type testing theory 之方式來做，利用 hypothesis testing 將 construct 中的 concept 做 operational definition 進而 test，所屬

construct 是否正確，其實是抽象概念印證 theory 之重要步驟，因為唯有透過針對單一 construct 做效度驗證才能得之複雜之 relationship 及 proposition 所得到之 Theory 是否正確。

例如在照顧 Alzheimer's disease 之人,其 stress 和病人之 level of function 有關，"Stress 及 level of function"是兩個 construct，如果欲了解這兩個 construct 之間關係是否有相關，必須藉由 Hypothesis:

level of function stress 如此透過 operational definition 將 stress 及 level of function 變成可以測量來 verify 其建構效度是正確。

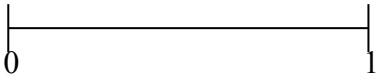
Construct Validity(**建構效度)**

簡而言之，因為建構效度是建立在 Theory 之下，在檢視其效度過程中必須從理論、relationship 及 proposition 出發，最後藉由 Hypothesis 來檢視其各種 construct 之效度,上述 ”Stress” 和 ”level of function” 就是一個實例，不過由於在檢視建構效度過程中經常會遭遇某種自變項的概念或特質(Independent construct)和依變項的概念或特質會受第三種概念或特質之影響。因此，一般而言，將建構效度分成兩種：Convergent validity 及 Divergent validity。

Campbell 和 Fiske (1959)提出的多種方法矩陣

(Multitrait-multimethod matrix，MTMM)是檢視上述二種效度之重要方法，以下舉一例來說明這個方法。

例：假設某研究分別在兩個不同時間測量緊張、焦慮、足球大小，每次都用不同測量方法，目測法及訪視法。



目測法：Visual Scale

訪視法：以量表或以工具實際測量足球

如果將兩次結果以矩陣表示：

		1st					
		緊張$_{目}$	緊張$_{訪}$	焦慮$_{目}$	焦慮$_{訪}$	足球$_{目}$	足球$_{訪}$
2nd	緊張$_{目}$	R	+				
	緊張$_{訪}$	+	R				
	焦慮$_{目}$	#					
	焦慮$_{訪}$		#				
	足球$_{目}$	#					
	足球$_{訪}$		#				

以這個矩陣為例，相同特質(Construct)和相同方法應該具有最高

相關(Reliability)，而相同特質不同方法其相關應為次高，這是檢視 convergent validity。如是不同特質其相關性則視兩個不同特質是否有共同特性(共變)，共同特性(共變)愈高則相關愈高，在此例中焦慮與緊張即是用不同測量法，應該會有相關，因此是用來檢視 divergent validity，而緊張或焦慮應該與足球大小無關，此時所得相關程度應該接近於 0。若以相關係數來表示，信度＞Convergent validity＞Divergent validity＞No correlation。

MTMM 之概念對於後來效度檢視貢獻相當大，因為它可以區分共變是來自“特質”或“方法”，也就是反映建構效度若要能夠真正被檢視則必須考慮這兩個截面。

例：某研究欲檢視醫護人員(M)與家屬(F)對 palliative care 之態度，研究者利用觀察法(O)及問卷(Q)法分別得到兩種結果，若利用 MTMM 得到下列結果：

		Observed		Question	
		M	F	M	F
Obs	M	(0.96)	0.45		
	F	0.45	(0.98)		
Que	M	0.75	0.15	(0.91)	0.45
	F	0.15	0.75	0.45	(0.90)

( )→Reliability

Convergent validity=0.75 > divergent validity=0.45 > 0.15 (No correlation)

上述 MTMM 可以宣稱是在檢視兩個 construct 表示醫護人員和家屬對於 palliative care 的態度不同。

如果 MTMM 所得結果如下：

		Observed		Ques	
		M	F	M	F
Obs	M	(0.96)	0.85		
	F	0.85	(0.98)		
Que	M	0.45	0.25	(0.91)	0.80
	F	0.25	0.45	0.80	(0.90)

Divergent validity=0.85 > Convergent validity=0.45，此表示這兩個 construct 可能非常相似。

Construct validity 是效度中最重要但難表達，這種型態的 validity 經常是經常是一種測量某個變項或工具之 Abstractive concept 而非以 statistical quantity 來表示常有 2 種型式包括 convergent validity 及 divergent validity.

Remark: Convergent validity 意味著對於相同特性或概念用不同方法所得到之結果含相同，convergent validity 和 alternate-form reliability 類似，然而 convergent validity 較為 theoretical 而 alternate-from reliability 是實際執行時工具或變項之一致程度。 Divergent (discriminant) validity 和 convergent validity 剛好相反，以相同方法測量不同特性其 correlation 不高，比較少使用。

效度之統計觀點：

假定某種測量或測驗是針對某個變項，則其結果之總變異量（σ_T）是來自三部份。

（1） 受試者與該變項相關之變異量σ_1^2

（2） 另一是與該變項沒有相關之其他變項σ_2^2

（3） 誤差變異σ_e^2

若以公式來表示：

$$\sigma_T^2 = \sigma_1^2 + \sigma_2^2 + \sigma_e^2$$

效度就是由測量結果之總變異量中由本次測量所關心變項所造成變異量所佔百分比：

$$Validity = \frac{\sigma_1^2}{\sigma_T^2}$$

信度(Reliability)

一般而言，在收集資料研究過程中經常含有 error 產生，為了反應真實狀況，研究者通過必須 minimize 這種 error，一般而言 error 可分為兩種：(1)Random error：通常來講，這是一種 by chance error，是屬於統計抽樣問題，降低 random error 最好方法是選擇大且具代表性之樣本，然而研究樣本數愈大則所須 cost 愈高，如何估計 optimal 樣本是研究法中 sample size 估計問題。

(2)Measure meat error：在選定 population 之後研究工具測量之好壞，現有一種工具是 perfect，測量是一門大學問，俗話說"Measurement is science" 來表示。

信度在研究設計是用來測量利用工具收集所欲映證之因果關係其結果重覆出現之可信程度，概言之有三種型式之信度 Test-retest，alternate-from 及 consistency。

Test-retest reliability

(a) 使用 Test-retest reliability 必須非常小心，因為其 correlation 變低或變高，常常不是反映其 Test-retest reliability 不好，而是本身特性所引起的反映真正變化,例如：在 Anxiety 研究中，研究者在 Time1，Time2，Time3 及 Time4 分別測量其 outcome 然後測量其兩次之間 correlation 發現 1&2 為 0.84、1&3 為 0.12 而 2&3 為 0.09。這樣之結果可能不是反映 poor Test-retest reliability 而是真正變化。

(b) Practice effect：有些時候 Test-retest reliability 之 correlation 很高並

非是 good test-retest reliability 而是因為先前受過 test，因此有記憶效果進而造成 good test-retest reliability，在這種情況之下，研究者可以調整 Test-retest reliability 之時間間隔，不過適當間隔視主題而定。

(2)複本信度(Alternate-form Reliability)

為了克服上述 practice effect，研究者可以使用測量相同性質但在 wording 或 order 上有差異之平行兩組問題來測量。如果兩組之間 correlation 高則有 good alternate-form reliability，複本信度可以同時連續實施，也可以相距一段時間分兩次測試，前者稱為 coefficient of equivalence，後者又稱為 coefficient of stability and equivalence，分別說明內容和時間變異所造成的誤差。

例：某研究以下列問題測量 Urinary function 研究，研究者分別使用 alternate form 來測量兩個問題

量化頻率

1.過去一星期中，您每天解尿之頻率

(1)1-2 次 (2)3-4 次 (3)5-8 次 (4)12 次 (5)12 次以上

2.過去一星期中，您解尿頻率

(1)每 12-24 小時 (2)每 6-8 小時 (3)每 3-5 小時 (4)每 2 小時 (5)

小於 2 小時

3.過去一個月中，您排尿狀況為何

(1) 困難 (2)有點困難 (3)還好 (4)算頻繁 (5)相當頻繁

上題之複本信度所對應問題為

(1) 沒有使用 pad

(2) 每天使用 1 次 pad

(3) 每天使用 2~3 次 pad

(4) 每天使用 4~ 6 次 pad

(5) 每天使用≥7 次 pad

(3)折半信度(Split-half reliability)

有些時侯研究者沒有平行問題而且只能實測一次，通常會使用折半信度來估計信度，一般而言，折半信度是將測量結果接題目單變號分成兩半計分再根據其相關係數來判定折半信度好壞因為折半信度只有半個測驗，因此必須使用公式校正一般使用 Spearman-Brown formula (J.C. Stanely and K.D. Hopkia 1972)

$$R = \frac{2 \times r}{1 + r}$$

r 是相關係數，R 是折半信度

而這個公式是建立在兩測驗分數其 variance 為 equal 之下因此最好

是使用下列公式

$$R = 2\left(1 - \frac{S_1^2 + S_2^2}{S^2}\right)$$

S^2 及 S_2^2 分別兩組總變異量而 S^2 是整體總變異量

(4)內在一致性(Internal consistency)

內在一致性效度是用來檢視同一 concept 但以不同問題來測量其間之一致性，在社會行為科學上經常會使用

例：SF-36 針對不同 medical conditions 之病人測量 quality of life，其中一項測量 physical function 總共有 10 個問題

以下是每天可能會做之活動，您是否會因為健康狀況而有所不便

	非常不便	稍不便	非常方便
1.劇列運動如跑步	1 □	2 □	3 □
2.中度運動	1 □	2 □	3 □
3.提雜物	1 □	2 □	3 □
4.爬樓梯	1 □	2 □	3 □
:			
:			
9. 走一段路	1 □	2 □	3 □
10. 自已梳洗	1 □	2 □	3 □

內在一致性因素測量這 10 個反映 physical function 其一致性如何，而所使用之統計值為 Conbach's α 以下舉一例來示範其計算

例：下例三個問題是用來反映 mental health scale

	是	否
(a) 您相當容易緊張嗎?	1	0
(b) 您相當容易憂愁或失望	1	0
(c) 您相當容易沮喪而不快樂	1	0

今有 5 位病人其回答狀況如下：

	(a)	(b)	(c)	總分
1	0	1	1	2
2	1	1	1	3
3	0	0	0	0
4	1	1	1	3
5	1	1	0	2
Yes 之比例	3/5	4/5	3/5	

$\overline{X}=2$

$s^2=1.5$

$$\alpha = \left(1 - \frac{(\%Yes)\times(\%No)}{Sample\ varianl}\right)\left(\frac{K}{K-1}\right)$$

K 是 item

$$\alpha = \left(1 - \frac{0.6\times0.4 + 0.8\times0.2 + 0.6\times0.4}{1.5}\right) = 0.86$$

上述這個公式又叫做庫存信度(Kuder-Richardson)這個公式僅適用於二分變項(binary variable)而一般的 Crondach 之公式為

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum\sigma_i^2}{\sigma_y^2}\right)$$

σ_y^2是總分變異量　　σ_i^2是單一題目之變異量

(5)Interobserver Reliability

在許多醫學研究研究者為了了解不同人對於同一現象判斷是否一致，這樣之信度稱為 Interobserver Reliability，這種信度測量可以使用

(A) Correlation coefficient

(B) Reliability index

(C) Kappa value

例：在利用 Mammography 早期偵測 breast cancer 研究者使用為了了解是否不同 reader 會有不同結果，兩位 Radiologist 分別針對 100 個，mammographam 做判讀以下是其結果

Reader　1

\+　　-

Reader 2 +25(12.24)9(21.76)34

-11(23.76)55(42.24)60

36 64 100

Reliability index $= \frac{25+55}{100} = 80\%$

Kappa $= \frac{Po - Pe}{1 - Pe} = 56.06\%$

十、 傳染病流行病學之護理觀

I 傳染病流行病學及防治簡介－以禽流感為例

當傳染病流行發生，一般大家喜歡問「是否發生社區流行？」禽流感一定也不例外。

禽流感在感染的接觸型態上相當廣泛，因此感染源不易界定，一般在此狀況下，會使用平均再生數（Reproductive Number, R0），也就是每個第一代個案（Primary Case）在傳播過程中產生第二代個案的平均人數，或稱為 R0 值來界定流行，且 R0 愈大流行越會發生。

決定禽流感平均再生數有三個要素，若以天為單位，則平均再生數就等於：

接觸率 × 傳播成功機率 × 被感染個案傳染期（天）長短。

根據上述決定平均再生數之三要素，禽流感防治措施可有多種設計。依有效程度而言，分述如下：

對付此種傳播力強之禽流感最佳方法為於未流行前施打疫苗，減少可感染宿主進而降低接觸率、減低傳播成功機率及減低傳染力或減低變成

臨床症狀機率。然而由於禽流感病毒變異性相當高，疫苗製造速度可能不及流行發生速度，因此需其他措施加以輔助。

由於禽流感病毒的高變異性，且疫苗生產需要時間，退而求之是使用治療流感藥物，包括瑞樂沙及克流感等來減低流行傳播速度，提供較充裕的疫苗製造時間。就 R0 值降低而言，使用流感用藥，能同時減低可感染宿主之機率及傳染力進而降低平均再生數。以流行於泰國的 H5N1 型為例，僅採取抗流感用藥預防可以降低百分之三十感染及減低百分之六十傳染，如此可應付至少 R0 值為三點六之禽流感流行菌株。

一般而言，因為流感用藥相當昂貴而且必須施與抗病毒藥達三週（每個個案至少需兩個療程），若對全部族群施予投藥似乎不盡可行，常見作法是使用高危險標定。高危險標定方法有兩種， 第一種目標族群標定（Social Targeting），即針對與指標個案接觸之同一家戶、學校、職場或醫院施予流感用藥，通常在新發生個案不超過十個的情況下，採用此方法施予流感用藥可以有效控制流行。第二種作法是地理區域目標族群設定（Geographic Targeting），即以指標個案出現地區的方圓一定範圍（視該地區人口密度而定）的地理區域進行用藥，一般而言 R0 值愈大用藥範圍應愈大。若以 H5N1 病毒為例，方圓五公里之內給予預防性用藥，可以控制 R0 值為一點五之禽流感流行。

若前述疫苗及流感用藥無法取得的情況下，第三種預防禽流感的方法為使用隔離（Isolation） 及檢疫（Quarantine）等措施，其目的主要是降低接觸率進而降低平均再生數。然而由於禽流感會有隱性感染之問題，因此隔離及檢疫實施上較不易產生良好效果，除非前兩者無法實施，否則隔離及檢疫應該僅是配套措施。

2. 10 價流感嗜血桿菌疫苗評估

(Effect of the 10-valent pneumococcal nontypeable*Haemophilus influenza* protein D-conjugate vaccine on nasopharyngeal bacterial colonization in young children: A randomized controlled trial.)

背景: 本研究旨在比較利用 10 價流感嗜血桿菌 (*Haemophilus influenza*)疫苗相較於 7 價疫苗是否對於幼兒之鼻咽部無法分型(non-typeable)之流感嗜血桿菌帶菌株是否有所不同，其主要證明對於其集團免疫是否有效。

方法: 本研究利用在荷蘭進行之隨機分派對照試驗設計研究進行評估。研究開始於 7 價疫苗引進市場的兩年後，研究期間涵蓋 2008 年 4 月 1 日至 2010 年 12 月 1 日。研究對象包含 780 名嬰兒，以 2:1 的比例分別在第 2,3,4 以及 11 到 13 個月接受 7 價疫苗或 10 價疫苗。受試者之鼻咽部檢體分別於第 5,11,14,18 以及 24 個月採取並以培養之方式檢測樣本中是否有流感嗜血桿菌、肺炎鏈球菌、黏膜莫拉克氏菌 (*Moraxella catarrhalis*)以及金黃色葡萄球菌。並配合聚合酶鏈鎖反應 (PCR)對流感嗜血桿菌與肺炎鏈球菌進行定量分析以及判定流感嗜血桿菌是否為屬於無法分型之菌株。研究之主要觀察結果為疫苗在無法分型之流感嗜血桿菌於鼻咽之帶菌株比例上的效益 (Vaccine efficacy, VE)。

結果: 在接受 7 價疫苗以及 10 價疫苗組在鼻咽部無法分型之流感嗜血桿菌帶菌株比例皆隨年齡進展而增加：5 個月時為 33%, 至 24 個月時達 65%。在增效接種(booster)的三個月後，對於帶菌株之疫苗效益(VE)為 0.5% (95% 信賴區間: -21.8%,18.4%)。而對於罹病之疫苗效益(VE)為 10% (95% 信賴區間: -31.3%,38.9%)。在上述的鼻咽檢體採樣時點中，兩組間的無法分型之流感嗜血桿菌盛行率以及發生率皆無差異。肺炎鏈球菌、黏膜莫拉克氏菌以及金黃色葡萄球菌之帶菌狀況在兩組間無差異。

結論:

在無法分型之流感嗜血桿菌鼻咽帶菌株比例與發生率上，10 價疫苗與 7 價疫苗於本研究族群(2 歲以上之荷蘭健康兒童)之效益沒有差異。此結果顯示了 10 價疫苗可能無法達到族群集團免疫(herd effect)的效果。其餘的菌株帶菌在兩組間亦無差異。

II 統合分析:

1. **Mantel-Haenszel method: 以 Mantel-Haenszel 估計式計算統合勝算比 (pooled OR)** (ebook 01. Ch 10.8. p.871-884)

臨床實例：為了解在懷孕期使用類固醇藥物 (corticosteroid) 是否與新生兒先天異常(pregnancy outcome: child malformation)有關, Wyllie 等人收集了 4 個病例-對

照研究利用 Mantel-Haenszel 方法進行統合分析，資料如下:

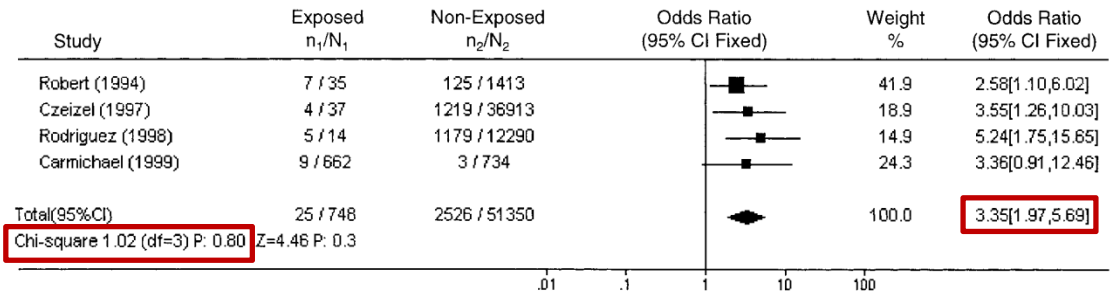


Fig. 2. Individual and cumulative Mantel-Haenszel summary odds ratio for corticosteroid-exposed case-control studies focusing on oral clefts.

對於研究想探討的臨床問題,可能已經有其他研究者進行類似的研究並提出結果，因此藉由統合分析，研究者可以綜合這些已知的結果進行評估,免於再次重複大規模的研究 (例如範例中的懷孕期類固醇使用與新生兒異常之相關性), 進行統合分析的好處之一，是可以藉由綜合已知的研究，達到增加樣本數的目的: 例如在範例的資料中，若每個研究的樣本數都不多，則雖然暴露與結果可能是相關的，但卻受限於樣本數不足無法達到顯著的程度，在這樣得情形下，統合分析就可以提供解決的方法.

以 Mantel-Haenszel 方法進行統合分析，如同分層分析中合併不同分層結果的概念一般: 將不同的研究視為不同的分層，利用 Mante-Haenszel 方法給予各個分層權重後得到加總 (統合, pooled) 的估計值，若每一個分層 (即各個研究) 資料如下:

	E	$\overline{E}$	
D	a_i	b_i	m_{1i}
$\overline{D}$	c_i	d_i	m_{2i}
	n_{1i}	n_{2i}	n_i

則 Mantel-Haenszel 估計式即為:

$$\hat{OR}_{MH} = \frac{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)\left(\frac{a_i d_i}{b_i c_i}\right)}{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)} = \frac{\sum_{i=1}^{k}\left(\frac{a_i d_i}{n_i}\right)}{\sum_{i=1}^{k}\left(\frac{b_i c_i}{n_i}\right)}$$

在臨床實例的表格中,“Odds Ratio” 欄以圖示方法標記出每個研究的估計值以及範圍, 圖中的點估計值以不同大小的方塊表示; 方塊的大小即表示各個研究估計值權重的大小, 權重大則貢獻在統合估計值的分量也就多.

將上述資料整理如下:

	研究 1		**研究 2**		**研究 3**		**研究 4**	
	E	$\bar{E}$	E	$\bar{E}$	E	$\bar{E}$	E	$\bar{E}$
D	7	125	4	1219	5	1179	9	3
$\bar{D}$	28	1288	33	35694	9	11111	653	731

研究 1 到研究 4 可以看成 4 個分層, 利用 Mantel-Haenszel 估計式計算統合勝算比($\hat{OR}_{MH}$)及 95%信賴區間為: 3.35 (1.97, 5.69)

統計報表:

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	Mantel-Haenszel	3.3472	1.9677	5.6939
	Logit	3.4065	2.0299	5.7165
Cohort (Col1 Risk)	Mantel-Haenszel	2.9336	1.8790	4.5801
	Logit	2.9955	1.9851	4.5203
Cohort (Col2 Risk)	Mantel-Haenszel	0.9682	0.9494	0.9874
	Logit	0.9892	0.9794	0.9992

Breslow-Day Test for Homogeneity of the Odds Ratios

Chi-Square	1.0322
DF	3
Pr > ChiSq	0.7934

Total Sample Size = 52098

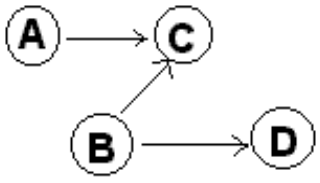
在臨床實例的表格中，除了統合風險勝算比外，左下角處還報導了各研究的勝算比之間是否同質的驗證結果: “Chi-square 1.02 (df=3) P:0.8”. 再利用 Mantel-Haenszel 方法合併個研究(分層)的結果得到統合估計值之前，必須確定各分層間的勝算比是否同質 (homogenous)，若同質才能進行合併估計，統計報表中也有相同的結果: “Breslow-Day Test for Homogeneity of the Odds Ratios, $\chi^2 = 1.032 \sim \chi^2_3, P = 0.793$. 這表示各研究的勝算比之間具有同質性: 暴露與結果的勝算比不因為研究的不同而有所不同. 這也表示暴露的效果與各個不同的研究之間並無異質性 (Heterogeneity) 存在. 各個研究設計的情境 (如收案以及排除標準，研究中暴露以及結果的定義)，相差不多，所以利用統合方法合併各個研究的結果是合理的. 同質性的成立也表示各研究設計的情境相差不多即表示收集到研究樣本類似，若同質性檢定不成立，就必須要用更複雜的方法調整各個研究間的差異才能得到正確的統合估計結果.

2. **統合分析貝氏羅吉斯隨機效應迴歸模型 (Meta-Analysis with Bayesian Logistic Random-effects Regression Model)**

Bayesian Directed Acyclic Graphical Models

補充說明

當考慮之共變數(covariate)增加時，模型變得複雜，此時 full conditional distribution 最好用 Bayesian directed acyclic graphical model (DAG) model 較直接這樣之圖形模型(graphic model)常用來處理條件獨立(conditional independence)之假設如下圖 D 只條件於 B 但獨立於 A,C，



the full joint distribution P(V)=P(A)P(B)P(C|A,B)P(D|B)

也就是 $P(V) = \prod_{\upsilon \in V} P(\upsilon \mid parents[\upsilon])$，

如此只要將欲處理之研究其 acyclic graphical model 弄清楚，則 joint

distribution 就可以在既定之參數得到。

例：統合分析貝氏隨機效應模型(Meta-analysis with Bayesian random-effects model)：

假定在 22 個研究中探討治療組及對照組其呼吸道感染得下表：

22 個呼吸道感染研究之個別及利用 Mantel-Haenszel 方法估計之統合勝算比(odds ratios)及 95%信賴區間

研究	感染人數/總人數		原始勝算比	估計勝算比(隨機效應)
	治療組	對照組	(95% 信賴區間)	(95% 信賴區間)
1	7/47	25/54	0.20 (0.08-0.53)	0.22 (0.08-0.46)
2	4/38	24/41	0.08 (0.03-0.28)	0.13 (0.04-0.29)
3	20/96	37/95	0.41 (0.22-0.78)	0.39 (0.20-0.68)
4	1/14	11/17	0.04 (0.004-0.40)	0.13 (0.02-0.37)
5	10/48	26/49	0.23 (0.10-0.57)	0.25 (0.10-0.50)
6	2/101	13/84	0.11 (0.02-0.50)	0.17 (0.04-0.41)
7	12/161	38/170	0.28 (0.14-0.56)	0.28 (0.14-0.50)
8	1/28	29/60	0.04 (0.01-0.31)	0.11 (0.02-0.28)
9	1/19	9/20	0.07 (0.01-0.61)	0.16 (0.03-0.45)
10	22/49	44/47	0.06 (0.02-0.20)	0.10 (0.03-0.23)
11	25/162	30/160	0.79 (0.44-1.42)	0.71 (0.38-1.21)
12	31/200	40/185	0.67 (0.40-1.12)	0.61 (0.36-0.98)
13	9/39	10/41	0.93 (0.33-2.61)	0.68 (0.25-1.54)
14	22/193	40/185	0.47 (0.27-0.82)	0.44 (0.25-0.73)
15	0/45	4/46	-	0.19 (0.02-0.60)

16	31/131	60/140	0.41 (0.25-0.70)	0.40 (0.23-0.64)
17	4/75	12/75	0.30 (0.09-0.96)	0.29 (0.09-0.68)
18	31/220	42/225	0.72 (0.43-1.19)	0.66 (0.39-1.04)
19	7/55	26/57	0.17 (0.07-0.45)	0.20 (0.08-0.40)
20	3/91	17/92	0.15 (0.04-0.53)	0.19 (0.06-0.43)
21	14/25	23/23	-	0.11 (0.02-0.31)
22	3/65	6/68	0.50 (0.12-2.09)	0.40 (0.10-1.06)

＊ 上表之統合估計值

(A) 固定效應模型(Fixed-effect Model): 勝算比 OR=0.35 (0.30-0.42)

(Mantel-Haenszel 方法)，同質性檢定(Homogeneous test) $\chi^2_{(21)}$=70.01, p<0.0001

(B) 隨機效應模型(Random-effect Model) (假定每個研究不一定相同而是從特定不同族群抽到之樣本): 勝算比 OR=0.228 (0.138-0.353)

我們想用 full Bayesian model 來探討在隨機效應模型(random-effect model)下其治療組與對照組之對比值(Odds ratio)。

令 γ_i^c 和 n_i^c 代表第 i 個試驗(trial)對照組之感染人數和總人數，P_i^c 代表其感染

之機率， γ_i^T, n_i^T 及 P_i^T 則代表治療組，

假如 δ_i 是第 i 個之治療效應(true treatment effect)(以 log-odds 之 scale)，如此

$$\delta_i = \text{logit}(P_i^T) - \text{logit}(P_i^C)$$
$$\mu_i = \left[\text{logit}(P_i^T) + \text{logit}(P_i^C)\right]/2$$

且假定個別試驗效應(individual trial effects)是從某個 Gaussian population ~ $N(d,\sigma^2)$ 抽出，如此 full model 可寫為

$$r_i^C \sim Binomial(P_i^C, n_i^C)$$

$$r_i^T \sim Binomial(P_i^T, n_i^T)$$

$$\text{logit}(P_i^C) = \mu_i - \frac{\delta_i}{2}$$

$$\text{logit}(P_i^T) = \mu_i + \frac{\delta_i}{2}$$

$$\delta_i \sim N(d,\sigma^2)$$

如果用 directed acyclic graphical model 來建構上述之關係得下圖，

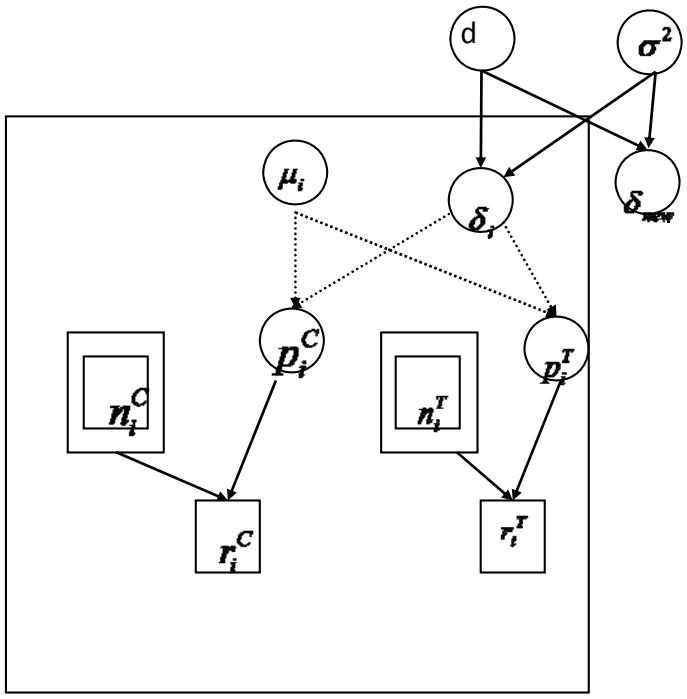


Figure 1　Directed Acyclic Graphical model for random-effects meta-analysis

圖中矩形□代表觀察變項(γ_i^C,γ_i^T)，雙矩形 □ 代表因研究設計之固定值(n_i^C,n_i^T)，而圓圈◯代表 unobserved variable$(P_i^C,P_i^T,\mu_i,\delta_i,\delta_{new},d,\sigma^2)$，箭頭由父母結節(parent nodes)至子系結節代表條件獨立假設(conditional independence assumption)，直接連結(the directed links)表示在母結節 pa(v)之下，每個結節(v)除了和父母結節有關參數之外和其他參數間是獨立的，如δ_{new}只和 d 和σ^2有關但和其他參數間是獨立。聯結(link)可能代表 stochastic dependence (實線箭頭) 或

只是一個 logical function (虛線箭頭)。

為了表示 full joint probability distribution of all quantity V，我們只須 specify the conditional parent-child distribution $P(v \mid pa(v))$

$$P(V) = \prod_{\upsilon \in \mathrm{V}} P(\upsilon \mid pa[\upsilon]) \quad \text{——————————} (1)$$

上述之模型利用上述條件獨立可表示為

$$P(\underline{r^C}, \underline{r^T}, \underline{\mu}, \underline{\delta}, \sigma^2)\alpha$$
$$\prod_i P(r_i^C \mid \mu_i, \delta_i) P(r_i^T \mid \mu_i, \delta_i) P(\mu_i) P(\delta_i \mid d, \sigma^2) P(d) P(\sigma^2) \quad \text{———}(2)$$

從式(2)欲得 appropriate posterior marginal distribution 會遭遇須要計算 high dimension 之 integration。

我們感興趣的是參數 d，我們必須得到邊際分布(marginal distribution) $P(d \mid \underline{r^C}, \underline{r^T})$，應用上式(2)將其他參數進行積分如下式

$$P(d \mid \underline{r^C}, \underline{r^T}) \ \alpha \ P(\underline{r^C}, \underline{r^T}, d) P(d)$$
$$\propto \iiint P(\underline{r^C}, \underline{r^T}, \mu, \delta, d, \sigma^2) d\mu \ d\delta \ d\sigma^2 \quad (3)$$

式(3)雖然有 simple form，然而這個積分無法用 closed form 表示，因此必須以數值積分(numerical integration)或模擬(simulation)來克服這個問題。

上述感興趣參數之估計結果如下 (參數估計是以 Bayesian Markov chain Monte Carlo (MCMC)方法進行)

參數	估計值	95%信賴區間
d	-1.479	(-1.980, -1.042)
Tau(=$1/\sigma^2$)	1.347	(0.534, 4.044)
$delta_{new}$	-1.483	(-3.372, 0.316)
OR_d	0.228	(0.138, 0.353)

流行病學方法論及實驗
(The Methods of Epidemiology and Practices)
：傳染病流行病學

授課教師：陳秀熙 教授/許辰陽 博士

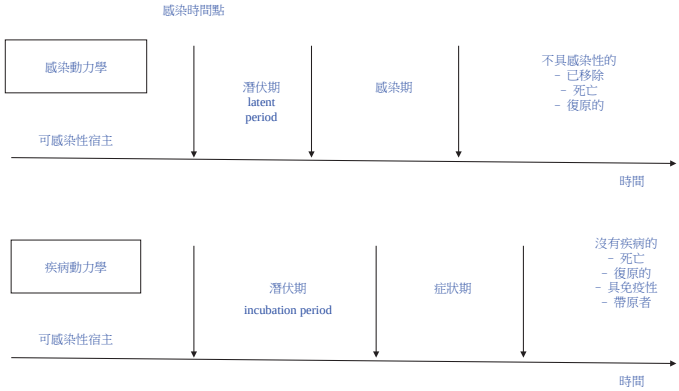
台灣大學流行病學與預防醫學研究所

應用實例

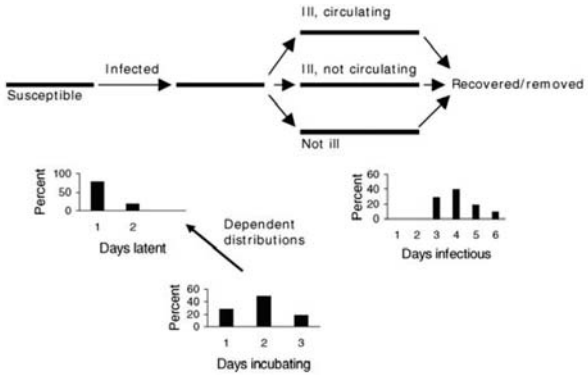
- (1)某國小發生1個Chickenpox 個案(case)，其最佳方法是將此個案隔離？
- (2)若無法確定明顯暴露接觸時，如何判定社區感染？
- (3)一般社區感染決定因素為何？如何量化？
- (4)如何量化〝流行〞(Epidemic)及〝地方性〞(Endemic)？
- (5)若欲達到〝流行〞不發生，疫苗接種率(集團免疫)應達到多少%？

(1)某國小發生1個Chickenpox 個案(case)，其最佳方法 是將此個案隔離？

感染和疾病於不同時間之表現



The Natural History of Influenza



(3)一般社區感染決定因素為何？若無法確定明顯暴露接觸時 ？ 如何判定社區感染？如何量化？

基礎再生數

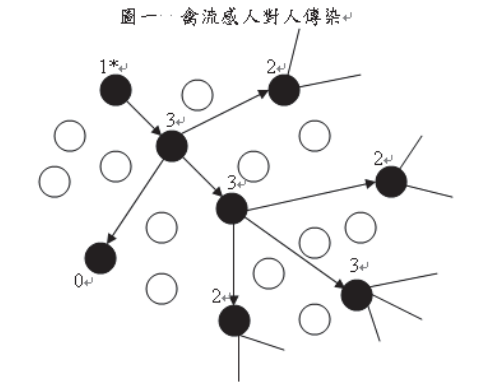
定義:

- 一個傳染病發病個案（Primary Case）在其可傳染期間於一個完全易感受的族群中預期可產生的新個數 (Secondary case)

R_0 = 每單位時間接觸次數 (Contact Rate) X 每次接觸的傳播機率 Transmission Probability X 可傳染期間 (Duration)

R_0 : 基礎再生數

$R_0 > 1$, 一個易感受族群將發生流行
$R_0 < 1$, 流行將不會發生
$R_0 = 1$, 可能有小波例群聚的情形



Transmissibility of 1918 pandemic influenza

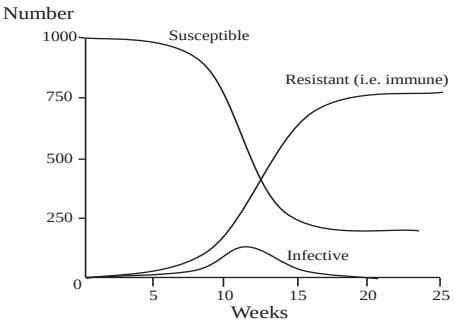
letters to nature

Christina E. Mills[1], James M. Robins[1,2] & Marc Lipsitch[1,3]

The 1918 influenza pandemic killed 20–40 million people worldwide, and is seen as a worst-case scenario for pandemic planning. Like other pandemic influenza strains, the 1918 A/H1N1 strain spread extremely rapidly. A measure of transmissibility and of the stringency of control measures required to stop an epidemic is the reproductive number, which is the number of secondary cases produced by each primary case. Here we obtained an estimate of the reproductive number for 1918 influenza by fitting a deterministic SEIR (susceptible-exposedinfectious-recovered) model to pneumonia and influenza death epidemic curves from 45 US cities: the median value is less than three. The estimated proportion of the population with A/H1N1 immunity before September 1918 implies a median basic reproductive number of less than four. These results strongly suggest that the reproductive number for 1918 pandemic influenza is not large relative to many other infectious diseases2. In theory, a similar novel influenza subtype could be controlled. But because influenza is frequently transmitted before a specific diagnosis is possible and there is a dearth of global antiviral and vaccine stores, aggressive transmission reducing measures will probably be required.

NATURE | VOL 432 | 16 DECEMBER 2004 | www.nature.com/nature

(4)如何量化〝流行〞(Epidemic)及〝地方性〞(Endemic)?



(5)
若欲達到〝流行〞不發生，疫苗接率(集團免疫)應達到多少%？

有效再生數和免疫

$R = R_0 f$

R: 有效再生數
R_0: 基礎再生數
x: 易感受宿主佔族群之比例

f 族群中免疫的比例 (疫苗注射)
1-*f* 族群中易感受宿主比例

降低疾病傳播之條件:

$$R = R_0 (1-f) < 1$$

一族群要降低疾病傳播需要疫苗注射之比例:

$$f > 1-1/R_0$$

較高的 R_0 值需要較高的疫苗注射比例以降低疾病傳播

Estimated values of the basic reproductive rate, R_0

Infection	Geographical location	Time period	R_0
Mumps (腮腺炎)	Baltimore, USA England and Wales Netherlands	1943 1960-80 1970-80	7-8 11-14 11-14
Rubella (德國麻疹)	England and Wales West Germany Czechoslovakia Poland Gambia	1960-70 1970-7 1970-7 1970-7 1976	6-7 6-7 8-9 11-12 15-16
Poliomyelitis (小兒麻痺)	USA Netherlands	1955 1960	5-6 6-7
Human Immunodeficiency Virus (Type I) (HIV)	England and Wales (male homosexuals) Nairobi, Kenya (female prostitutes) Kampala, Uganda (heterosexuals)	1981-5 19981-5 1985-7	2-5 11-12 10-11

Approximate estimates of the vaccination coverage (the degree of herd immunity) required to eradicate a variety of viral, bacterial, and protozoan infections in developed and developing countries

Infectious disease	Critical proportions (p_c) of the population to be immunized for eradication
Malaria (P. falciparum in a hyperendemic region)	99%
Measles	90-95%
Whooping cough (pertussis)	90-95%
Fifths disease (human parvovirus infection)	90-95%
Chicken pox	85-90%
Mumps	85-90%
Rubella	82-87%
Poliomyelitis	82-87%
Diphtheria	82-87%
Scarlet fever	82-87%
Smallpox	70-80%

SARS傳染病
流行病學

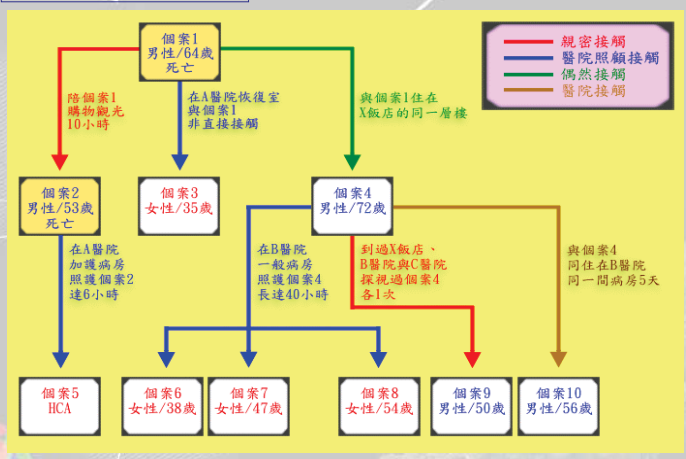
台大預防醫學研究所

陳秀熙教授

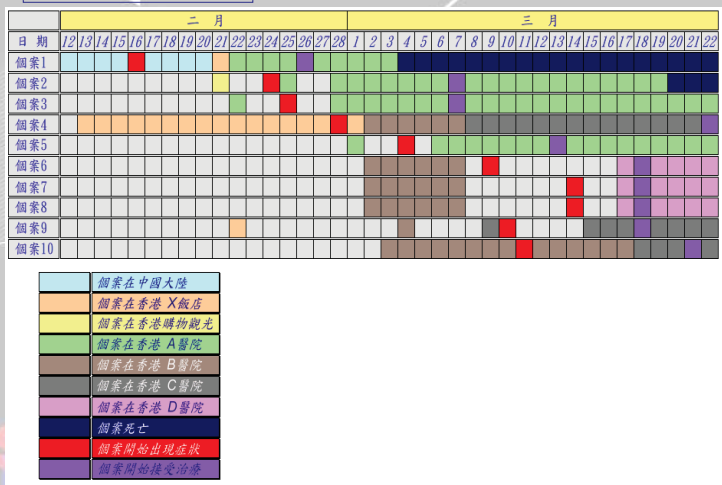
SARS傳染病
醫院、家庭感染途徑

以香港、加拿大等國經驗為例

香港醫院感染接觸史



香港醫院感染接觸史



香港醫院感染簡述

在香港被發現的首例SARS個案，是一名來自中國大陸的腎臟科醫師，然而早在他到達香港(02/21/2003)的前五日，他就已經開始出現症狀了；到達香港之後他還相當正常地跟他的小舅子一同去逛街購物，然而就在隔天，他就感到身體不適而就醫，並且因為呼吸衰竭被直接轉到A醫院的加護病房。

第二例SARS個案，一名香港居民，與前述個案1是連襟關係，就在個案1到達香港後，個案二還陪同他一起逛了10個小時的街。

第三例個案是一名A醫院急診部的護士，他在恢復室與個案一有過約1公尺的接觸，但是並不是診斷個案一的醫療團隊成員，個案三並未與個案一有直接的接觸，並且在當時有戴口罩。

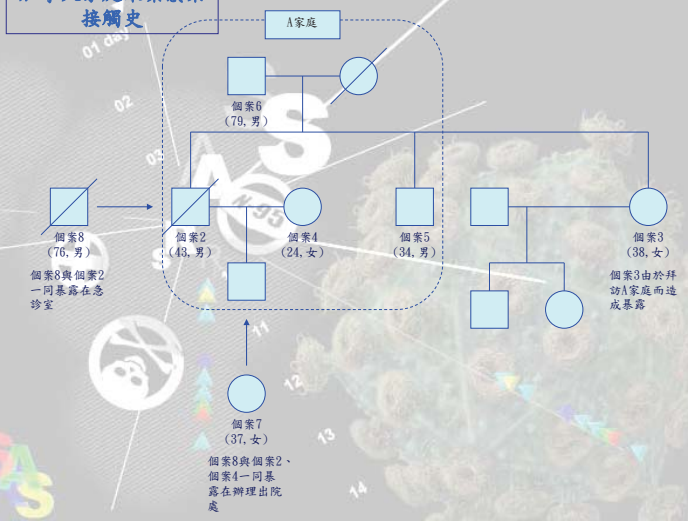
個案五是一名A醫院加護病房的助理，與個案二曾有不安全的接觸(沒有加以防護)將進6小時。個案四是一名加拿大華商，回香港參加02/13/2003的家族聚會，在此之前他有一年的時間沒離開過多倫多，這次回香港其中一天他與個案一住在同一家飯店，但是並不知道她們是否有過直接的接觸，而個案四也想不起來他何時再飯店裡見過個案一。

個案六、七、八三例都是B醫院的的護理人員，個案四在轉往C醫院之前曾因急性肺炎在B醫院住院治療6天。在這段期間裡，一些輪班護士與個案四在病房中有過接觸，上述三名護士在03/03/2003因為替個案四清理失禁穢物時有過較靠近的接觸，當時她們並未戴口罩或隔離衣。

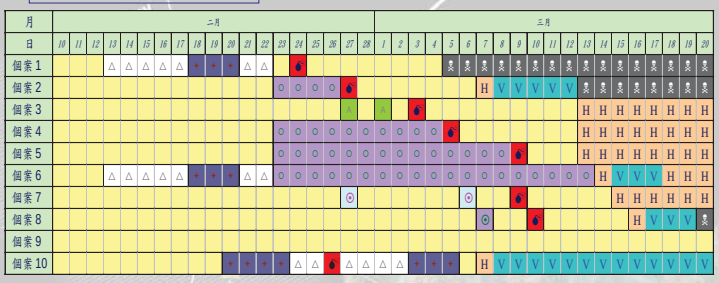
個案九是個案四的外甥，曾經到過X飯店、B醫院與C醫院的加護病房探視過個案四，每次約十分鐘；就在最後一次探視時，個案四正在接受NPPV。個案十因為腎臟切除而與個案四有5天住在同一間病房，在他們兩床中間的那名患者沒有與個案四、十有過接觸。

這些接觸關係，除了個案一外，沒有人有認識有呼吸道疾病的人，也沒有人在三個月內到過中國、越南或新加坡等地。

加拿大家庭聚集感染接觸史



加拿大家庭聚集感染接觸史



圖例	代表意義
△	旅行暴露
+	個案在香港 A 飯店
○	地區暴露
H	這段時間個案在醫院
V	個案使用呼吸器
△	這段時間個案拜訪 A 家
⊙	與個案 2,4 有接觸
☠	個案死亡
●	個案開始出現症狀

加拿大家庭聚集感染接觸史

個案一，78歲婦人，患有有第二型糖尿病與心血管疾病，在回家兩天之後出現發燒、食慾減退、肌肉疼痛、喉嚨痛以及輕微的咳嗽；三天後家庭醫師發現她除了咽喉黏膜出現紅斑外其他一切正常，就開給她口服抗生素之後讓個案一回家。兩天後，個案一出現嚴重的窒息性咳嗽，三天後(03/05/2003)個案一在家中過世，剛好發病後九天，死後並沒有解剖。

個案二是個案一的兒子，43歲，患有第二型糖尿病與高血壓，在他的母親出現症狀的兩天後(02/27/2003)他也出現發燒和大量出汗的症狀，數天後雖然發燒狀況解除，但依然因為持續咳嗽、胸疼及呼吸困難的情況而接受抗生素治療。由於他的症狀並沒有減緩，因此他被送進醫院治療，在他入院後兩天，他的呼吸器宮狀況明顯惡化，於是他開始被插管與使用呼吸器。儘管如此，他仍然在(03/13/2003)過世，正好是發病15天、入院後第6天，死因是多重器官衰竭病逝。

03/08/2003，個案一的家人包括五名成年人及三名小孩(五個月、九歲、十七歲大)在接受胸部X光檢查後，也發現肺部有異常現象，由於經研判SARS可能為個案一及個案二的死因。因此03/13/2003個案四被診斷為SARS患者，個案五、個案六跟個案三被診斷為疑似個案，四人均住院接受治療直到完全康復。

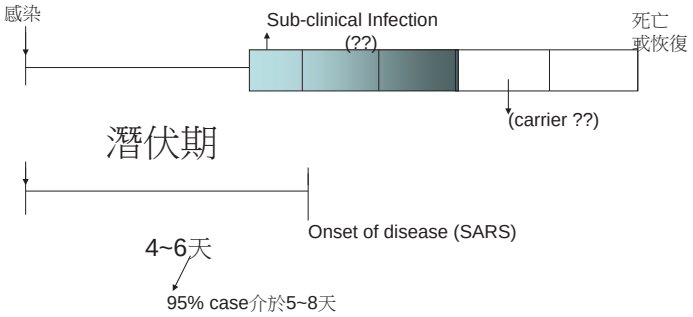
個案七，亞裔的37歲女性內科醫生，由於03/06/2003曾與已發病的個案二及個案四見過面，三天後她也出現嚴重頭痛發燒等症狀，隨後病情加劇並被診斷為SARS病患，經治療後目前已康復。

個案八是一位亞非裔的76歲男性，患有第二型糖尿病及高血壓，於03/07/2003在加護病房過夜觀察時僅與個案二的病床隔著一幅屏風，兩人相距僅約一到二公尺，隔天個案八出院回家，兩天後就出現發燒盜汗等症狀，伴隨而來的是乾咳和呼吸困難，肺部也有浸潤現象；住院後，被診斷為疑似SARS病患，接受插管跟隔離等各種治療，但是仍然在發病後12天病逝(03/21/2003)。

個案九是一位亞非裔的62歲男性，與個案一家族無關。03/14/2003返回多倫多，他就感到胸疼、喉嚨痛跟輕微頭痛，隨後更出現咳嗽、呼吸困難和發燒，就醫後被診斷為疑似SARS病患，目前仍持續治療中。

個案十是一位健康的55歲男性，他與妻子於二月底到三月初期之間到香港旅遊，2月20至24之間下榻香港A旅館，但與個案一、個案六及香港最初的發病者住不同樓層，期間沒有在旅館用餐也沒有拜訪其他客人，03/07/2003返回溫哥華後即發病被診斷為疑似SARS病患住進醫院持續治療中。

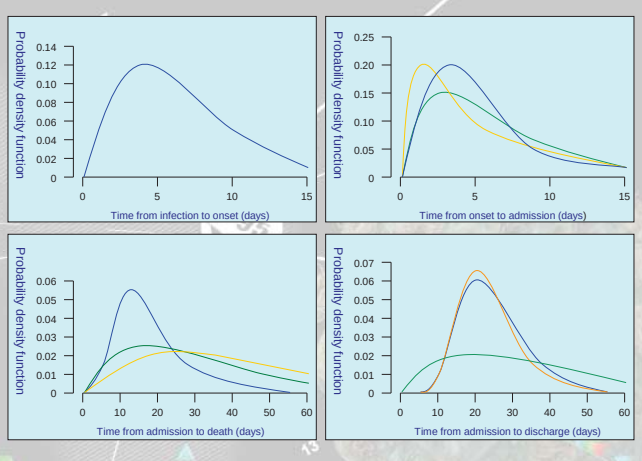
傳染期



慢性病：4天

非慢性病 :7天

SARS潛伏期與傳染期之關係



慢性病&SARS

Infection	→	Onset
糖尿病		2.00倍
高血壓		5.64倍
心血管疾病		1.67倍
任何一種慢性病		4.00倍

P.S. 香港經驗

Cumulative number of cases of severe acute respiratory syndrome reported to the Department of Health, Hong Kong, from 14 March to 1 April 2003

Date	Healthcare workers			Patients, families, and visitors	Cumulative total
	Hospital where outbreak started	Other healthcare workers	Total		
14 March	3	0	3	0	3
15 March	36	0	36	0	36
16 March	36+13*	0	49	0	49
17 March	44+16*	12	72	23	95
18 March	47+17*	20	84	39	123
19 March	54+17*	21	92	58	150
20 March	58+17*	24	99	74	173
21 March	66+17*	27	110	93	203
22 March	68+17*	29	114	108	222
23 March	73+17*	32	122	125	247
24 March	78+17*	34	129	136	265
25 March	82+17*	35	134	156	290
26 March	88+18*	37	143	176	319
27 March	NA	NA	149	221	370
28 March	NA	NA	153	272	425
29 March	NA	NA	156	314	470
30 March	NA	NA	162	368	530
31 March	NA	NA	164	446	610
1 April	NA	NA	168	517	685

NA=Not assessed. *Medical students

R_0=3

香港SARS流行有效再生數推估(R_0)

世代	日期	個案數(I_j)	
初級個案	2/21	1	I_0
第一代	2/21~2/23	9	I_1
第二代	2/24~2/28	106	I_2
第三代	2/29	237	I_3

$$R_0 = \frac{I_0 + \ldots + I_3}{I_0 + I_2} = \frac{353}{116} = 3$$

SARS感染率推估-香港

SARS有效再生數 (R_0)

1個初級個案

↓

n個初級個案

$$R_0 = 3 = 1 + \frac{\text{感染率}}{\text{死亡率}}$$

香港死亡率：2.25%

香港感染率：4.5%

感染數推估-香港

Back-calculation SARS累積個案

= $\int_0^t$ 感染數$(I(s))$×潛伏期分佈$(F(t-s))$

➔推估 *14/March – 1/April*

➔推估 1,284感染數

➔推估 28,533接觸個案

*F(t-s)*是 Weibull分佈

λ： 0.04

γ： 1.62

SARS病毒之探索—新型冠狀病毒

A novel Coronavirus Associated with Severe Acute Respiration Syndrome

New England Journal of Medicine 2003; 348: 1953-1966.

前言

傳染模式：疑似飛沫或接觸傳染

➢傳播速度快

範圍:廣東、越南、香港、加拿大、新加波及台灣等

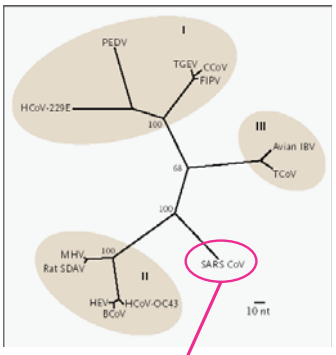
➢跨國際合作

衝擊:家庭成員及健康醫療照護者

➢社會及醫療嚴重衝擊

迫切需要➔快速直接檢驗方法及疫苗開發

第一要件:對於病毒的瞭解



系統演化樹分析結果SARS病毒與原來之冠狀病毒不同種

研究方法

✸生物研究安全性: 第三級生物安全(Biosafety Level 3)

✸一般研究: 一般病毒、細菌、披衣菌、立克次體等等,
排除一般感染可能性

✸病毒培養及純化: 幼鼠及多種細胞株進行培養

✸血清免疫法:從CDC血清庫中受過已知人類冠狀病毒感染之血清進行檢驗比對

✸分子技術: 以Real Time PCR及RT-PCR進行序列定序

✸核甘酸序列以UNIX系統之CLUSTAL進行序列比對，並進行系統發展演化樹(phylogenetic tree)進行距離及極大近似值之運算

Urbani strain SARS-associated Coronavirus

找出作為SARS診斷病毒基因之前端子(primer):

Cor-p-F2(+)5'CTAACATGCTTAGGATAATGG3',

Cor-p-F3(+)5'GCCTCTCTTGTTCTTGCTCGC3',及

Cor-p-R1 (–) 5'CAGGTAAGCGTAAAACTCATC3'

✸2003/2/28經由越南河內法國醫院通報, 院內多人染病
WHO- Dr. Carlo Urbani前往了解並5人小組進駐醫院
✸2003/3/9向WHO通報為新型疾病—急性非典型肺炎
✸2003/3/29 Dr. Carlo Urbani死於SARS感染

紀念因發現SARS而死亡之Dr. Carlo Urbani，研究團隊決定以”Urbani”命名該初次定序出來之序列

流行病學方法論及實驗
(The Methods of Epidemiology and Practices)

護理與基因遺傳

授課教師：陳秀熙 教授/許辰陽 博士

台灣大學流行病學與預防醫學研究所

族群基因池(gene pool)與基因頻率

- 例: 假設今有200對男女，其中AA為128對，Aa為64對，在這400位個體中含A是320而a是80所以

  A之頻率=80%, a之頻率=20%

如此我們想知道等位基因(Alleles)A及a之頻率是否會隨時間演化而不同，如同族群基因池(族群內所有成員之全部基因座所有等位基因)之演化相同

Hardy-Weinberg Law of Equilibrium (哈溫平衡定律)

- 英國數學家 G.H. Hardy與德國物理學家W. Weinberg 在1908年所提出。

		Romeo	
		A	A
Juliet	A	p^2	pq
	a	pq	q^2

- p及q分別代表基因A及a之頻率，則基因頻率分布AA:Aa:aa=p^2:2pq:q^2

哈溫平衡定律之檢定

- 假設「愛」之基因為共顯基因(Codominant Gene)，分別為AA,AB及BB，某地抽樣之樣本其三種型態分別有233,385及129，是否符合H-W平衡?

	AA	AB	BB	Total
觀察值	233	855	129	747
期望值	242.37	366	138.37	747
$\frac{(觀察值-期望值)^2}{期望值}$	0.3622	0.9588	0.6345	1.9555
	(p^2=0.3245)	(2pq=0.4903)	(q^2=0.1852)	

$A\% = 851/1494 = 0.5696$

$B\% = 643/1494 = 0.4304$

$\chi^2_{(1)} = 1.9555, p = 0.162$

符合H-W之平衡定律

哈溫定律與自體隱性遺傳模式

- 他們居住地區族群有苯酮氏尿症，Romeo與Juliet結婚前欲知苯酮氏尿症帶原之機率？

- 哈溫定律對於計算隱性遺傳模式之帶原基因機率非常重要 $q^2 = I, q = \sqrt{I}, 2pq = 2\times\sqrt{I}$

I:疾病發生率

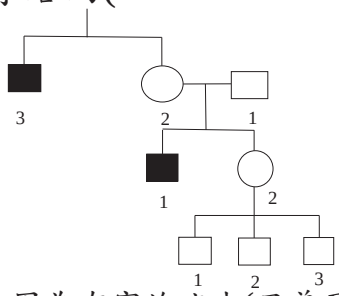
疾病發生率(q^2)	基因頻率(q)	帶原率(2pq)
1/100	1/10	2/11
1/1000	1/32	1/16
…	…	…
1/100000	1/316	1/158

哈溫定律與隱性遺傳模式

- 實例

發生率(q^2)	帶原率 2pq	
1/400	1/10	Sick Cell disease(鐮刀型貧血) Thalassemia (地中海貧血)
1/10000	1/50	Phenylketonuria (苯酮氏尿症)

貝氏定理(Bayes' Theorem)與遺傳諮詢(Genetic Counseling)



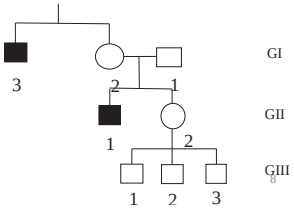
- GII Juliet 因為有家族病史(兄弟及舅舅感染)，想知道她帶某疾病基因之機率？

貝氏定理(Bayes' Theorem)與遺傳諮詢(Genetic Counseling)

- GI (2) 必定是帶原(Carrier, 簡寫C)，所以GII (2) 帶原之機率為1/2 (事前機率來自上一代)，且三位兒子都是健康者 $\bar{D}=(\bar{D}_1,\bar{D}_2,\bar{D}_3)$

$$P(C\mid\bar{D})=\frac{P(C)\times P(\bar{D}\mid C)}{P(C)P(\bar{D}\mid C)+P(\bar{C})P(\bar{D}\mid\bar{C})}$$

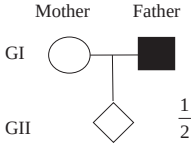
$$=\frac{\frac{1}{2}\times\frac{1}{8}}{\frac{1}{2}\times\frac{1}{8}+\frac{1}{2}\times 1}=\frac{1}{9}$$



自體顯性遺傳(Autosomal Dominant Inheritance)

例: 視網膜母細胞瘤(Retinoblastoma)、硬化性結節症(tuberous sclerosis)

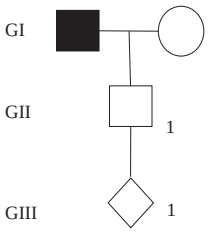
1. 通常是異型合子(heterozygote)
2. 相關疾病表現常受到穿透力(penetrance, 簡寫為p)程度之影響



Romeo 患有視網膜母細胞瘤，欲知小孩會得視網膜母細胞瘤的機率？

$\frac{1}{2}p$(遺傳下基因且致病之機率)

自體顯性遺傳(Autosomal Dominant Inheritance)



H : *Heterozygote*

$$P(\text{H for GII(1)}\mid\bar{\text{D}}\text{ for GII(1)};p)$$

$$=\frac{P(\text{H})\times P(\bar{\text{D}}\text{ for GII(1)}\mid\text{H})}{P(\bar{\text{D}}\text{ for GII(1)})}$$

$$=\frac{\frac{1}{2}(1-p)}{\frac{1}{2}(1-p)+\frac{1}{2}}=\frac{1-p}{2-p}$$

The risk of D for GIII (1)

$$=\frac{1-p}{2-p}\times\frac{1}{2}\times p=\frac{p(1-p)}{2(2-p)}$$

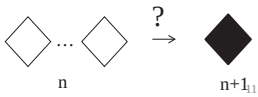
自體顯性遺傳(Autosomal Dominant Inheritance)

致病力 (p)	GIII (1) 致病之機率(p)
0.2	4.4%
0.4	7.5%
0.6	**8.6%**
0.8	6.7%

- 若第三代(GIII)有n個健康者，則第(n+1)個得病之機率 $=W\times\frac{p}{2}$, n=3 其機率為4.7%

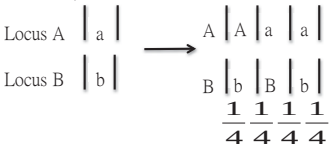
	H for GII (1)	$\bar{H}$ for GII(1)
事前機率P(H)	(1-p)/(2-p)	1/(2-p)
$P(\bar{D}_1,\bar{D}_2,\bar{D}_3\mid\text{H})$	$[1/2+(1-p)/2]^2$	1
聯合機率	(1-p)/(2-p) $[1/2+(1-p)/2]^2$	1/(2-p)

事後機率 $P(H\mid\bar{D}_1,\bar{D}_2,\bar{D}_3)=\frac{(1-p)\left(1-\left(\frac{p}{2}\right)^n\right)}{(1-p)\left(1-\left(\frac{p}{2}\right)^n\right)+1}=W$

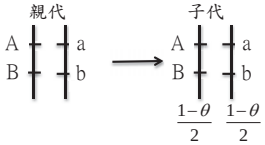


連鎖(Linkage)及互換(Recombination)

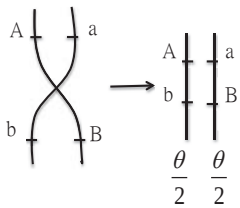
1. 基因座(Loci)在不同染色體



2. 基因座(Loci)在相同染色體且不互換

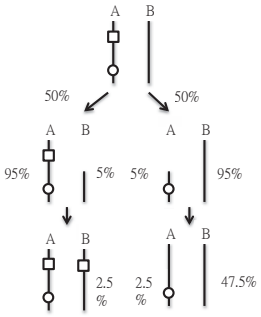


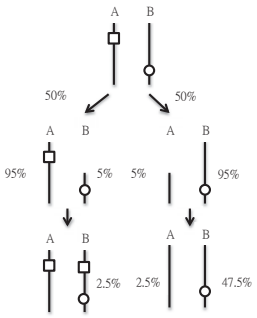
3. 基因座(Loci)在相同染色體但有互換



同位及不同位之連鎖及互換

1. 同位(Coupling)　　2. 不同位 (Repulsion)



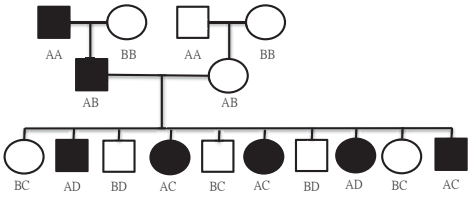


親代與子代基因標記(A及a)和疾病標記(D及d)之互換

親代基因型	傳至子代之半套型			
	da	dA	Da	DA
da|da	1	0	0	0
da|dA	1/2	1/2	0	0
dA|dA	0	1	0	0
da|DA	(1-θ)/2	θ/2	θ/2	(1-θ)/2
da|Da	θ/2	(1-θ)/2	(1-θ)/2	θ/2
dA|DA	0	1/2	0	1/2
Da|Da	0	0	1	0
Da|DA	0	0	1/2	1/2
DA|DA	0	0	0	1

只有da|DA 或dA|Da (雙異型合子) 對於連鎖才有訊息

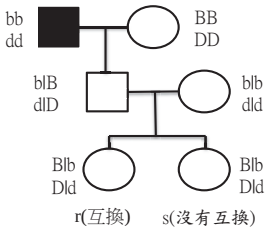
連鎖分析(Linkage Analysis)-Lod Score Analysis



1) 標記基因A和疾病連鎖，因此互換率(θ)≈0，如此標記基因A和疾病一起傳下來第三代10個子女之機率$(1-\theta)^{10}\approx 1$
2) 如果沒有連鎖互換率(θ)=1/2，則其機率為$(1/2)^{10}$

$\log_{10}\frac{1}{(1/2)^{10}} \geq 3$　一般而言Lod>3表示有連鎖

二項分佈方法



如果現觀察100次減數分裂，35次互換則觀察互換次數(r)次會遵循二項分佈

$r \sim Bin(100,\theta)$二項分佈

$P(r\,|\,100)=\binom{100}{r}\theta^{r}(1-\theta)^{n-r}$

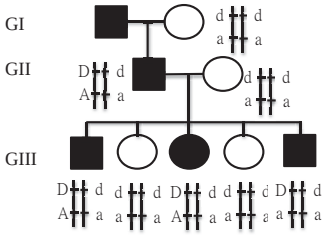
$\hat{\theta}=0.35$

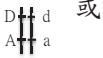
虛無假設 $H_0:\theta=\frac{1}{2}$

使用McNemar test

$$\chi^2=\frac{\left(|\text{互換次數}-\text{無互換次數}|-\frac{1}{2}\right)^2}{100}=8.7(df=1),\qquad p<0.0001\text{表示連鎖}$$

祖父母代訊息未知顯性遺傳互換



但如果祖母之基因型態未知，則父親之基因型態就有可能是  或 

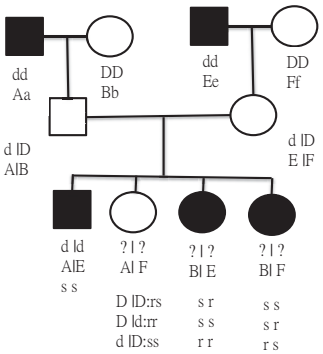
則其互換訊息為r=4，如此二項方法就無法使用。

$\frac{1}{2}[\theta(1-\theta)^4+\theta^4(1-\theta)]$

代入不同 θ 可得到下列圖在父親是不同基因型態

產生雙峰在θ=0.21及0.79, 若考慮整體貢獻　$\hat{\theta}=0.21$

隱性遺傳互換



$(1-\theta)^2\times\left[\theta^2+\theta(1-\theta)+(1-\theta)^2\right]^2\times\left[2\theta(1-\theta)+(1-\theta)^2\right]$

連鎖不平衡檢定 (Transmission disequilibrium test,TDT)

下傳基因	未下傳基因		
	A	a	Total
A	22	78	100
a	46	54	100
Total	68	122	200

下傳基因個數

	A	a	Total	χ^2	Significance (p)
觀察值	78	46	124	8.26	0.004
期望值	62	62			

$$\chi^2_{td} = \frac{(78-46)^2}{124} = 8.26,\ p = 0.004$$

DNA序列與潛在馬可夫模式

- ACTATAGGACTTAGCCTT

  $\pi(A,C)$ $\pi(T,A)$ $\pi(T,T)$

- 所有可能轉移之機率

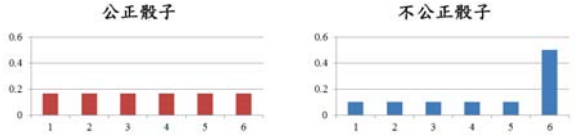
$$P = \begin{array}{c} \\ A \\ C \\ G \\ T \end{array} \begin{array}{c} \begin{array}{cccc} C & G & T \end{array} \\ \begin{pmatrix} \pi(A,A) & \pi(A,C) & \pi(A,G) & \pi(A,T) \\ \pi(C,A) & \pi(C,C) & \pi(C,G) & \pi(C,T) \\ \pi(G,A) & \pi(G,C) & \pi(G,G) & \pi(G,T) \\ \pi(T,A) & \pi(T,C) & \pi(T,G) & \pi(T,T) \end{pmatrix} \end{array}$$

- 平衡分佈
  - 因DNA序列通常長所以一般會有 A,T,C,G 平衡分布 (μ)存在

$$\mu = [\mu(A), \mu(C), \mu(G), \mu(T)]$$

$$\mu_N = [0.3, 0.3, 0.2, 0.2]$$

潛在馬可夫模式 (Hidden Markov Model)



$$Z = \begin{cases} 0 & \text{公正} \\ 1 & \text{不公正} \end{cases} \qquad Y \in \{1,2,3,4,5,6\}$$

(Z,Y)構成一個潛在馬可夫模式,假設Z(潛在變項)之轉移機率如下

$$P = \begin{pmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{pmatrix}$$

$P(Y_n = i \mid Z_n = 1) = 0.1$對$i = 1\ to\ 5$，$P(Y_n = 6 \mid Z_n = 1) = 0.5$

$P(Y_n = i \mid Z_n = 0) = \frac{1}{6}$對$i = 1\ to\ 6$

CpG 島在DNA序位

- Y{TAGTGGAATGCGACG}-DNA序位
- Z 0 0 00 0 0 0 00 01 1…
- DNA序位決定如果一個公正骰子有兩面是不用的而其他四面分別是ACTG，而且並非是公正骰子，因為一般CG對出現頻率低，因此通常如果有CpG島出現在外顯遺傳(epigenetic)扮演重要角色

- 我們想問若在不公正骰子下(CpG島)，無數次遊戲(n)後，出現點數之機率

$$\lim_{n\to\infty} P(Y_n = i \mid Z_0 = 1)$$

$$= P(Y_n = i \mid Z_n = 0)v(0) + P(Y_n = i \mid Z_n = 1)v(1)$$

$v = (v(0), v(1)) = (\frac{5}{6}, \frac{1}{6})$ 因此CpG島出現

對$i = 1$ to 5

$$\lim_{n\to\infty} P(Y_n = i \mid Z_0 = 1) = \frac{1}{6}\times\frac{5}{6} + \frac{1}{10}\times\frac{5}{6} = \frac{7}{45}$$

對$i = 6$

$$\lim_{n\to\infty} P(Y_n = 6 \mid Z_0 = 1) = \frac{1}{6}\times\frac{5}{6} + \frac{1}{2}\times\frac{1}{6} = \frac{2}{9}$$

DNA序列與潛在馬可夫模式

- 但如果在CpG區域(不公正骰子，Z=1)，則

$$\mu_C = [0.1, 0.1, 0.4, 0.4]$$

- 如此潛在狀態(Z)之轉移機率

$$P_H = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix} \end{array}$$

- 由此可以計算AGCCG序列之機率

DNA序列與潛在馬可夫模式

$\pi_{Z_2}(A,C)$ $\pi_{Z_8}(T,A)$

觀察到*DNA*序列: A G C C G A T A G G A C T T A G C C T T

隱藏狀態轉移 : $Z_1 Z_2 Z_3 \ldots$

$H(Z_1, Z_2)$

假定起始一定由公正骰子開始P(*A*|*Z*=0)，則會有16種可能

1. 如果AGCCG全部由Z=0產生

$$P(A|Z_1=0)\times P(G|Z_2=0)\times P(C|Z_3=0)\times P(C|Z_4=0)\times P(G|Z_5=0)$$
$$\times P(Z_2=0|Z_1=0)\times P(Z_3=0|Z_2=0)\times P(Z_4=0|Z_3=0)\times P(Z_5=0|Z_4=0)$$
$$\times P(Z_1=0)$$
$$=0.3\times 0.2\times 0.2\times 0.2\times 0.2\times 0.2\times 0.8\times 0.8\times 0.8\times 0.8=1.97\times 10^{-4}$$

DNA序列與潛在馬可夫模式

2. AGCCG除第一位置外由Z=1產生

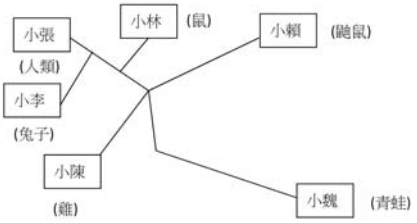
$$P(A|Z_1=0)\times P(G|Z_2=1)\times P(C|Z_3=1)\times P(C|Z_4=1)\times P(G|Z_5=1)$$
$$\times P(Z_5=1|Z_4=1)\times P(Z_4=1|Z_3=1)\times P(Z_3=1|Z_2=1)\times P(Z_2=1|Z_1=0)\times P(Z_0=1)$$
$$=0.3\times 0.4\times 0.4\times 0.4\times 0.4\times 0.7\times 0.7\times 0.7\times 0.2=5.27\times 10^{-4}$$

…

找出機率最大之潛在隱藏狀態

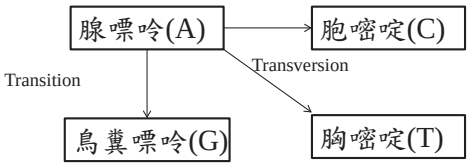
種系遺傳(Phylogenetics)

小魏	GCTTGACTTCTGAGGTT
小陳	GCGTAACTTCACATGAT
小張	GCGTCACTTGAGACGCT
小李	GCGTCACTTGAGACGCT
小林	GCGTCACTTGACAGGCT
小賴	GCGTCACTTGAGACGCT

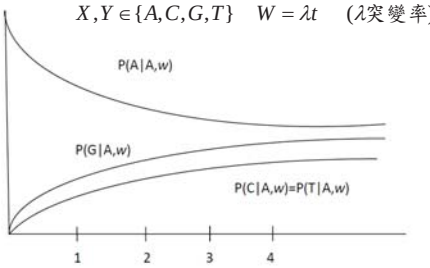


Branching Order
Branching Length

核酸置代數學模式



$X, Y \in \{A,C,G,T\}$ $W=\lambda t$ (λ突變率)



核酸置代數學模式

$$p(t)=\begin{bmatrix} P(y(t)=A|y(0)=A,...,y(t)=A|y(0)=T) \\ P(y(t)=G|y(0)=A,...,y(t)=A|y(0)=T) \\ P(y(t)=C|y(0)=A,...,y(t)=A|y(0)=T) \\ P(y(t)=T|y(0)=A,...,y(t)=A|y(0)=T) \end{bmatrix}$$

$$p(t)=e^{Rt}$$

$$R=\begin{array}{c} \\ A \\ C \\ G \\ T \end{array}\begin{array}{c} \begin{array}{cccc} A & C & G & T \end{array} \\ \begin{bmatrix} -2\beta-\alpha & \beta & \alpha & \beta \\ \beta & -2\beta-\alpha & \beta & \alpha \\ \alpha & \beta & -2\beta-\alpha & \beta \\ \beta & \beta & \alpha & -2\beta-\alpha \end{bmatrix} \end{array}$$

核酸置代數學模式

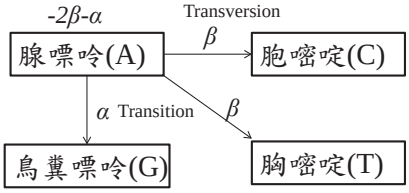
$$p(t)=e^{Rt}=\begin{bmatrix} d(t) & f(t) & g(t) & f(t) \\ f(t) & d(t) & f(t) & g(t) \\ g(t) & f(t) & d(t) & f(t) \\ f(t) & g(t) & f(t) & d(t) \end{bmatrix}$$

$$f(t)=\frac{1}{4}(1-e^{w})$$

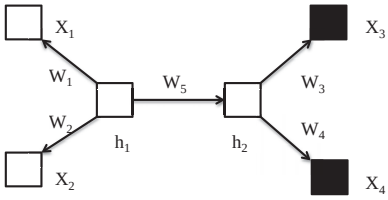
$$g(t)=\frac{1}{4}(1+e^{wt}-2e^{-\frac{r+1}{2}w})$$

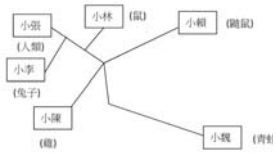
$$d(t)=1-2f(w)-g(w)$$

$$w=4\beta t$$



種系演化樹推估 (Phylogenetic trees)





$$P(x_1,x_2,x_3,x_4,h_1,h_2 \mid W,S)$$
$$= P(x_1 \mid h_1,w_1)P(x_2 \mid h_1,w_2)P(h_2 \mid h_1,w_5)$$
$$P(x_3 \mid h_2,w_3)P(x_4 \mid h_2,w_4)\pi(h_1)$$

我們可以改變箭頭，如$h_2 \rightarrow h_1$則中間條件機率變成$P(h_1|h_2,w_5)$如果轉移機率是可逆的，則兩者會相同

Polygenetic Inheritance

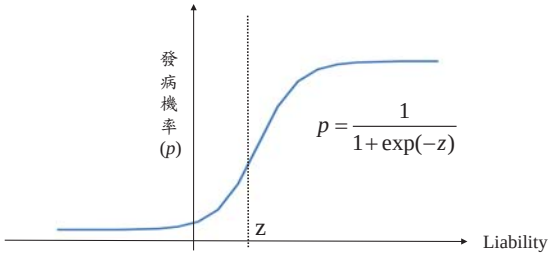
1. 不遵循孟德爾定律
2. 多基因參與但每個具加成性少許貢獻 (small additive contribution)
3. 多因性遺傳疾病(如:IDDM及Schizophrenia)
4. 通常遵循常態分佈
5. 許多基因少許加成變異所得到的整體性狀變異稱為heritability

Polygenetic Inheritance

$$y = \bar{y} + h^2(x - \bar{x})$$

- 例: Romeo (190cm) and Juliet (175cm) 欲預測所生兒子及女兒身高，假定h^2=0.8
  - ➢y (兒子)=176.5+0.8(189.5-176.5)

閾值模式(Liability-Threshold model)



- 通常多基因遺傳疾病是由基因與環境因子共同影響，以數學是來表示

  $Z = \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p x_p$, ($\beta_1$-$\beta_p$是加權係數)
- 因子越多，Z之分數越容易呈現常態分佈

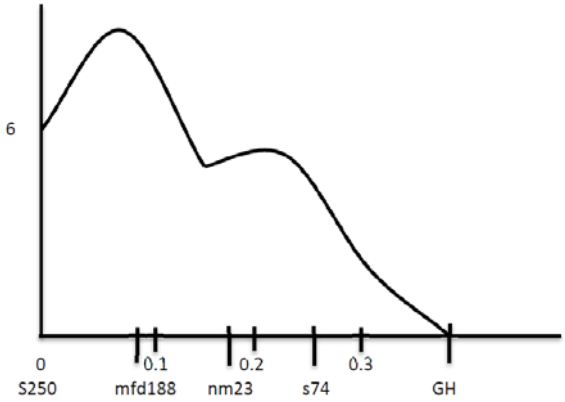
Risk Assessment Using Epidemiological Data

- Does Julia have the breast cancer??

危險性評估(Risk assessment)--族群乳癌危險分層

極高危險羣 (第一類)	高危險羣(4%) (第二類)	略高或平均危險羣 (第三類)	低於平均危險羣 (第四類)
(<1%)	(4%)	95%	
1. 罹患卵巢癌高危險羣 2. 二位一等親或二等親家屬罹患乳癌或卵巢癌加下列條件 (1) 另一位家族有乳癌或卵巢癌 (2) 乳癌發病年齡在35歲以前 (3) 雙側乳癌 (4) Ashkenazi Jewish後裔 (5) 男性 3. 一等親家屬乳癌發病年齡在50歲以前及另一位具有sarcoma (bone/soft tissue)在50歲以前 4. 家族成員出現BRCA	1. 一等親家屬乳癌發病年齡在50歲以前 2. 二位一等親家屬罹患乳癌 3. 二位一等親家屬罹患乳癌，但其中一位發病年齡在50歲以前 4. 上述1~3條件並不含第一類極高危險羣	1. 無家族病史-以其他傳統及個人生理特徵危險因子分層 2. 家族病史 (10~15%) (1) 一位一等親罹患乳癌但發病年齡超過50歲 (2) 二位二等親罹患乳癌但發病年齡皆超過50歲	以相關危險因子進行分類

BRCA1 gene



37

Autosomal Dominant Model
(自體顯性遺傳模式)

Mean Age for genotype AA (Aa)

$\mu_{AA}\left(-\mu_{Aa}\right) = 55.435\ (1.742)$

Mean Age for genotype aa

$\mu_{aa} = 68.990\ (1.532)$

Standard deviation for genotype AA, Aa, and aa

$\sigma_{AA}\left(=\sigma_{Aa}=\sigma_{aa}\right)$=15.387 (0.669)

Penetrance for genotype AA and Aa

$\lambda_{AA}\left(=\lambda_{Aa}\right) = 0.928\ (0.163)$

Penetrance for genotype aa

$\lambda_{aa} = 0.100\ (0.009)$

Gene frequency

$q = 0.003\ (0.001)$

38

Mathematical Algorithm

$$\Pr\left(\text{Women A affected at age } x \mid \text{Relative (R) is affected at age } x_r\right)$$

$$= \sum_j \sum_k \left\{ \Phi\left(x; \mu_j, \sigma_j^2\right) \times \Pr\left(\text{A has genotype j} \mid \text{R has genotype k}\right) \right.$$

$$\left. \times \frac{\Pr\left(\text{R has genotype k}\right) \times \frac{1}{\sigma_k} \exp\left(-\frac{\left(x_r - \mu_k\right)^2}{2\sigma_k^2} \right)}{\sum_k \left[\Pr\left(\text{R has genotype k}\right) \times \frac{1}{\sigma_k} \exp\left(-\frac{\left(x_r - \mu_k\right)^2}{2\sigma_k^2} \right) \right]} \right\}$$

Relative: mother or sister

$\Phi(x; \mu, \sigma^2)$: the cumulative distribution function of normal distributed random variable, x, with mean μ and variance σ^2

39

Breast Cancer Risk Assessment with Empirical Epidemiological Data Assuming Autosomal Dominant Mode

Age of woman (yr)	Age of onset (year) -*First-degree relative*					
	20-29	30-39	40-49	50-59	60-69	70-79
29	.007	.005	.003	.002	.002	.001
39	.025	.017	.012	.008	.006	.005
49	.062	.044	.032	.023	.018	.015
59	.116	.086	.064	.049	.040	.035
69	.171	.130	.101	.082	.070	.062
79	.211	.165	.132	.110	.096	.088
-Second-degree relative						
29	.004	.003	.002	.001	.001	.001
39	.014	.010	.007	.006	.005	.004
49	.035	.027	.021	.017	.017	.013
59	.070	.056	.045	.038	.038	.032
69	.110	.090	.076	.067	.067	.058
79	.142	.120	.104	.094	.094	.083

40

Genetic Testing

- Direct Testing
- Indirect Testing and Linkage Analysis
  RFLPs
  Microsatellite Markers
  SNPs
- LOD score Analysis
  Haplotype Analysis
- Linkage Disequilibrium – Genome-wide scans

41

Cumulative probability of breast cancer by given age for women who are heterozygous for a dominant susceptibility allele and for control

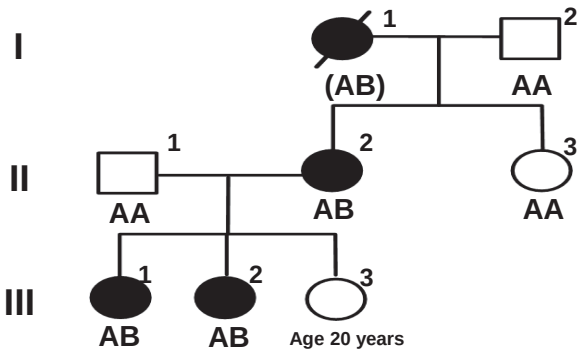
Age in years		
	Heterozygotes	Controls
20-29	0.02	0
30-39	0.14	0
40-49	0.38	0.01
50-59	0.55	0.03
60-69	0.67	0.05
70-79	0.95	0.08
80+	1	0.12

(1) If III 3 has AB genotype the probability that III 3 has inherited a breast cancer susceptibility allele=0.88×0.98 (AB genotype) +0.12 ×0.5=0.92

(2) If III3 has AA genotype, this probability is 0.08

42

Using DNA Linkage Analysis



1. Posterior probability that disease is linked with BRCA1=0.88
2. Posterior probability that disease is not linked with BRCA1=0.12

	Disease linked with BRCA1		Disease not linked with BRCA1
	Link with A in I 2	Link with B in I 2	
Prior	0.25	0.25	0.5
Likelihood			
II 2	θ	1-θ	0.5
II 3	θ	1-θ	0.5
III 1	1-θ	1-θ	0.5
III 2	1-θ	1-θ	0.5
Prior × Likelihood	$0.25\theta^2(1-\theta)^2$	$0.25(1-\theta)^4$	0.5^5
	a	b	c

Posterior

[1] Disease linked with BRCA1 (a+b) / (a+b+c)

[2] Disease not linked with BRCA1 c / (a+b+c)

Set θ=0.02

[1] 88%

[2] 12%

Cumulative probability of breast cancer by given age for women who are heterozygous for a dominant susceptibility allele and for control

Age in years		
	Heterozygotes	Controls
20-29	0.02	0
30-39	0.14	0
40-49	0.38	0.01
50-59	0.55	0.03
60-69	0.67	0.05
70-79	0.95	0.08
80+	1	0.12

(1) If III 3 has AB genotype the probability that III 3 has inherited a breast cancer susceptibility allele=0.88×0.98 (AB genotype) +0.12 ×0.5=0.92

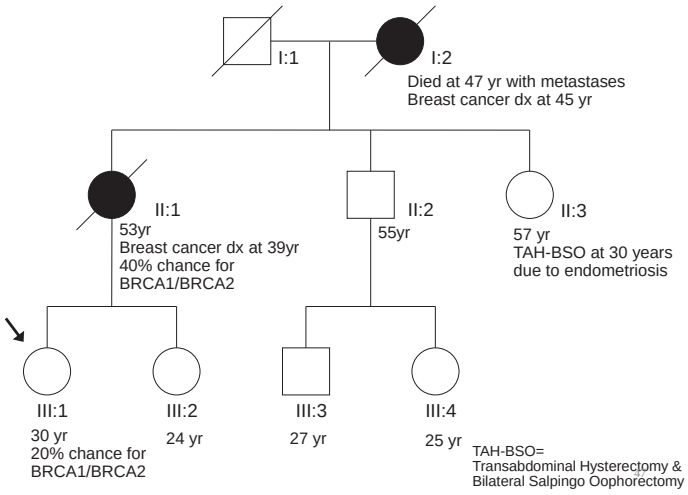
(2) If III3 has AA genotype, this probability is 0.08

The probability of BRCA1/2 mutation

BC Dx before 50 years in relatives	Ovarian Cancer in relatives	Bilateral Cancer	Breast Cancer	Probability for BRCA1/2 or BRCA2
·				25%
·			·	40%
		·		51%
·		·	·	76%
	·			35%
	·		·	35%
	·	·		71%
	·	·	·	71%
·	·			35%
·	·		·	59%
·	·	·		71%
·	·	·	·	89%

Note: Based on analysis of women with at least 1 first- or second-degree relative with ovarian cancer or breast cancer before 50 yr of age.

Cancer genetic susceptibility flow diagram



Polygenes, Risk Prediction, and Targeted Prevention of Breast Cancer

N Engl J Med, Vol. 358, No. 26. (26 June 2008), pp. 2796-2803.

Polygenes, Risk Prediction, and Targeted Prevention of Breast Cancer

N Engl J Med, Vol. 358, No. 26. (26 June 2008), pp. 2796-2803.

Table 1. Established Common Breast-Cancer Susceptibility Alleles.*

dbSNP No.	Gene†	Chromosome	Risk-Allele Frequency‡	Relative Risk per Allele‡	Fraction of Total Variance in Risk Explained§ (%)	Population Attributable Risk§ (%)	Study
rs2981582	*FGFR2*	10q	0.38	1.26	1.7	19	Easton et al.,[26] Hunter et al.[27]
rs3803662	*TNRC9, LOC643714*	16q	0.25	1.20	0.9	10	Easton et al.[26]
rs889312	*MAP3K1*	5q	0.28	1.13	0.4	7	Easton et al.[26]
rs3817198	*LSP1*	11p	0.30	1.07	0.1	4	Easton et al.[26]
rs13281615	None known	8q	0.40	1.08	0.2	6	Easton et al.[26]
rs13387042	None known	2q	0.50	1.20	1.2	19	Stacey et al.[28]
rs1053485	*CASP8*	2q	0.86	1.13	0.3	20	Cox et al.[25]

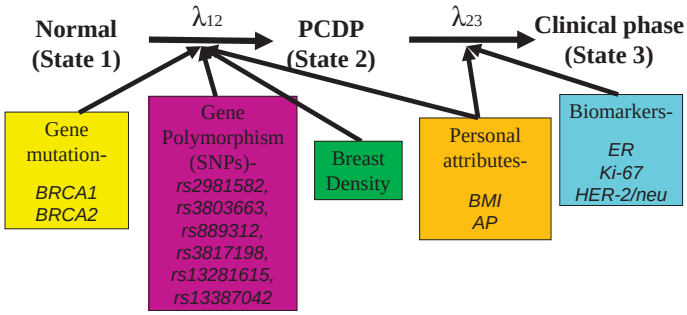
* *CASP8* denotes caspase 8, dbSNP database of single-nucleotide polymorphisms, *FGFR2* the fibroblast growth factor receptor 2 gene, *LOC643714* a hypothetical protein LOC643714, *LSP1* lymphocyte-specific protein 1, *MAP3K1* mitogen-activated protein kinase kinase kinase 1, and *TNRC9* trinucleotide repeat containing 9.
† These genes are within the linkage-disequilibrium block or blocks defined by the associated variant and are plausible candidates for the causal gene.
‡ Values are from published data cited in the Study column.
§ See the Supplementary Appendix for details.

49

Individual-tailored screening –integrate gene, polygenetic, conventional risk factors and biomarkers into multistate model

50

Risk Score (Genomic and Clinical Attributes) for Multi-step Progression of Breast Cancer

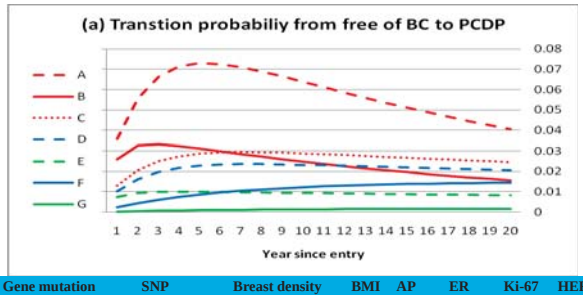


$$\lambda_{12} = \lambda_{12j} \times \exp\left(\sum_{i=1}^{3} \beta_i \times \text{Breast Density} + \sum_{i=4}^{10} \beta_i \times \text{SNP}_i + \beta_{11} \times \text{BMI} + \beta_{12} \times \text{Age at first pregnancy}\right)$$

j=0,1,2 for noncarriers, BRCA1 and BRCA2 carriers

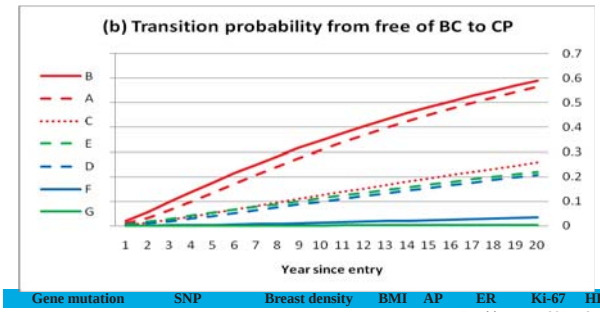
$$\lambda_{23} = \lambda_{230} \times \exp\left(\beta_{13} \times \text{BMI} + \beta_{14} \times \text{Age at first pregnancy} + \beta_{15} \times \text{ER} + \sum_{i=16}^{17} \beta_i \times Ki67 + \sum_{i=18}^{19} \beta_i \times HER2\right)$$

51



	Gene mutation	SNP	Breast density	BMI	AP	ER	Ki-67	HER-2
A	BRCA1 carrier	High risk-alleles	Heterogeneously	>23	>25	Positive	<10%	0 or 1+
B	BRCA1 carrier	High risk-alleles	Heterogeneously	>23	>25	Negative	10-30%	2+
C	BRCA1 carrier	Low risk-alleles	Heterogeneously	>23	>25	Positive	<10%	0 or 1+
D	Noncarrier	High risk-alleles	Heterogeneously	>23	>25	Positive	<10%	0 or 1+
E	Noncarrier	High risk-alleles	Heterogeneously	>23	>25	Negative	10-30%	2+
F	Noncarrier	High risk-alleles	Heterogeneously	<=23	<=25	Positive	<10%	0 or 1+
G	Noncarrier	Low risk-alleles	Almost entirely fat	<=23	<=25	Positive	<10%	0 or 1+

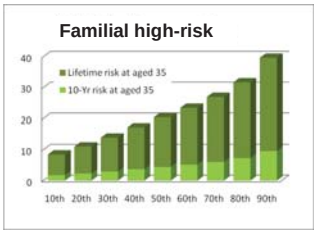
52

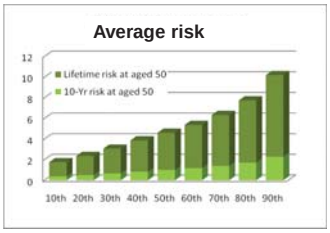


	Gene mutation	SNP	Breast density	BMI	AP	ER	Ki-67	HER-2
A	BRCA1 carrier	High risk-alleles	Heterogeneously	>23	>25	Positive	<10%	0 or 1+
B	BRCA1 carrier	High risk-alleles	Heterogeneously	>23	>25	Negative	10-30%	2+
C	BRCA1 carrier	Low risk-alleles	Heterogeneously	>23	>25	Positive	<10%	0 or 1+
D	Noncarrier	High risk-alleles	Heterogeneously	>23	>25	Positive	<10%	0 or 1+
E	Noncarrier	High risk-alleles	Heterogeneously	>23	>25	Negative	10-30%	2+
F	Noncarrier	High risk-alleles	Heterogeneously	<=23	<=25	Positive	<10%	0 or 1+
G	Noncarrier	Low risk-alleles	Almost entirely fat	<=23	<=25	Positive	<10%	0 or 1+

53

Ten-year risk and lifetime risk of Breast Cancer

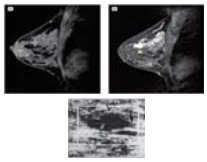




54

The recommend start screening age and inter-screening intervals in two types of screening at the different percentile of Risk Score

Percentile	Start screening age	Inter-screening interval (year)†	Alternative imaging tool (Threshold/Study) §
90-100	29	0.4 (4.8 months)	M+MRI (94th / Kuhl's Study)
80-90	34	1	M+U (85th / Berg's study)
70-80	39	1.5 → Reduce False Negative	M+U (76th / Kuhl'study)
60-70	44	2	
50-60	**50**	3	
40-50	57	4	
30-40	60	8 → Reduce False Positive	
20-30	60	10	
10-20	60	10	
0-10	60	10	



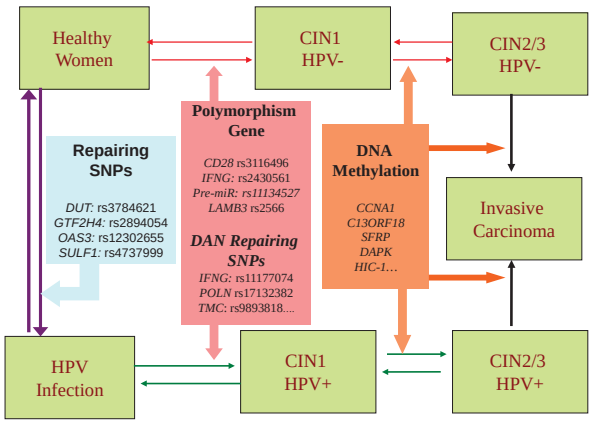
M: Mammography
U: Ultrasound

Cost-effectiveness Analysis

Screening strategy	RR	ICER
No screening	1.00	---
Universal Screening, 1-yearly	0.64	74,673
Universal Screening, 2-yearly	0.76	57,104
Universal Screening, 3-yearly	0.82	51,231
Risk-score approach with varying inter-screening intervals	0.80	59,624
Risk-score approach with alternative imaging technique, 3-yearly	0.80	46,077

** Assuming 70% attendance rate*

Public Health Genetics-Application to Cervical Cancer Screening



$$Inc_{CIN1} = Baseline_{CIN1} \times \exp(\sum_{i=1}^{17} SNP_i)$$

$$\lambda_{CIN1 \to CIN2/3} = Baseline_{CIN1 \to CIN2/3} \times \exp(\sum_{i=1}^{9} Methylation_i)$$

$$\lambda_{CIN2/3 \to InvCa} = Baseline_{CIN2/3 \to InvCa} \times \exp(\sum_{i=1}^{7} Methylation_i)$$

Evidence of epigenetic factors on susceptibility to cervical cancer

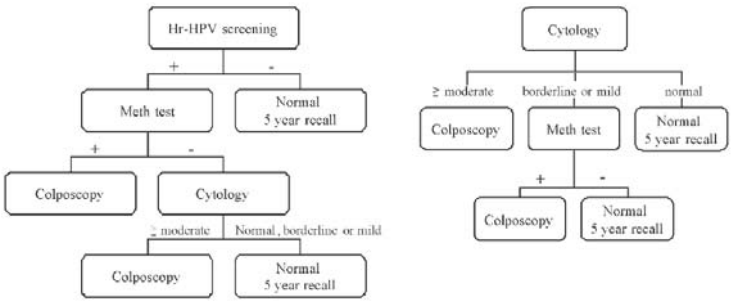
Methylation Markers for *CCNA1* and *C13ORF18* Are Strongly Associated with High-Grade Cervical Intraepithelial Neoplasia and Cervical Cancer in Cervical Scrapings

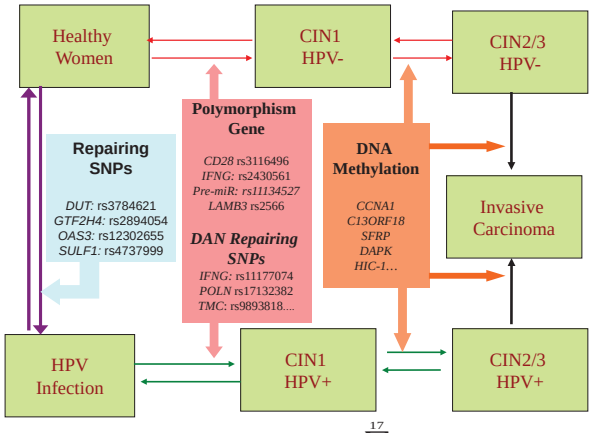
Nan Yang,[1] Jasper J.H. Eijsink,[1] Ágnes Lendvai,[1] Haukeline H. Volders,[1] Harry Klip,[1] Henk J. Buikema,[2] Bettien M. van Hemel,[2] Ed Schuuring,[2] Ate G.J. van der Zee,[1] and G. Bea A. Wisman[1]

Cancer Epidemiol Biomarkers Prev 2009;18(11). November 2009

Purpose: The aim of the present study was to perform an in-depth analysis of the methylation patterns of these 13 candidate genes in cervical neoplasia and to determine their diagnostic relevance.

Proposed scheme for the incorporation of a methylation test in cervical cancer screening if either hr-HPVor cytology will be used as primary screening.





$$Inc_{CIN1} = Baseline_{CIN1} \times \exp(\sum_{i=1}^{17} SNP_i)$$

$$\lambda_{CIN1 \to CIN2/3} = Baseline_{CIN1 \to CIN2/3} \times \exp(\sum_{i=1}^{9} Methylation_i)$$

$$\lambda_{CIN2/3 \to InvCa} = Baseline_{CIN2/3 \to InvCa} \times \exp(\sum_{i=1}^{7} Methylation_i)$$

61

Screening Strategy	Relative Risk									
	<20% HPV testing		20-40% HPV testing		40-60% HPV testing		60-80% HPV testing		>80% HPV testing	
	-	+	-	+	-	+	-	+	-	+
HPV vaccination + Pap Smear screening by inter-screening interval										
1 yr	0.32	0.28	0.15	0.19	0.21	0.20	0.24	0.23	0.29	0.30
3 yr	0.32	0.29	0.24	0.21	0.26	0.23	0.26	0.25	0.36	0.32
5 yr	0.39	0.34	0.27	0.20	0.25	0.24	0.28	0.26	0.36	0.32
Pap Smear screening by inter-screening interval										
1 yr	0.37	0.31	0.33	0.28	0.35	0.35	0.44	0.40	0.42	0.44
3 yr	0.45	0.35	0.35	0.33	0.40	0.33	0.46	0.46	0.51	0.45
5 yr	0.44	0.42	0.43	0.41	0.42	0.39	0.50	0.46	0.54	0.48
HPV Vaccination	0.94	0.96	0.55	0.61	0.57	0.58	0.64	0.58	0.73	0.70
Marginal effect of vaccination	3-8%		8-12%		10-17%		17-22%		13-18%	

Baseline (lacking of any Intervention (baseline): RR=1.00

Simulated Cohort with 300,000 Taiwan women age started from 0 years

62

Conclusion

- Cancer genetics play different roles by different disease natural history models.
- Complex Inheritance complicates risk assessment.
- Application to Personalized Medicine

63


國立臺灣大學
流病與預醫所
IEPM・CPH・NTU