

# Bias e confondenti



SAPIENZA  
UNIVERSITÀ DI ROMA

[annarita.vestri@uniroma1.it](mailto:annarita.vestri@uniroma1.it)

# Validità delle associazioni

---

Per determinare se una associazione osservata in uno studio è valida è necessario cercare spiegazioni alternative ai risultati trovati. Cosa sarebbe successo se...? In particolare dobbiamo considerare tre spiegazioni alternative :

- probabilità
- bias
- confondenti

# probabilità

---

**Il risultato che osserviamo potrebbe essere dovuto al caso – alla variabilità di campionamento**

**Un modo comune di misurare l'effetto del caso è di effettuare un test statistico**

- Definiamo l'ipotesi nulla ( $H_0$ ): non c'è nessuna differenza, nessuna associazione**
- Saggiamo l'ipotesi alternativa ( $H_1$ ): c'è una differenza, c'è una associazione**

# probabilità

---

- Calcoliamo il p-value: la probabilità che il risultato del nostro studio o uno più estremo potrebbe verificarsi per effetto del caso, dato che è vera  $H_0$
- Sulla base del p-value, possiamo rifiutare  $H_0$  (quindi dire che c'è una associazione statisticamente significativa), o non rifiutare  $H_0$  (l'associazione non è statisticamente significativa)

- $p = 0.05$  è il livello di cut-off convenzionale (arbitrario) per la significatività statistica.
- se  $p \leq 0.05$ , concludiamo che il caso è molto improbabile per il nostro risultato e quindi affermiamo che esiste una associazione statisticamente significativa tra esposizione e malattia.
- se  $p > 0.05$ , concludiamo che il caso non può essere escluso come spiegazione del nostro risultato, non rifiutiamo  $H_0$ , e concludiamo che non esiste una associazione statisticamente significativa tra esposizione e malattia.

# BIAS

---

Nella determinazione dell'associazione tra esposizione e malattia ci possono essere fonti di errore. Potrebbe accadere che i partecipanti allo studio siano stati selezionati in modo non corretto (selection bias) oppure non è corretto il modo con il quale è ottenuta l'informazione (observation bias). La parola chiave rispetto al bias è "differente".

# SELECTION BIAS

---

**Può accadere che la selezione di un soggetto nello studio sia influenzata dalla conoscenza dello stato di esposizione o di malattia. Il selection bias non può essere corretto**

## ESEMPIO DI SELECTION BIAS

---

**Studio caso controllo con base ospedaliera tra uso di contraccettivi orali e tromboembolismo. Se le donne con sospetto di tromboembolismo sono più facilmente ricoverate se dichiarano di fare uso di CO, quindi il rischio osservato potrebbe essere distorto, ovvero sovrastimato rispetto all'uso di CO. Questo non può essere controllato nell'analisi ma solo prevenuto nel disegno.**



## ESEMPIO DI SELECTION BIAS

---

Ad esempio si vuole studiare la prevalenza di infezione cervicale da *Human papillomavirus* (HPV) nelle donne in età fertile in una regione.

Si sottopongono all'HPV DNA test tutte le donne sintomatiche che afferiscono agli ambulatori di ginecologia: la stima della prevalenza dell'infezione da HPV in questo caso sarà distorta rispetto a quella della popolazione generale che comprende tutte le donne in età fertile, sintomatiche (in cui la probabilità di isolare l'HPV è maggiore) e asintomatiche.

# OBSERVATION BIAS

---

Può accadere quando c'è mancanza di comparabilità tra l'accuratezza o la completezza delle informazioni tra i gruppi di studio.

- Recall bias:  
ricordo differenziale dell'evento tra i gruppi di studio.
- Misclassification:  
Errore nell'assegnazione delle informazioni in studio.  
Se è casuale viene compensata nell'associazione
- Response bias: questo occorre quando si intervistano persone che hanno una predisposizione a riferire eventi personali in maniera sistematicamente diversa da quella della popolazione generale

Si possono introdurre correttivi nel disegno dello studio ma non possono correggere i bias nell'analisi

# OBSERVATION BIAS

---

## Recall bias:

donne che hanno avuto un nato pretermine (casi), se intervistate, riferiranno con maggiore precisione e dettaglio gli eventi della gravidanza (come per esempio gli episodi febbrili), rispetto a donne con neonati a termine (controllo);

le donne con neonato a termine infatti non ripercorrono mentalmente la gravidanza altrettanto scrupolosamente delle donne con neonato pretermine, che tentano di risalire nella maniera più precisa possibile all'evento all'origine dell'esito negativo.

# OBSERVATION BIAS

---

## Misclassification:

Es Si realizza tutte le volte in cui il paziente è posto nel gruppo sbagliato a causa della errata rilevazione del parametro considerato (es. misurazione errata di peso ed altezza e collocazione erronea nel gruppo con BMI basso)

# OBSERVATION BIAS

---

## Response bias

Un ampio studio osservazionale prospettico condotto in Scozia ha studiato l'associazione fra livello di stress psicologico percepito (esposizione) in un gruppo di persone non caratterizzate da svantaggio socioeconomico e sintomatologia anginosa percepita (esito soggettivo), episodi di ischemia rilevati tramite ECG (esito oggettivo), ricoveri per cause generali (esito dipendente dai sintomi lamentati), ricoveri per coronaropatia e mortalità associata (esiti oggettivi).

Sono stati arruolati oltre 5mila uomini di età media pari a 48 anni; oltre 2mila uomini sono stati disponibili per una nuova valutazione a distanza di 5 anni dall'arruolamento. Le morti e i ricoveri sono state rilevate per un arco di tempo di 21 anni. Un elevato livello di stress psicologico (autodefinito) si associa a maggiore incidenza e prevalenza di angina (**odds ratio**, OR aggiustata: 2.66; **intervallo di confidenza al 95%**, IC 95%: 1.61, 4.41) e a maggior frequenza di ricovero per cause generali (OR: 1.13; IC 95%: 1.01, 1.27).

# OBSERVATION BIAS

---

## Response bias

Non c'è invece alcuna associazione significativa con incidenza e prevalenza di ischemia, con ricovero per coronaropatia, con mortalità per coronaropatia o per malattia cardiovascolare.

L'associazione fra la percezione del livello di stress psicologico e il rischio di angina non è confermata da misure oggettive di danno cardiaco; gli autori concludono che fra gli uomini arruolati, quelli caratterizzati da assenza di svantaggio sociale, hanno una elevata percezione di stress e riferiscono più frequentemente degli altri uomini sintomi di angina.

Ma le due variabili non sono associate causalmente; sono solo associate a una maggior inclinazione alla lamentelela.

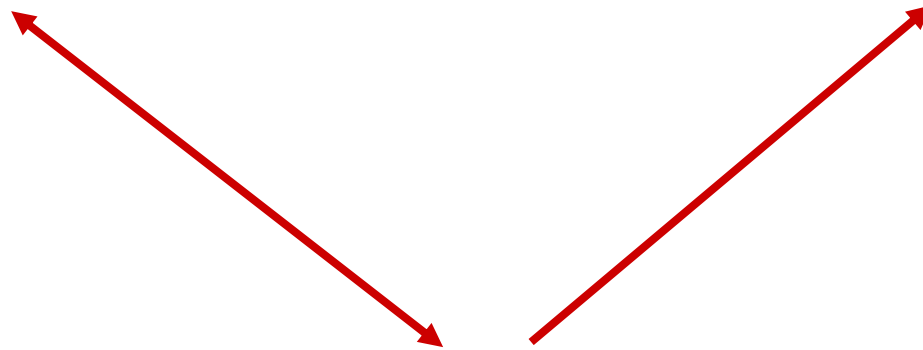
# CONFONDENTE

---

**Effetto tra l'associazione in studio e una terza variabile. Il terzo fattore (il confondente) deve essere associato sia all'esposizione e indipendentemente dall'esposizione deve essere causa, o correlato, con la malattia.**

**Il confondente può essere responsabile in parte o totalmente all'associazione trovata nei dati.**

**ESPOSIZIONE** → **MALATTIA**



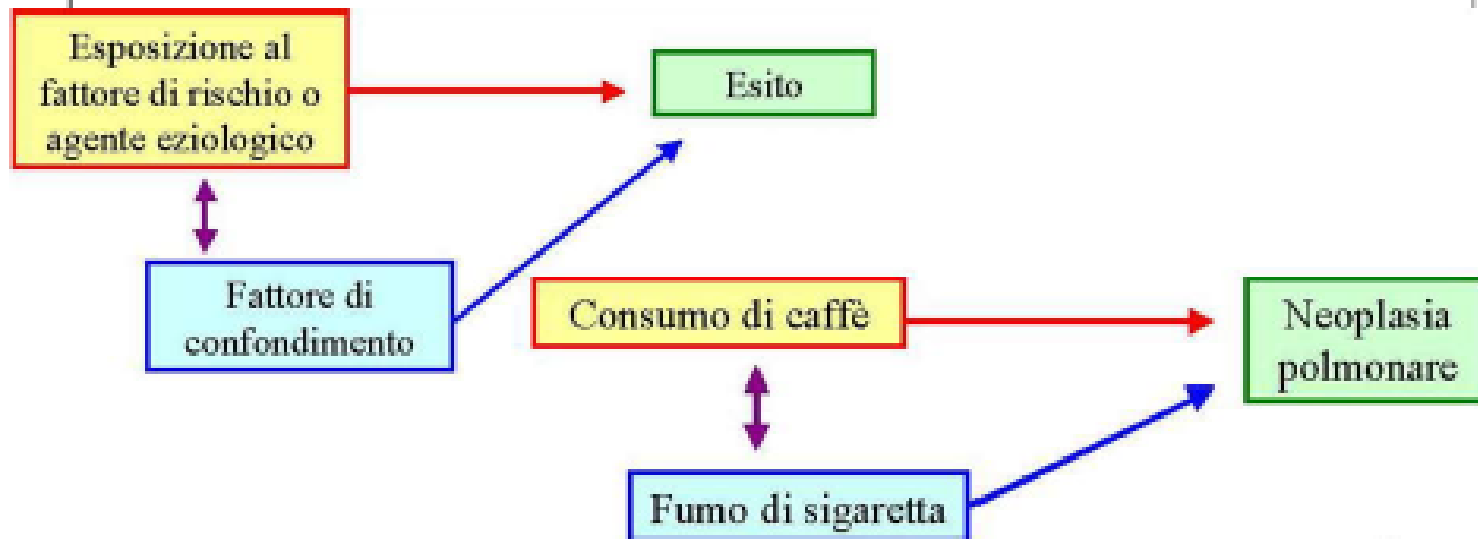
**CONFONDENTE**

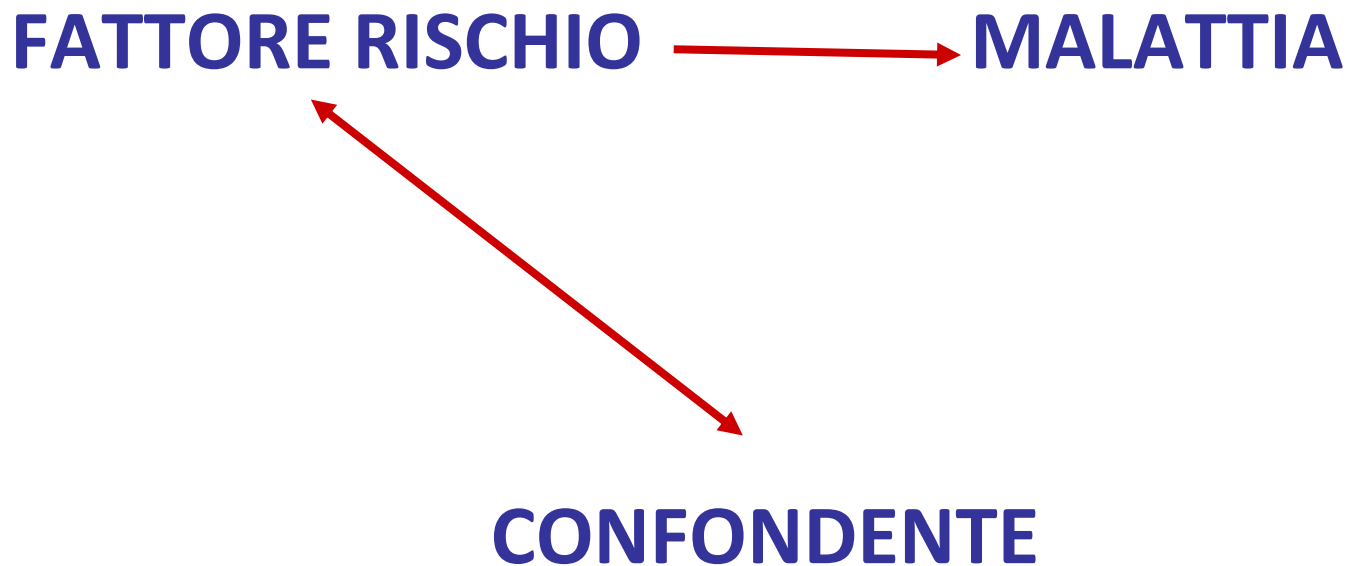


# CONFONDENTE

Esiste una relazione tra l'assunzione di caffè e l'insorgenza del tumore del polmone?

- I fattori di confondimento sono le variabili associate sia all'esposizione, sia all'esito, che potrebbero spiegare qualsiasi relazione osservata tra questi due fattori.





**NOTA:** se associato con l'esposizione ma non con la malattia, non è un confondente (es., fumo e tumore al polmone, alcool non è un potenziale confondente perchè l'alcool non è un fattore di rischio indipendente per il tumore al polmone).

**FATTORE RISCHIO** → **MALATTIA**

**CONFONDENTE**




```
graph TD; A[FATTORE RISCHIO] --> B[MALATTIA]; C[CONFONDENTE] --> B;
```

**NOTA: se associato con la malattia ma non con l'esposizione non è un confondente (es. Uso moderato di alcool e CHD, se l'esercizio fisico non è associato con l'alcool non è un confondente)**

**FATTORE** → **CONFONDENTE** → **MALATTIA**

**NOTA: questo è un intermediario, non un confondente (es., obesità e malattia coronarica, il diabete non è un confondente).**

## Possibili fattori di confondimento negli studi etiologici per importanti malattie

Malattie respiratorie		Fumo
Diabete		Familiarità
Malattie cardiache		Fumo, familiarità
Tumore alla cervice		Uso di contraccettivi
Tumore all'esofago		Fumo, alcol
Tumore al fegato		Alcol, HBV, HCV
Tumore al colon		Poliposi familiare
Tutte le malattie		Età, genere, etnia

# QUALI SONO I POTENZIALI CONFONDENTI?

---

- **Fattori conosciuti che sono legati sia all'esposizione che alla malattia di interesse, ma non sono il meccanismo con il quale si suppone che agisca l'esposizione**
- **Se si sospetta che un fattore di rischio possa essere un potenziale confondente si raccolgono informazioni nel disegno dello studio**

# METODI PER IL CONTROLLO DEI POTENZIALI CONFONDENTI

---

## 1. NEL DISEGNO DELLO STUDIO:

- Restrizione: (si restringe lo studio ai soggetti di uno strato del fattore confondente)
- Matching: (appaiare i gruppi di studio in modo da rendere sovrapponibili i fattori confondenti)
- Randomizzazione: (se è un trial)

## 2. NELL'ANALISI DELLO STUDIO:

- Analisi appaiata: (se il disegno è appaiato)
- Stratificazione: (analizzare l'associazione separatamente per ogni livello del fattore confondente e standardizzazione)
- Analisi multivariata: (modelli matematici per controllare i confondenti simultaneamente)



# COME RICONOSCERE SE UN POTENZIALE CONFONDENTE E' UN CONFONDENTE?

---

- **Confronto tra RR grezzo e RR aggiustato: la differenza tra i due è dovuta al confondente. Esempio:**
  - **Dieta a basso contenuto di grassi e CHD, RR = 0.60 confrontato con dieta usuale.**
  - **Aggiustato per BMI, RR = 0.80.**

# **MODIFICATORE D'EFFETTO (INTERAZIONE)**

---

**Quando la grandezza della relazione tra esposizione e malattia differisce in dimensione (è modificata) dal livello di una terza variabile (chiamata modificatore d'effetto)**

- **ESEMPIO:      contraccettivi orali e IMA in donne in età fertile.**

**RR grezzo= 2.0**

**RR (CO and IMA, tra non fumatori) = 1.9**

**RR (CO and IMA, tra fumatori) = 41.0**

**RR (CO and IMA, tra bevitori) = 2.0**

**RR (CO and IMA, tra non bevitori) = 2.0**

- **Il fumo modifica l'effetto dei CO su IMA, l'associazione è differente tra fumatori e non fumatori L'alcool non è un modificatore d'effetto su questa associazione.**