

# 干擾與 去干擾-剷除潛在變項

106031125 張舒植 材料21  
107072117 廖姝鈞 經濟22

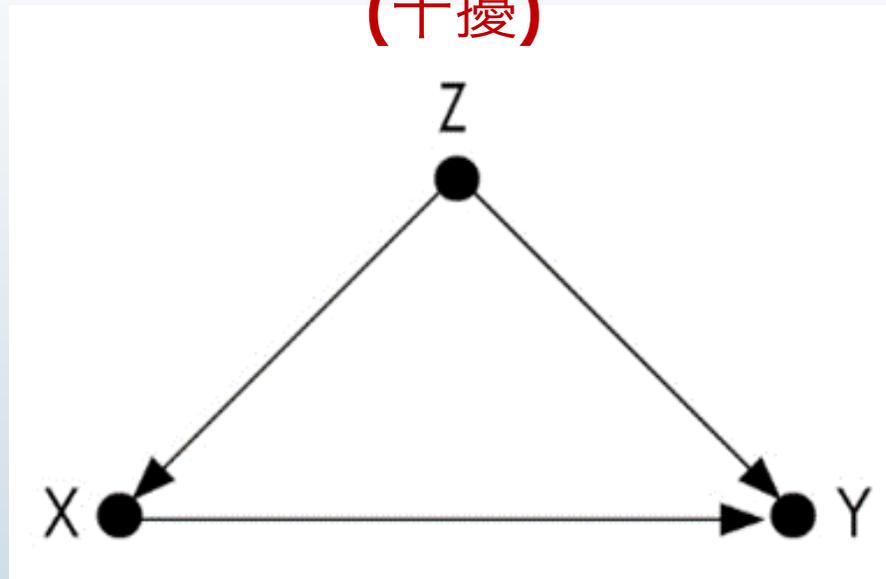
## 簡易定義干擾

### ■ 直觀:

由於某個因素的影響，導致原先結果改變，也可以說是把真正的因果效應與假相關混和在一起。

# 因果圖 比較素食者與葷食者身體狀況

實驗前身體狀況  
(干擾)



飲食  
(待觀測參數)

實驗後身體狀況  
(待觀測參數)

## 試比較素食者與葷食者身體狀況 (續)

變項條件化  
(conditioning)

- 對照實驗: 觀察兩組之間的差異
- Z 執行對照: 改變 Z, 觀察變化

目的

- 消除第三變項影響
- 使變項間資訊流通或中斷

# 干擾的各種定義

## 敘述性

- 同時與 **X**, **Y** 相關變項
- 流行病學經典定義:
  - I. 群體中與 **X** 相關
  - II. 未接受 **X** 處理的民眾與 **Y** 相關
  - III. **Z** 不應該在 **X** 和 **Y** 的因果路徑上

※ **X** (處理)、**Y** (結果)

## 程序性

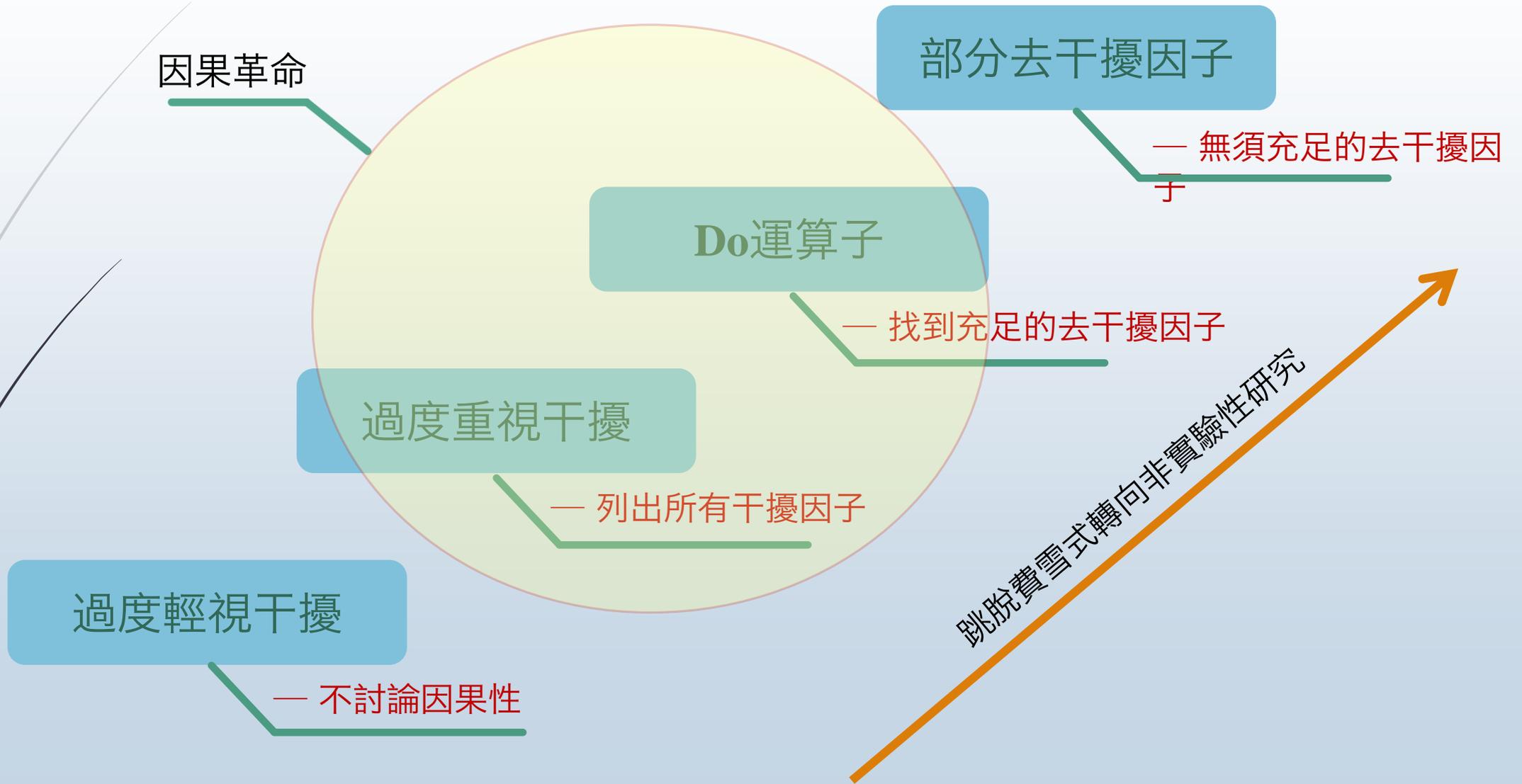
- 不可崩潰性 (noncollapsibility)
- 評估相對風險:
  - I. 原始值
  - II. 調整值

## 筆者定義

- 敘述性與程序性定義的問題:  
未去除 **Z** 對照對 **X** 的影響
- **Do** 運算子判斷:

$$P(Y|X) \neq P(Y|\text{do}(X))$$

# 去除干擾的四大進程



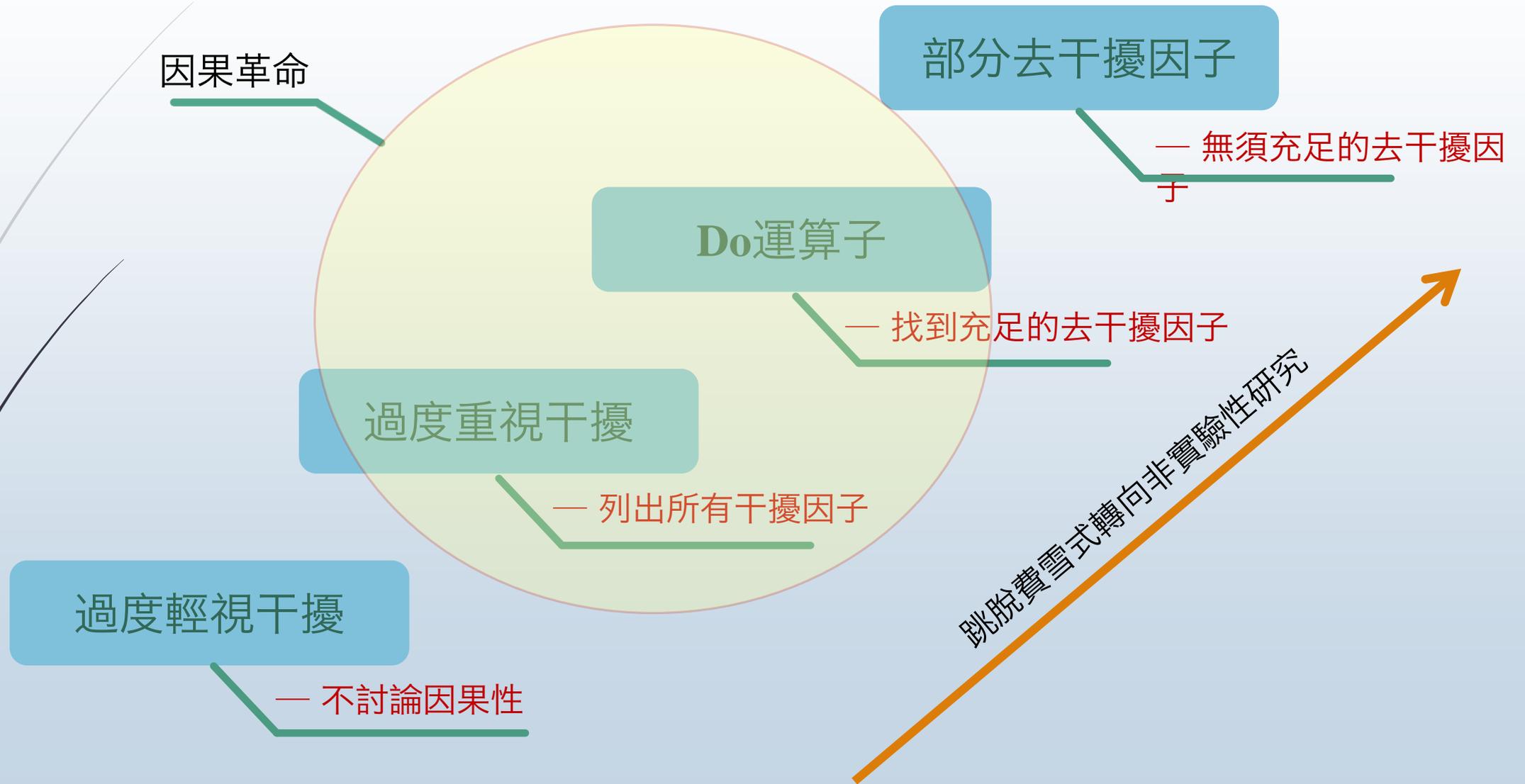
## 去干擾因子作用

- 去除某部分的干擾因子

- 充足的去干擾因子則可忽略部分干擾因子

- 解除過度重視干擾因子的問題

# 去除干擾的四大進程



# 過度重視干擾 — Robert Abbott 實驗

(經常步行與降低死亡率的關聯)

## 尋找干擾因子

- 列出所有可能的干擾因子(年齡、身體狀況、飲酒、飲食....)

## 調整干擾因子

- 針對每個可能的干擾因子進行調整
- 作者認為調整是為了要消除關聯中虛假的部分

## 結論

- 無法給出一個肯定的結論 (由於干擾因子太多無法一一討論)

## 過度重視干擾 — 費雪拉丁實驗方陣

- ➔ 系統化整理變數，比較不同樣本結果
- ➔ 缺點：
  - 無法預見所有干擾因子
  - 缺乏效率
- ➔ 轉向隨機實驗發展



實驗區依據肥料種類及土壤種類劃分

## 過度重視干擾 — 費雪隨機實驗 (RCT)

- ➡ 隨機組合實驗變項，避免討論所有干擾因子
- ➡ 優點：
  - a. 消除干擾偏差
  - b. 量化不確定性
  - c. 切斷所有與隨機化變項的連結
- ➡ 缺點：
  - a. 特定狀況無法介入 (例:受試者體重)
  - b. 違反實驗倫理 (例:要求受試者抽菸)

# 因果革命 — Do運算子及後門準則

目的:找到去干擾因子來消除干擾

(1) Do 運算子: 消除所有指向X的所有箭頭，使所有關於X的資訊都無法朝非因果方向流動

條件機率:  $P(Y|X) \neq P(Y|\text{do}(X))$

(2) 後門路徑: 由X到Y且已指向X的箭頭為起點的路徑

(3) 後門準則:

a. 執行對照並去除後門路徑(非因果路徑)

b. 阻斷資訊流通，遵守資訊流通的規則

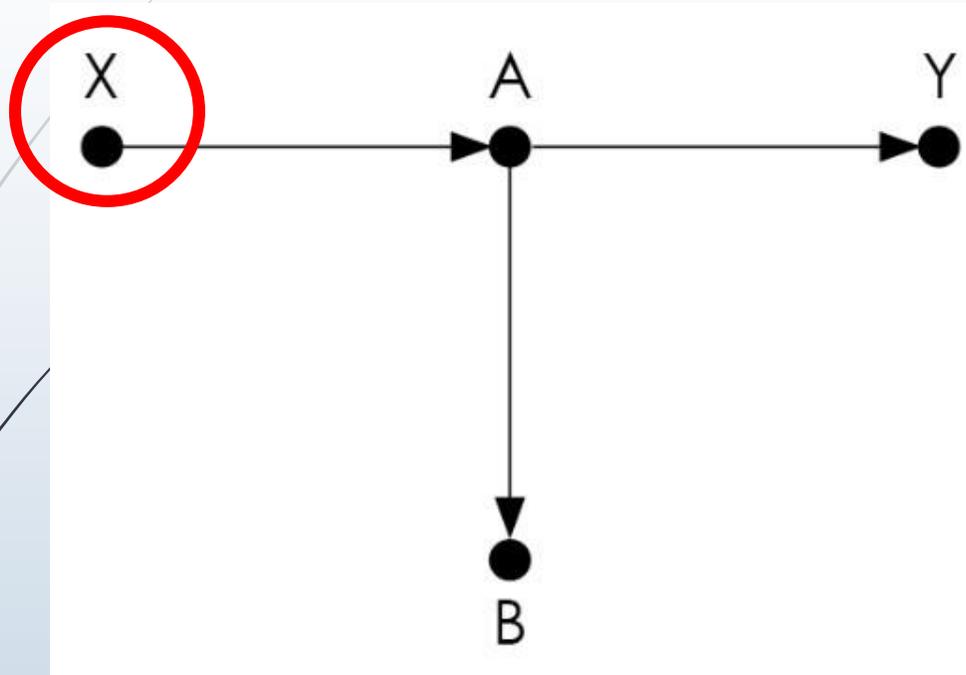
## 因果革命 — Do運算子及後門準則

### 資訊流通規則:

- $A \rightarrow B \rightarrow C$ : 以B執行對照，防止A的資訊傳到C
- $A \leftarrow B \rightarrow C$ : 以B執行對照，防止A的資訊傳到C
- $A \rightarrow B \leftarrow C$ : 互相獨立的A、C，以B執行對照反而使資訊流通
- $A \leftarrow B \rightarrow C \leftarrow D \rightarrow E$ :

用任何一個執行對照，都能使AC之間的資訊被切斷

## 遊戲驗證 — 遊戲 —



無後門路徑:

→ 沒有指向X的路徑

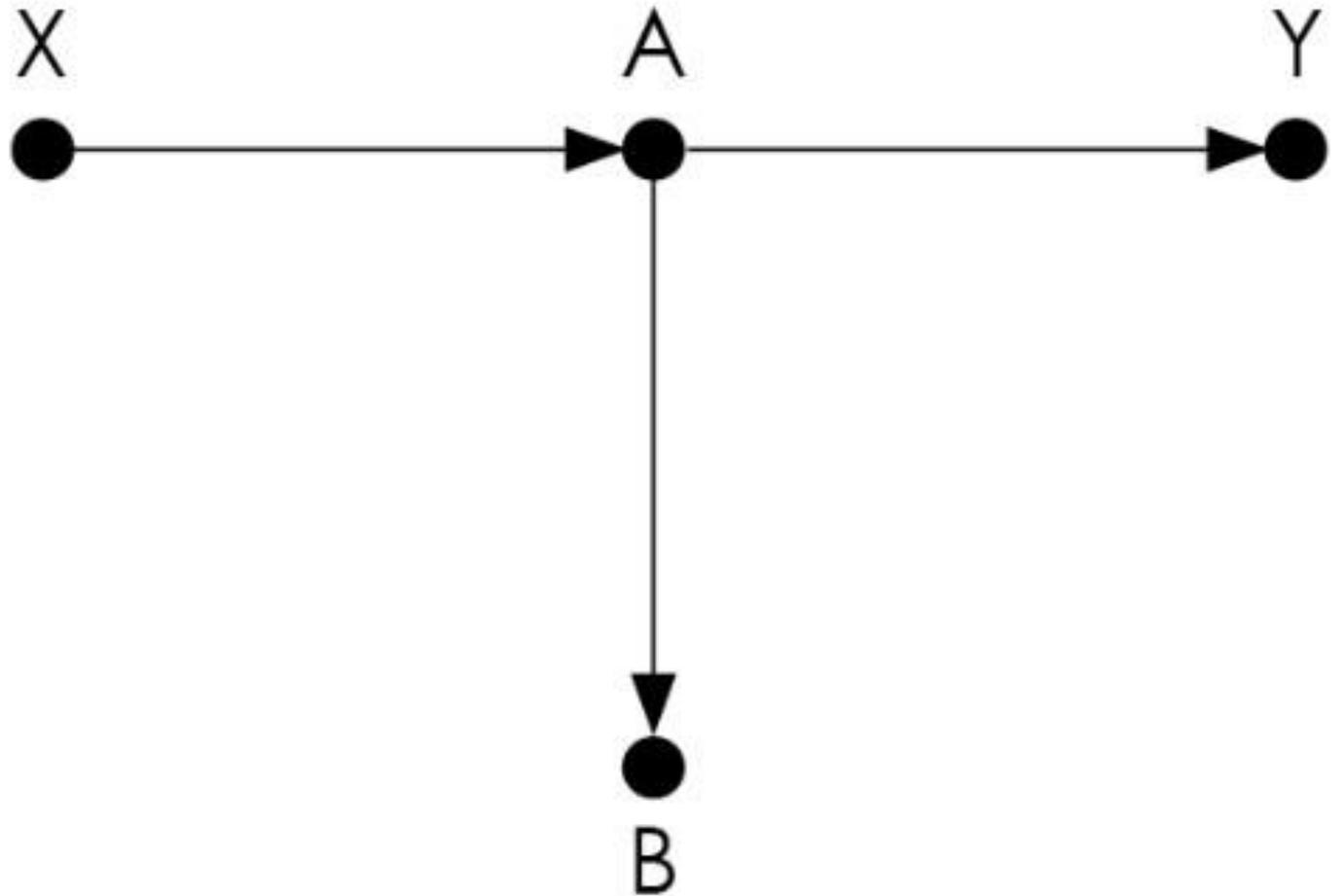
→ 無須對照

## 干擾的各種定義

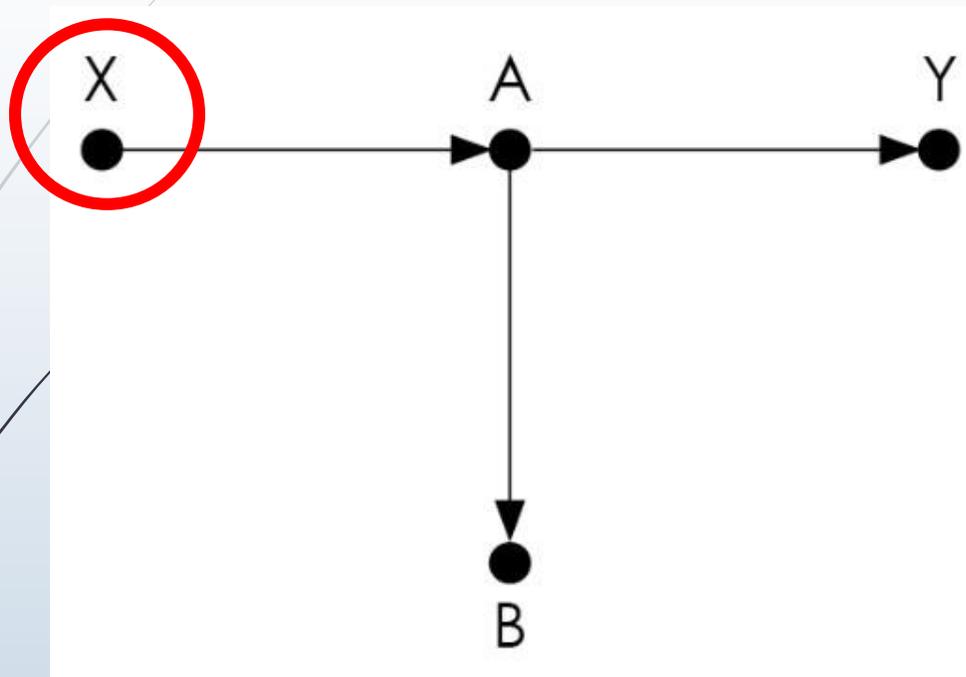
敘述性

- 同時與 X, Y 相關變項
- 流行病學經典定義:
  - I. 群體中與 X 相關
  - II. 未接受 X 處理的民眾與 Y 相關
  - III. Z 不應該在 X 和 Y 的因果路徑上

※ X (處理)、Y (結果)



## 遊戲驗證 — 遊戲一



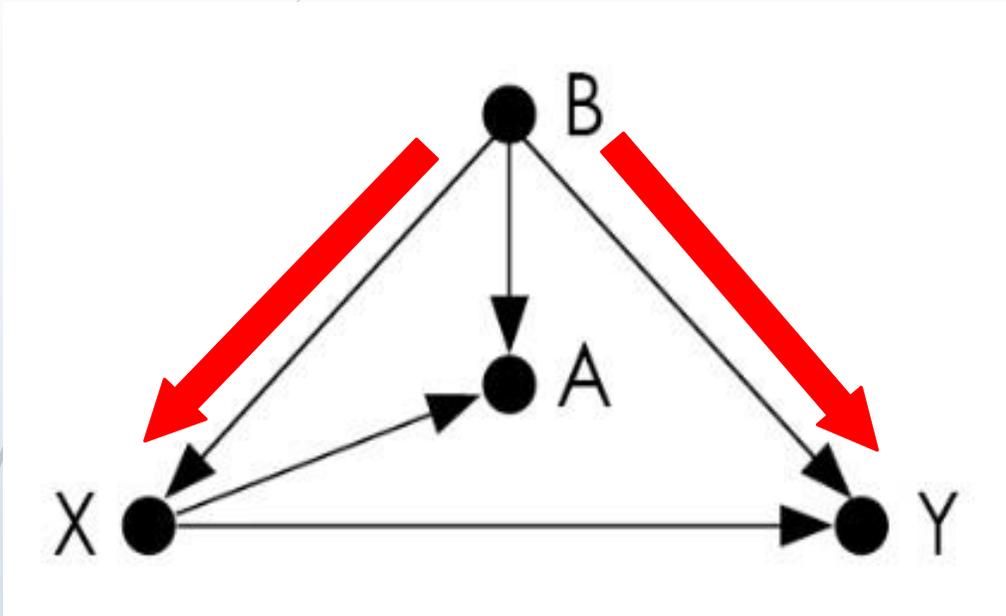
比較流行病學經典定義：

變項B符合干擾的定義，所以  
應該要對他執行對照。

但是，變項B是變項A的後代，  
執行對照B等同執行A部分對照。

→違反後門準則

## 遊戲驗證 — 遊戲二



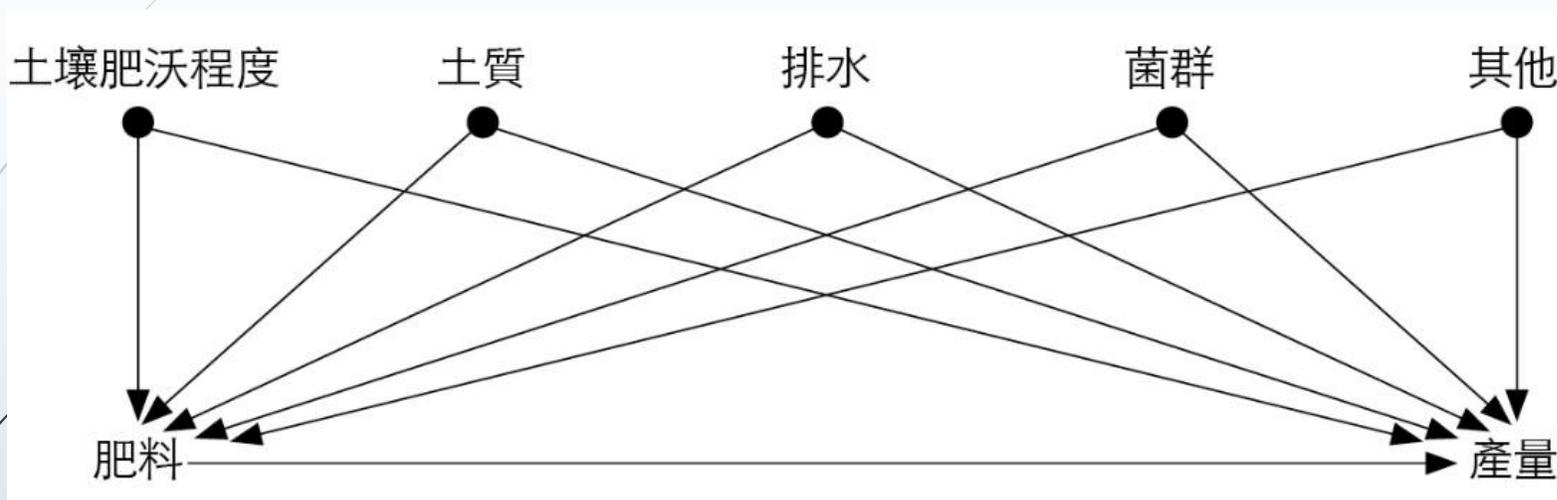
後門路徑： $X \leftarrow B \rightarrow Y$

依據資訊流通的規則，要阻斷此後門路徑需執行B對照。

執行A對照：

→ 僅部分消除干擾

## 比較RCT實驗與do運算子 — 因果圖模型—



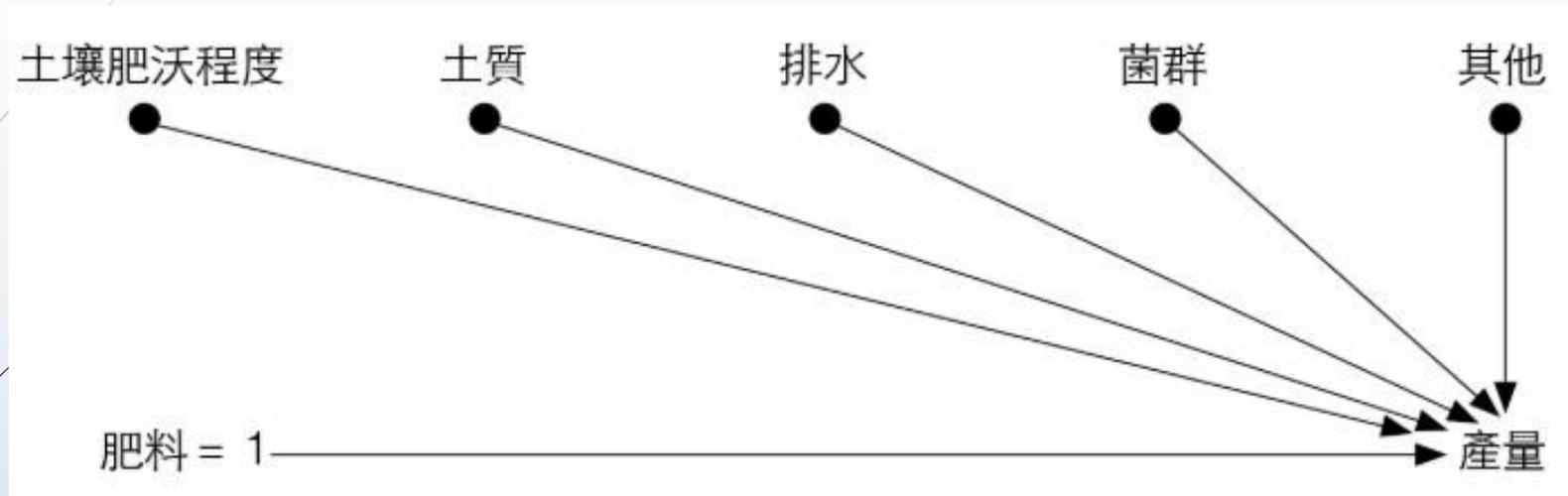
### 討論:

原想探討肥料與產量關係，卻包含眾多干擾因子，易出現偏差。

### 行動:

應想辦法去除干擾因子的影響 → RCT / do運算子

## 比較RCT實驗與do運算子 一因果圖模型二



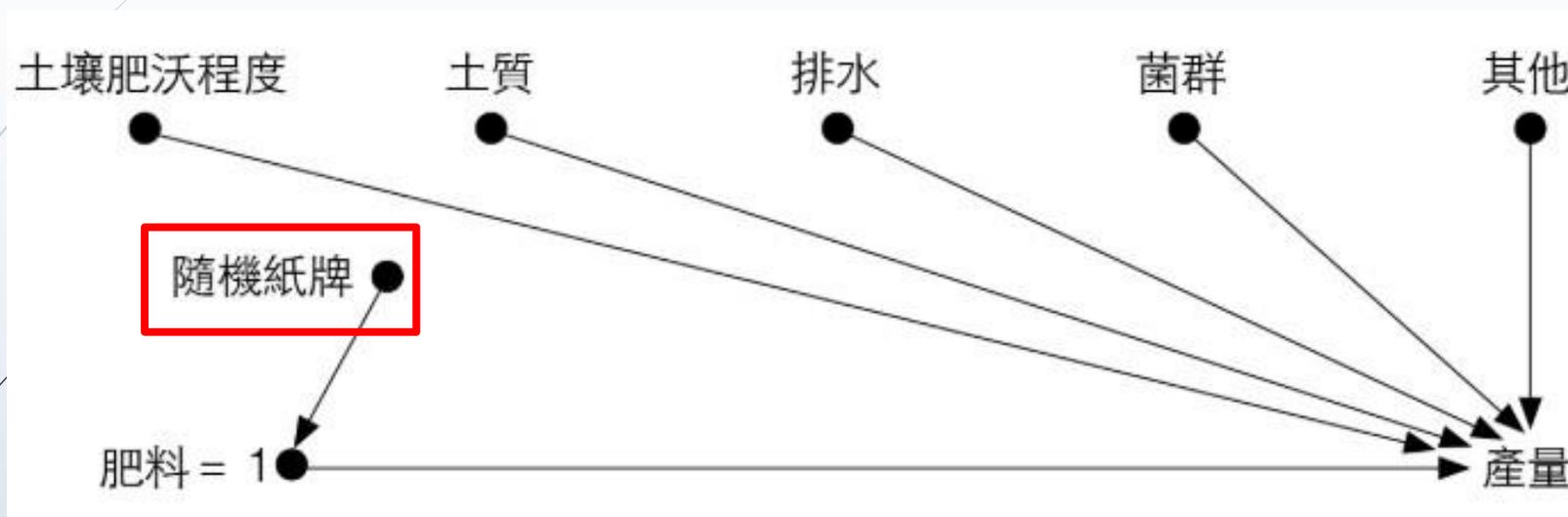
### 討論:

此模型是執行do運算子的結果，成功消除干擾。

### 行動:

比較RCT結果

## 比較RCT實驗與do運算子 一因果圖模型三



### 討論:

- RCT與do運算子模型呈現的世界相同
- 隨機化是模擬模型二的方法
- 消除干擾且不會生成新的干擾因子

## 比較RCT實驗與do運算子 一小結

### 相同:

- 可由隨機化完成模型，消除干擾。
- 非討論所有干擾因子，效率較高。

### 相異:

- do**運算子為非實驗研究推論，可避免**RCT**中可能遇到實驗倫理議題。

## 總結

### 定義干擾:

敘述法、程序法、作者定義

### 去除干擾的方法比較:

對照法

直觀卻會忽略第三變項與待測變項間的關係

RCT

可消除干擾偏差且量化誤差，無法討論違反實驗倫理之變項

do運算子

擁有隨機化模擬結果且屬非實驗研究推論