

# LA STATISTICA PER IL MEDICO

## L'interpretazione dei dati epidemiologici

Professor Carlo La Vecchia

## ■ LA CAUSALITÀ IN EPIDEMIOLOGIA

- L'obiettivo di uno studio epidemiologico ed in particolare degli studi sui fattori di rischio, è quello di stabilire un'inferenza di **causalità** a livello di popolazione, tra questi fattori e la patologia, oltre che di quantificarne i rischi.
- A causa della mancanza di randomizzazione, negli studi epidemiologici è importante che il ruolo della **distorsione**, del **confondimento** e del **caso** sia accuratamente valutato prima di poter formulare ogni inferenza di causalità tra esposizione e patologia.

## ■ LA CAUSALITÀ IN EPIDEMIOLOGIA: UNA NOTA IMPORTANTE

Ogni volta che si parla di **causalità** nel contesto di studi epidemiologici, ci si riferisce ad una specifica accezione del termine, non esportabile al di fuori dell'epidemiologia.

In tal senso, la causalità definisce **l'aumento del rischio** di una data patologia a livello di una popolazione.

Deve essere infatti ben chiaro che la causalità epidemiologica nulla può dire sulla causa di malattia del singolo caso, che potrà invece essere o meno stabilita solo applicando i metodi della diagnostica clinica.

## I TERMINI PIÙ FREQUENTI IMPIEGATI NELLA DISCUSSIONE EPIDEMIOLOGICA

### Il *bias*

- Il *bias*, o errore sistematico, deriva generalmente da imprecisioni nel disegno di uno studio non sperimentale e nella raccolta dei dati.
- Il *bias* di selezione comporta differenze sistematiche nell'esposizione oggetto dello studio, tra coloro che vengono selezionati per essere inclusi nello studio e coloro che non vengono selezionati.
- Il *bias* di informazione comporta differenze sistematiche nella misurazione dell'esposizione in oggetto tra i gruppi di confronto e comprende il *bias* di ricordo e di intervista. Laddove qualità e dettaglio delle informazioni raccolte risultino eterogenee, il confronto tra casi e controlli non sarà adeguato.

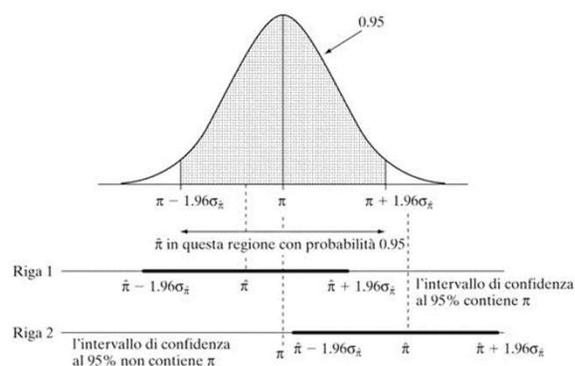
# GLI INTERVALLI DI CONFIDENZA

Negli studi epidemiologici, il ruolo del caso viene valutato per mezzo di **test di significatività**, nonché del calcolo degli **intervalli di confidenza**.

Se una particolare stima di rischio è statisticamente significativa a un livello definito (usualmente al 95%, ammettendo un margine di errore casuale del 5%), il caso può essere escluso come probabile spiegazione dei risultati, con maggiore o minor margine di errore.

Il **livello di significatività** comunemente adottato è di  $p < 0.05$ , corrispondente a un intervallo di confidenza (IC) del 95%.

Quando l'IC include l'**unità**, il risultato non è significativo. Quando entrambi i limiti di confidenza (quello inferiore e quello superiore) sono sotto o sopra l'unità, il risultato è significativo nell'indicare rispettivamente una riduzione o un eccesso di rischio relativo (RR).



## I TRADIZIONALI E DISCUSSI CRITERI DI HILL

### **Forza del nesso.**

In generale, più elevata è la stima del rischio, minore è la probabilità che il risultato possa essere spiegato con la distorsione, il confondimento o il caso.

In linea di massima, le associazioni forti (in base al valore del RR) sono ritenute causali, poiché è poco verosimile che tali associazioni derivino soltanto da distorsioni o da errori sistematici.

D'altra parte, è molto difficile, per non dire impossibile, definire una relazione di causalità attraverso studi epidemiologici osservazionali che abbiano rilevato rischi relativi dell'ordine di 1.2-1.5 o più in generale  $< 2$ .

## ■ COERENZA

- La **plausibilità** di un nesso causale è notevolmente maggiore se risultati simili vengono descritti in studi epidemiologici condotti su popolazioni diverse utilizzando vari schemi di studio.
- L'esistenza di una **coerenza interna** è uno degli argomenti di base per stabilire la **causalità di un'associazione**.
- Quando non vi sia associazione significativa nella globalità della popolazione, ma soltanto in alcuni sottogruppi (mentre, evidentemente, altri sottogruppi mostrano associazioni in senso inverso), è necessario essere scettici nell'interpretare uno studio.
- In linea generale, soltanto in presenza di un'**associazione riferibile alla globalità** della popolazione esaminata, sarà consentito passare alle analisi di sottogruppi e sempre sulla base di ipotesi biologiche a priori.
- Le associazioni casualmente emerse a posteriori in uno studio devono invece restare nel campo delle formulazioni di ipotesi e non in quello delle inferenze di causalità

## ■ EFFETTO DOSE-RISPOSTA

- Se il rischio di malattia aumenta all'aumentare dell'esposizione, un'interpretazione causale del nesso diviene più plausibile.
- La presenza di **gradienti dose-rischio** e **durata-rischio** rappresenta quindi una forte evidenza a favore della causalità di un'associazione.
- Occorre ricordare, però, che è sempre necessario tenere **separati** i due fattori della **dose** e della **durata**, perché il loro effetto biologico è spesso quantitativamente molto diverso.

## ■ SEQUENZA TEMPORALE e PLAUSIBILITÀ BIOLOGICA

### SEQUENZA TEMPORALE

- L'esposizione deve precedere lo sviluppo della malattia.
- Durata dell'esposizione, tempo dalla prima esposizione (**latenza**) e dall'ultima esposizione (**recenza**) devono essere considerate separatamente in uno schema patogenetico globale.

### PLAUSIBILITÀ BIOLOGICA

- È necessario chiedersi, in termini rigorosi, se il legame tra esposizione e patologia abbia senso dal **punto di vista biologico**, alla luce di ciò che le conoscenze scientifiche ci dicono sulla malattia.
- L'osservazione può suonare banale, ma acquista immediatamente valore considerando che sono stati pubblicati numerosi studi nei quali le associazioni tra determinati fattori di rischio e l'insorgenza di alcune malattie, pur raggiungendo la significatività statistica, lasciano assai perplessi circa il loro significato biologico.
- È poi indispensabile verificare se il legame è compatibile con le **prove sperimentali**. Purtroppo, le informazioni necessarie per giudicare la **plausibilità biologica** sono limitate per molte patologie.

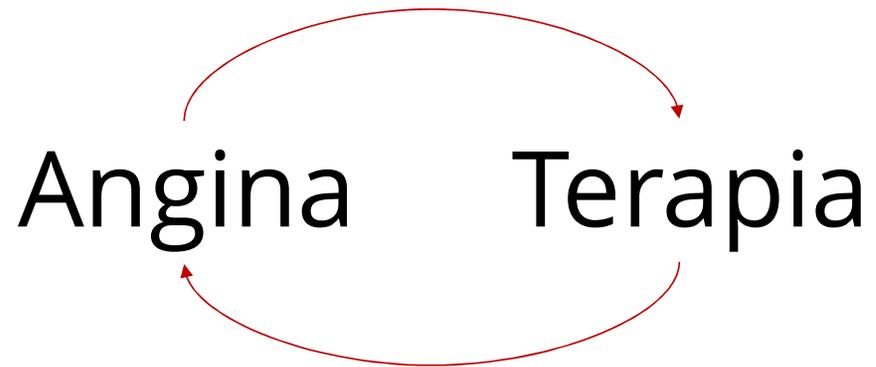
## ■ Fonti utili

- Trichopoulos D, Lipworth L, Petridou E, Adami H-O (1997). Epidemiology of cancer. In: De Vita VT Jr, Hellman S, Rosenberg SA, editors, Cancer: Principles and Practice of Oncology, 5<sup>th</sup> edition. Philadelphia: J.B. Lippincott Co.
- Hill AB (1965). The environment and disease: association or causation? Proc Soc Med 58: 295-300.
- US Department of Health, Education, and Welfare (1964). Smoking and Health: Report of the Advisory Committee to the Surgeon General. Washington, DC: US Government Printing Office, Public Health & Service Publi.
- MacMahon B., Trichopoulos D., 1996. Epidemiology. Principle and Methods, second ed. Little Brown and Company, Boston.

■ **Esempio**



Agranulocitosi

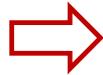


## ■ ESEMPIO

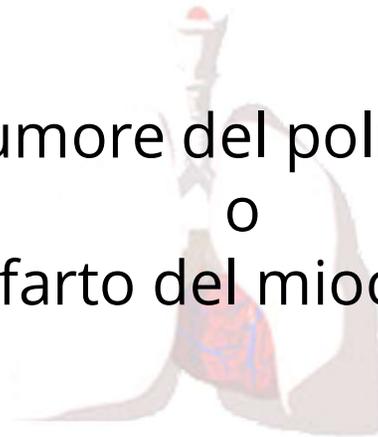
# Malattie croniche



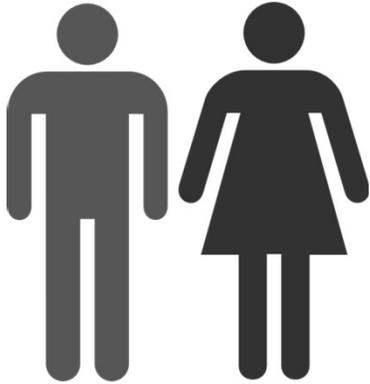
Fumo



Tumore del polmone  
o  
Infarto del miocardio



## ■ Studio di coorte



Carcinoma  
prostatico

Carcinoma  
mammario

Livelli ormonali al basale