

ASSISTENZA SANITARIA
&
TECNICHE DELLA PREVENZIONE NELL'AMBIENTE E NEI
LUOGHI DI LAVORO

STATISTICA MEDICA

gbarbati@units.it

A.A. 2024-25



UNITÀ DI BIOSTATISTICA

Dipartimento Universitario Clinico di
Scienze Mediche Chirurgiche e della Salute

CALENDARIO/ARGOMENTI DELLE LEZIONI

<u>orario 11 – 15:30</u>	Pausa 12:30 - 13:15
Novembre	
5	Lezione Introduttiva
26	Statistica Descrittiva 1
Dicembre	
3	Statistica Descrittiva 2
10	Statistica Inferenziale 1
17	Statistica Inferenziale 2



Scaricare il materiale didattico da Moodle: <https://moodle2.units.it/>

228ME-3 - STATISTICA MEDICA 2024

Di volta in volta
troverete qui tutto il
materiale svolto a
lezione

Vi dovete iscrivere al corso in Moodle per poter fare i quiz !!

Dipartimento Universitario Clinico di Scienze mediche, chirurgi... / Laurea triennale (DM270) / ME16 - TECNICHE DELLA PREVENZIONE NELL'AMBIENTE E NEL... / A.A. 2024 - 2025

228ME-3 - STATISTICA MEDICA 2024

[Corso](#) [Impostazioni](#) [Partecipanti](#) [Valutazioni](#) [Report](#) [Altro](#) ▾



Materiale didattico:

Lezioni (.pdf), esercitazioni, quiz di autovalutazione

Modalità di Esame:

Esame scritto:

- Il quiz su Moodle (da svolgere in aula) sarà costituito da **10** domande. Avrà una durata massima di **45** minuti.
- Il quiz sarà superato se si otterranno **almeno 5** risposte corrette.

< 5 : non superato

5: 20/30

6: 22/30

7: 24/30

8: 26/30

9: 28/30

10: 30/30

Nel quiz ci sono domande di teoria e qualche domanda sull'interpretazione di output dal **software R/JAMOVI**

APPELLI PARZIALI DEL MODULO DI STATISTICA

Descrizione Appello	Data ora aula
<u>Secondo Appello Invernale Statistica</u>	 19/02/2025 09:00
<u>Primo Appello Invernale Statistica</u>	 24/01/2025 09:00

Vi dovete iscrivere all'appello parziale in ESSE 3, nella data in cui intendete sostenere l'esame.

I voti al quiz appariranno in Moodle subito dopo la conclusione della prova.

Chi intende RIFIUTARE il voto deve comunicarmelo via mail.

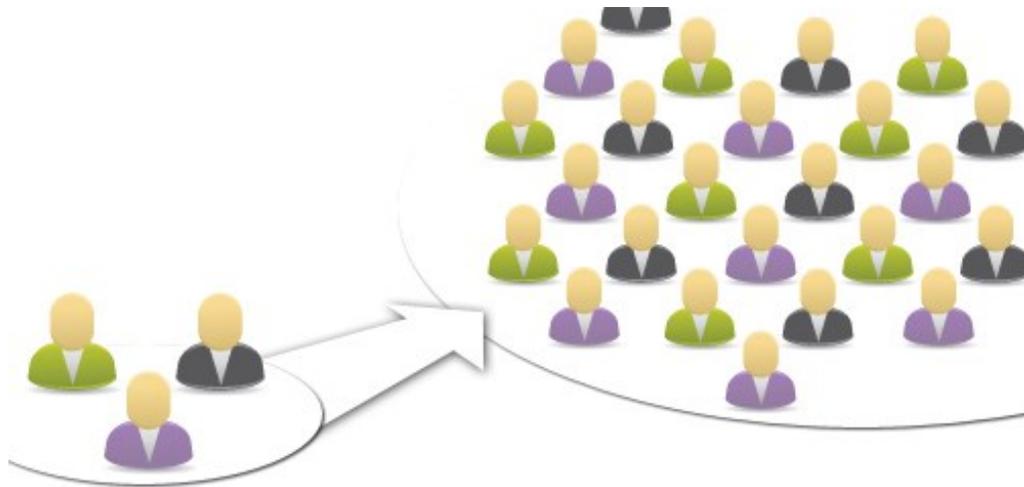
Ricevimento studenti: nella **pausa pranzo** o tramite videochiamata in MS teams

**Previo avviso
tramite e-mail o
comunicazione
nella classe
Teams**



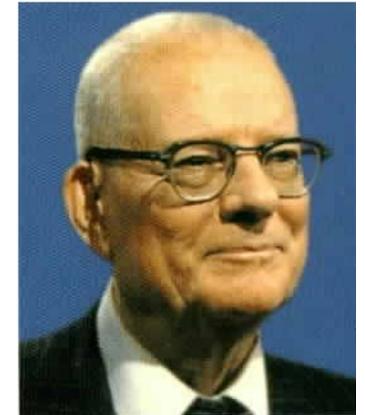
ARGOMENTI:

- Introduzione alla Statistica
- Ruolo della Statistica nella Ricerca
- Cenni ai Disegni dello Studio



In God we trust, all
others bring data.

William E. Deming
(1900 – 1993)



Che cos'è la statistica?

La statistica ha come scopo la conoscenza **quantitativa** (e l'interpretazione...) dei **fenomeni collettivi**.

Es: **Quante** persone partecipano a questa classe?

L'operazione che si deve fare è forse la più importante, ma certo la più semplice, operazione della statistica: il **conteggio**.
In base ad esso si perviene alla conoscenza del numero delle persone che costituiscono il gruppo.

I fenomeni che la nostra mente non può conoscere con una sola osservazione, ma che invece apprende tramite la *sintesi* delle osservazioni di fenomeni più semplici, si possono definire *fenomeni di massa* o *fenomeni collettivi*.

Essi sono composti da una collettività di osservazioni di fenomeni più semplici che sono detti quindi, *fenomeni individuali* o *singoli*.

Tramite la statistica è possibile sostituire ad un'impressione *qualitativa* di un fenomeno la sua misura **quantitativa**.

Che cosa si intende in statistica per 'fenomeni collettivi'?

- Collettività di casi singoli:

Caratteristiche di una popolazione osservando gli individui che compongono.
(Es: nascite, matrimoni, morti...)

- Ripetizioni dell'osservazione di un unico fenomeno singolo:

La lunghezza di un segmento, la massa di un corpo, la temperatura di un materiale ecc...: un'unica misurazione non può fornire l'esatta misura del fenomeno, poiché eseguendo misure della stessa grandezza si ottengono generalmente misure diverse (anche se di poco). Se le differenze sono effetti di cause che sfuggono ad ogni controllo, si può avere una misura verosimilmente esatta **ripetendo** l'osservazione un certo numero di volte e sintetizzando in un unico numero i risultati ottenuti. Alla misura si perviene quindi mediante una **collettività di misurazioni**.

La statistica nelle applicazioni biomediche

Evidence-based Medicine: la pratica professionale basata sulle evidenze scientifiche.



- 1) Raccolta delle evidenze scientifiche
- 2) Interpretazione critica



La statistica **sintetizza** ed **interpreta** i dati, e contribuisce alla progettazione e al disegno della ricerca stessa.



I **fruitori** di queste ricerche (chimici, medici, biologi, tecnici sanitari..) devono sapersi *orientare* sull'utilizzo e l'interpretazione dei metodi statistici (di base).

Chi usa la statistica?

- Le scienze che studiano un aspetto di una **collettività di individui** o enti:
per es: demografia, economia, sociologia, ***epidemiologia*** (cioè quel settore della medicina che studia la frequenza, la distribuzione e il ruolo di fattori causali di un determinato evento nelle popolazioni), la **ricerca biomedica** sull'efficacia dei **farmaci**;
- Le scienze cosiddette 'sperimentali' quelle cioè che **analizzano serie di dati** misurati con molta precisione: la fisica, la chimica, l'ingegneria, la biologia, ***la ricerca biomedica di laboratorio***...ecc...

Qual è il ruolo della statistica nella ricerca?

Lo studio dei fenomeni collettivi avviene attraverso diverse fasi:

FASE 1:

Schematizzazione: definizione del fenomeno, individuazione della **collettività** in cui esso si realizza, scelta delle **caratteristiche** della collettività che interessano.

Si intende studiare la *propensione al consumo di alcolici* (= fenomeno) nella *popolazione europea* di età compresa tra i 14 e i 30 anni (= collettività). Si decide di misurare *la spesa media per alcolici in una settimana* (= caratteristica).



(A) TUTTA LA POPOLAZIONE VIENE STUDIATA (CENSIMENTO):

Se la collettività di individui è *interamente da osservare* alla fase di schematizzazione segue la:

FASE 2A: Osservazione: raccolta, ordinamento e classificazione delle informazioni.

Alla collettività di individui si sostituisce un ***insieme di dati*** (dati sulla spesa media settimanale per alcolici per individuo) su cui si può operare con procedimenti statistici per la fase successiva:

FASE 3A: Descrizione dei dati raccolti.

FASE 1: Schematizzazione, definizione del fenomeno, individuazione della collettività (popolazione target), scelta delle caratteristiche che interessa rilevare.



(B) VIENE STUDIATA UNA PARTE DELLA POPOLAZIONE (CAMPIONE):

Si osserva solo una parte della popolazione. A causa di questa osservazione parziale, ci sono più fasi:

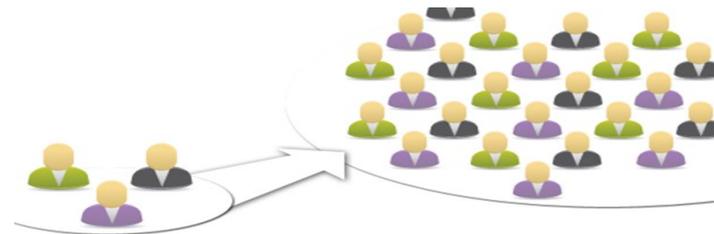
FASE 2B: **Formulazione di ipotesi** sulla popolazione target;

FASE 3B: **Estrazione** del campione, che è una parte popolazione target;

FASE 4B: **Descrizione dei dati campionari** con metodi della statistica descrittiva;

FASE 5B: **Induzione** o di **Inferenza** : risalire dalla descrizione del campione a quella della popolazione target -> **verifica delle ipotesi** della fase 2B.

Step del progetto di ricerca:



quesito scientifico -> ipotesi -> **disegno dello studio** -> raccolta dati [campione] -> analisi statistica

quesito scientifico -> ipotesi -> **disegno dello studio** -> analisi dei dati

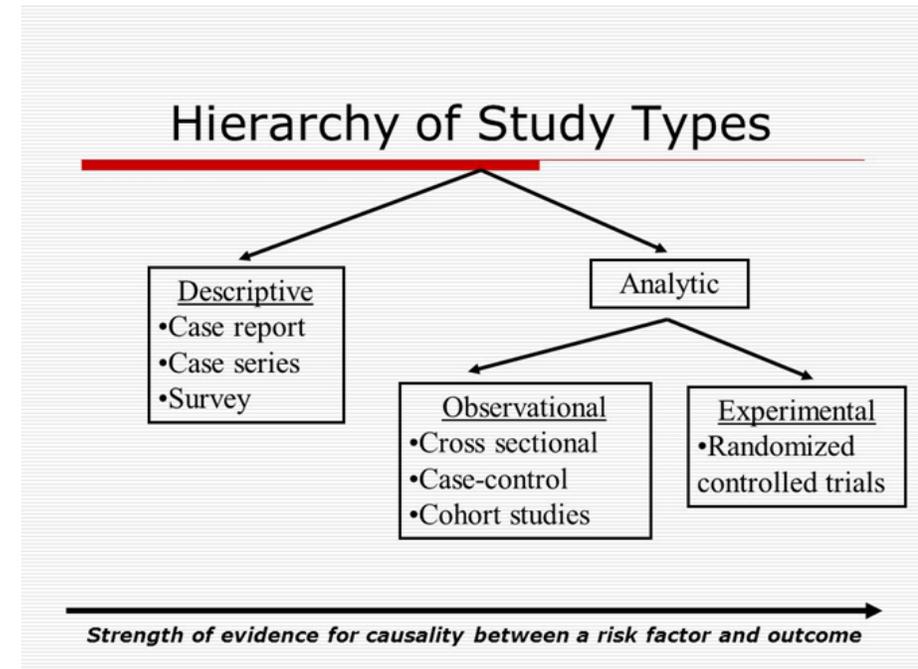
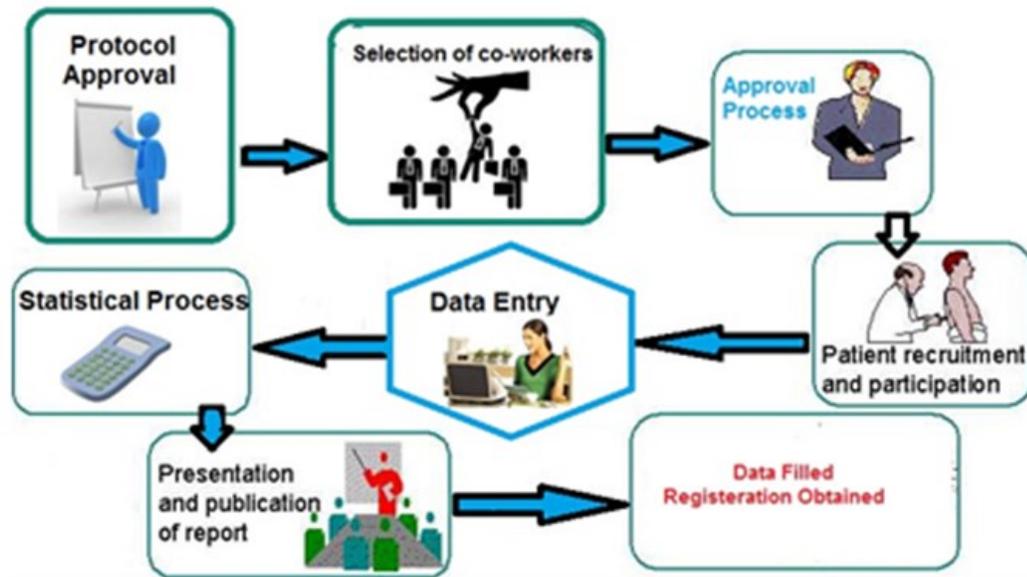


Lo statistico deve lavorare in un **team** con i ricercatori esperti dell'argomento per poter indirizzare verso il disegno di studio più adeguato e poter poi condurre in modo appropriato le analisi sui dati raccolti.

- Quale è la popolazione target?
- Quali gruppi si vogliono **descrivere/confrontare**?
- Quali parametri si vogliono **misurare** ?
- **Quando**/in che occasione/i si vogliono misurare ?
- Quali parametri si ha bisogno di “**controllare**” per confrontare i gruppi ?
- ...

Concetti Generali sul **Disegno dello Studio**: terminologia

- Studi Sperimentali (cenni)
- Studi Osservazionali (cenni)



TERMINOLOGIA:

Outcome : il risultato di interesse per i soggetti (**unità statistiche**) dello studio (=> fenomeno di interesse).

- ◁ es : evento (diagnosi di cancro al cavo orale)
- ◁ es : la resistenza di un materiale per protesi dentarie

Predittore/i: la/le variabile/i che si pensa siano possibili “**cause**” dell’outcome

- ◁ es : l’abitudine al fumo
- ◁ es : la composizione chimica del materiale

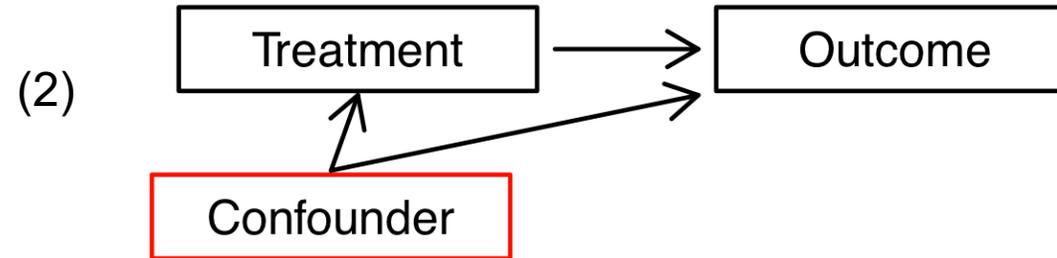
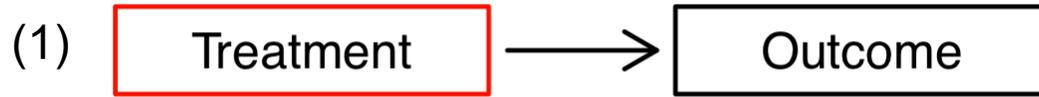
Confondente: ciò che può “**confondere**” la relazione tra il predittore e l’outcome (associata sia all’outcome che al predittore).

- ◁ es: consumo di alcolici
- ◁ es: il tipo di rivestimento utilizzato per la protesi

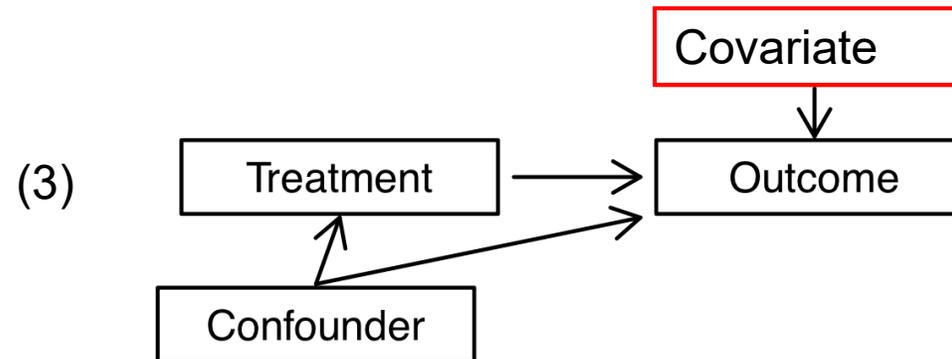
Covariata: spiega parte della **variabilità** dell’outcome (ma non è associata al predittore di interesse).

- ◁ es: età dell’individuo
- ◁ es: abitudini alimentari dei soggetti con protesi...

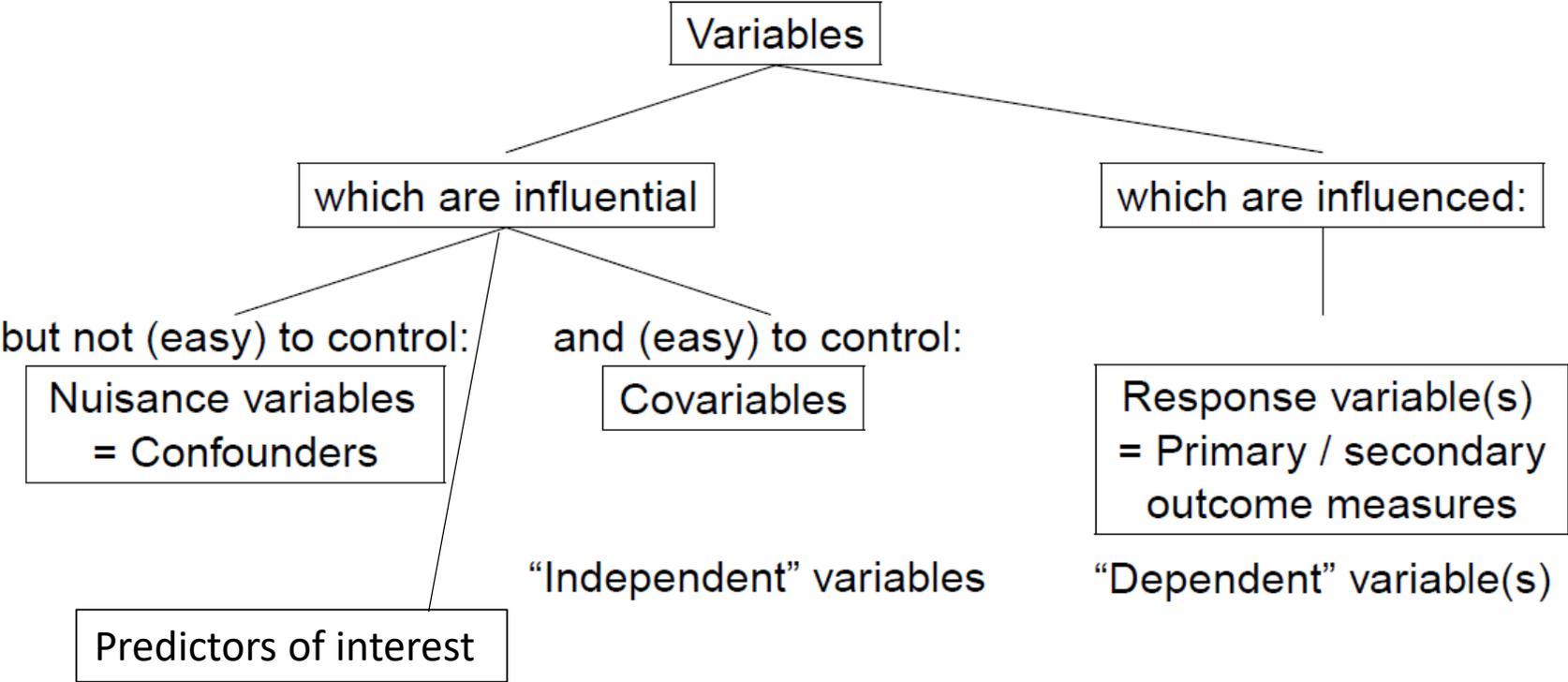




Livello crescente di
«**complessità**»
del disegno dello studio
da (1) a (3)



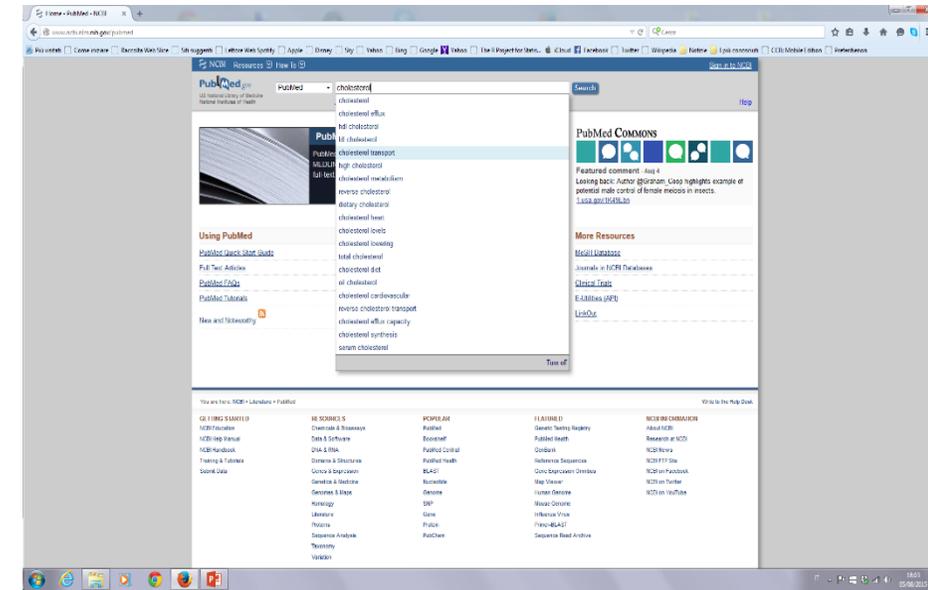
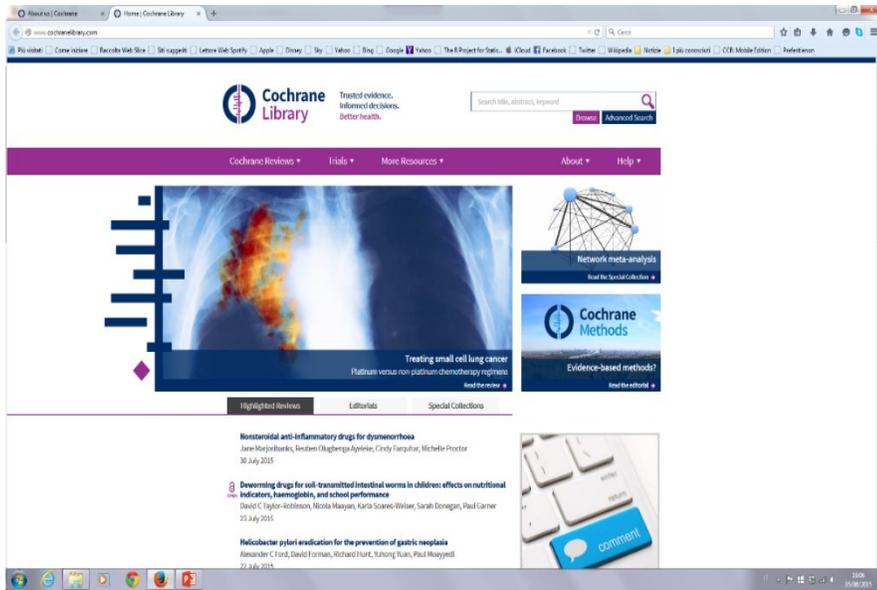
Le 'Variabili' in un Disegno di Studio:



Prima di cominciare: revisione sistematica della letteratura !!

Revisione **aggiornata** della letteratura sull'argomento da studiare:

- **Background** sullo stato dell'arte della ricerca sul fenomeno, (referenze -> **bibliografia**) ; esaminare **in modo critico** risultati e metodi degli studi pubblicati.
- **Stime quantitative** sull'outcome..."effetto atteso" (-> servirà per determinare la **dimensione campionaria** opportuna)



Una prima distinzione tra due tipi di studi:

(1) **STUDI SPERIMENTALI**

(2) **STUDI OSSERVAZIONALI**

STUDI SPERIMENTALI: il ricercatore «interferisce» con il corso naturale degli eventi cosicché l'effetto sperimentale possa essere osservato.

STUDI OSSERVAZIONALI : il ricercatore è un osservatore passivo e si limita a raccogliere dei dati sul fenomeno che deve essere studiato.

- Indagine sui tipi di veicoli a motore che passano per un casello stradale durante un intervallo di tempo; (*studio osservazionale*)
- Sondaggio di opinione (*studio osservazionale*)
- Confronto tra tempi di sopravvivenza di cavie dopo avere inoculato una dose di una sostanza tossica; (*studio sperimentale*)
- Confronto tra intervento chirurgico vs farmaci per pazienti in una certa condizione, avendo assegnato in modo *casuale* i soggetti ai due trattamenti (*studio sperimentale*)

Gli studi sperimentali

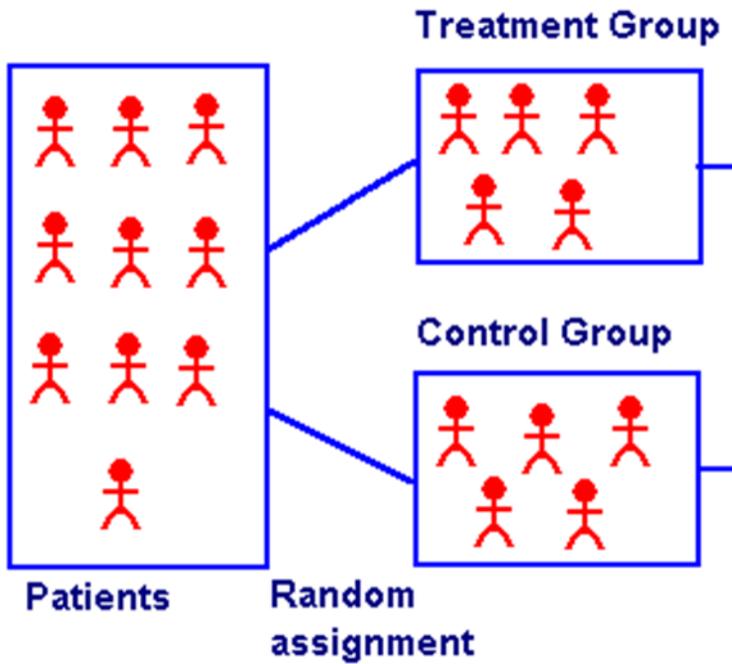
Studi che vengono pianificati con lo scopo di confrontare gli effetti di diversi trattamenti sullo stesso tipo di unità sperimentali.

Sono studi di tipo **analitico** (si indaga la presenza di **relazioni** fra fenomeni).

In particolare in medicina tali esperimenti si definiscono con il termine inglese: ***clinical trials***.

- confronto sull'effetto in diversi animali di diverse dosi una sostanza chimica;
unità=gli animali, **trattamento**= le diverse dosi della sostanza.
- studio di profilassi per confrontare l'efficacia sui bambini di diversi vaccini contro il morbillo;
unità=bambini, **trattamento**=vaccino.
- confronto in pazienti che soffrono di emicranie di diversi metodi per alleviare il dolore;
unità=i pazienti, **trattamento**= i vari metodi anti-dolore.
- studio sui vantaggi e svantaggi relativi a diversi programmi di educazione alla salute;
unità=aree, **trattamento**= i vari programmi di educazione sanitaria

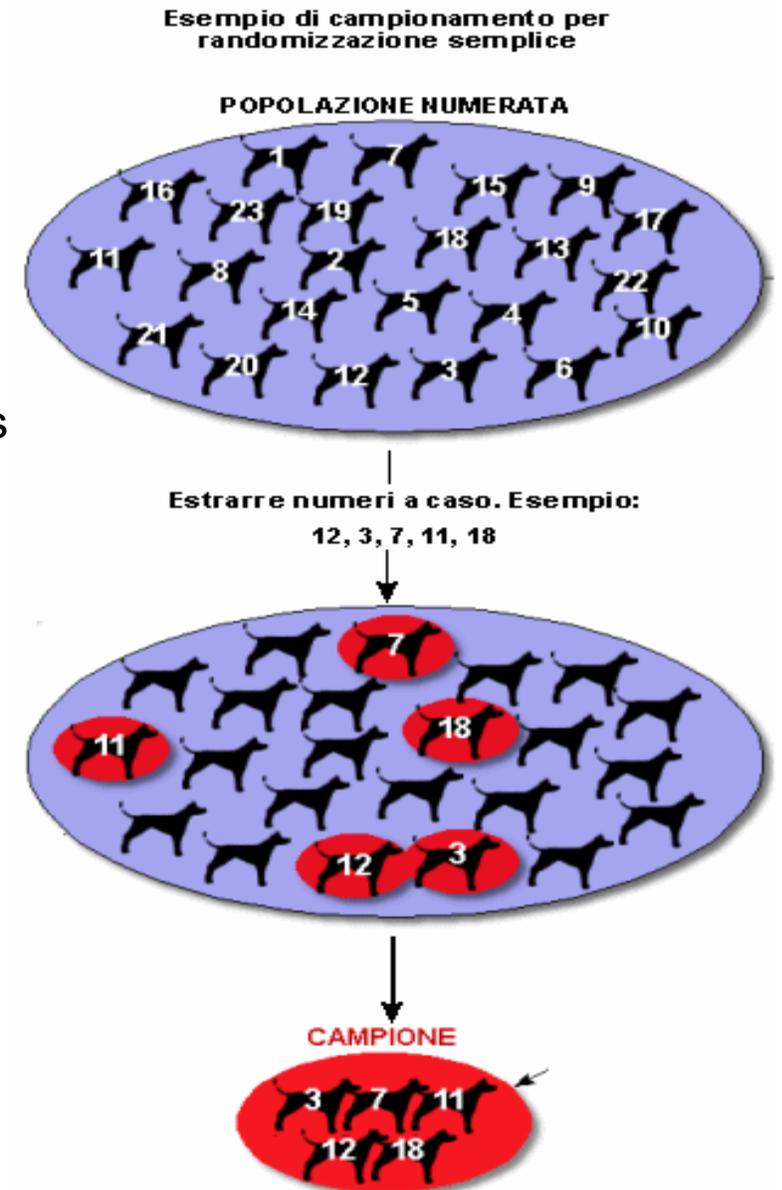
...come **assegnare** i trattamenti alle unità disponibili?



<https://youtu.be/MmpF1zxfQZ8>
 Explaining Randomization in Clinical Trials

I gruppi di unità sperimentali a cui vengono assegnati i diversi trattamenti devono essere il più *simili* possibile, almeno negli aspetti più rilevanti per la sperimentazione...

Soluzione (semplice) = **randomizzazione**



Quante unità assegnare ad ogni trattamento?
Quanto grande deve essere il campione?

In generale occorre infatti la **replicazione** delle osservazioni, più di una unità sperimentale per ogni trattamento.

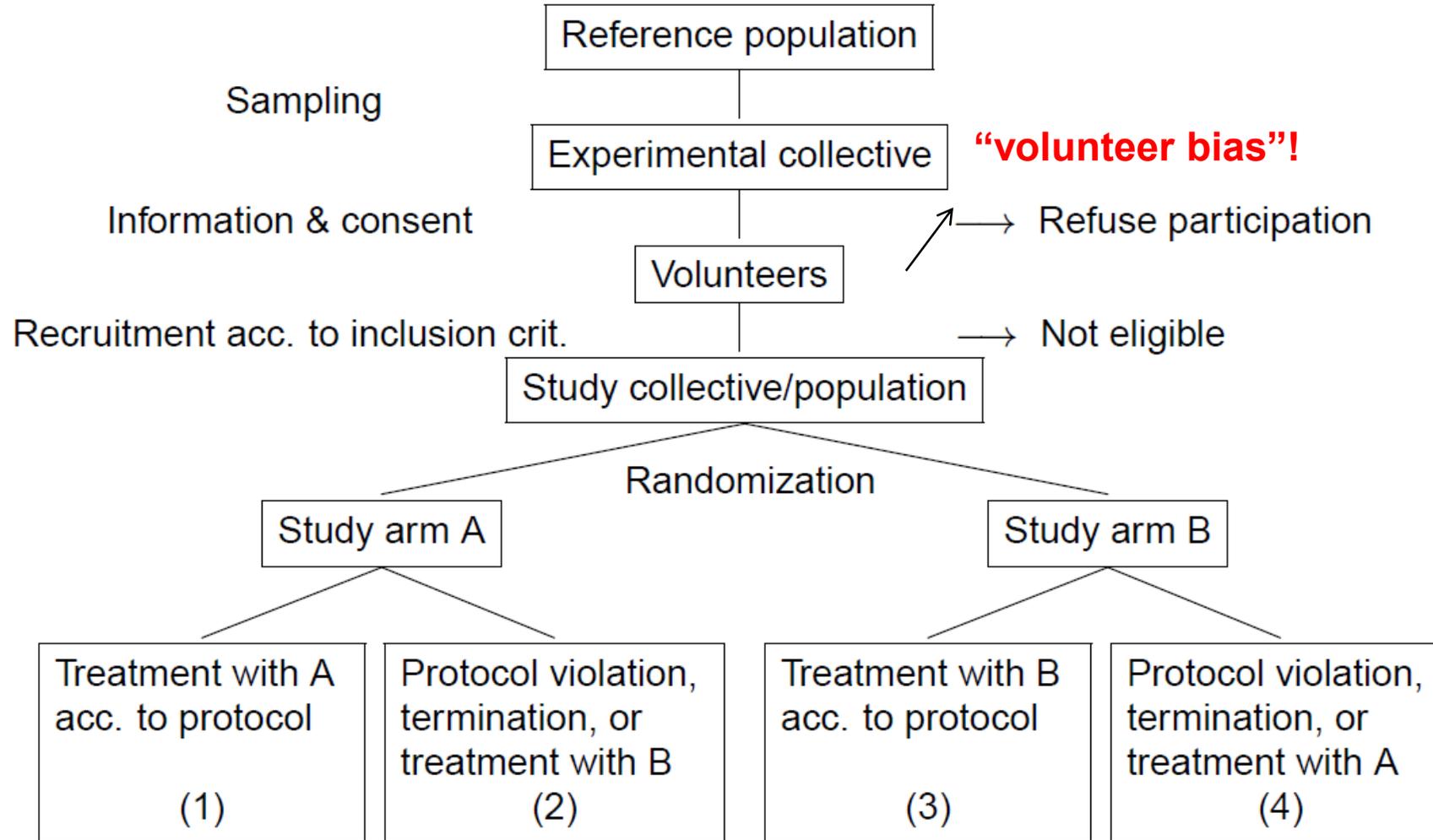
Problema complicato... la «dimensione campionaria»*...

*Per risolverlo serve conoscere alcuni metodi (base) di inferenza statistica

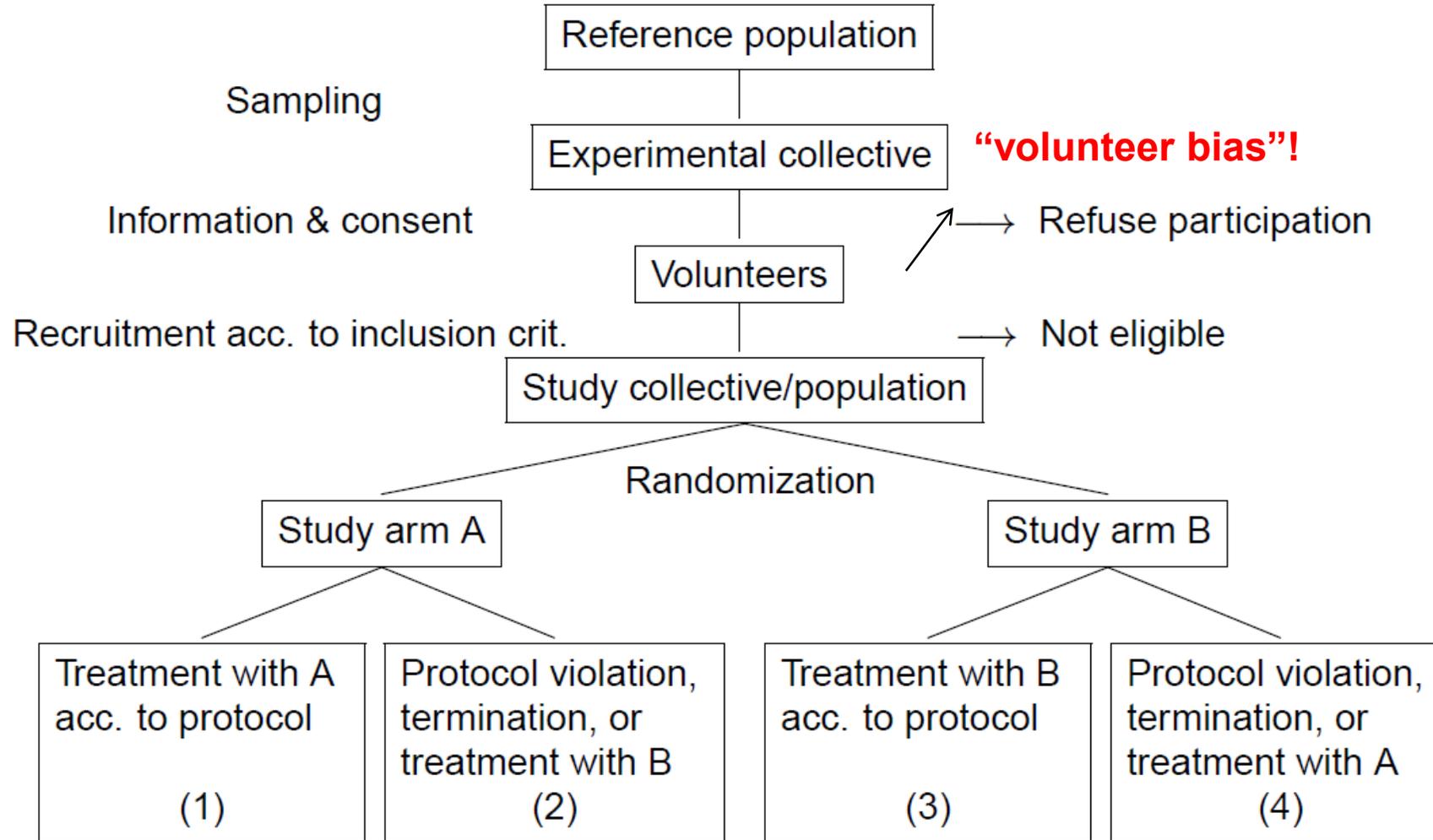
...



Studi Clinici Sperimentali (Clinical Trials): SCHEMA



Studi Clinici Sperimentali (Clinical Trials): SCHEMA



Studi Clinici Sperimentali (Clinical Trials): BLINDING (MASKING)

Per minimizzare la soggettività di chi valuta i pts o analizza i dati.

- **SINGLE BLIND:** i soggetti dello studio non sanno quale trattamento ricevono
- **DOUBLE BLIND:** i soggetti e i ricercatori (medici/biologi) non sanno
- **TRIPLE BLIND:** i soggetti, i ricercatori, gli statistici non sanno

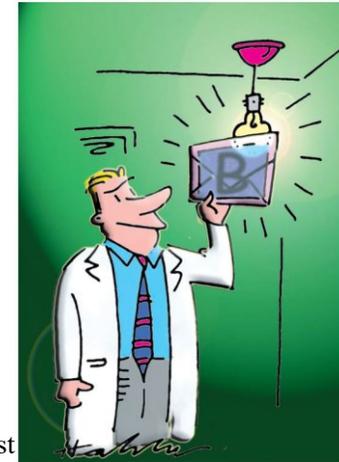


Figure 1: The authors: double blinded versus single blinded

Schulz & Grimes. Lancet 2002

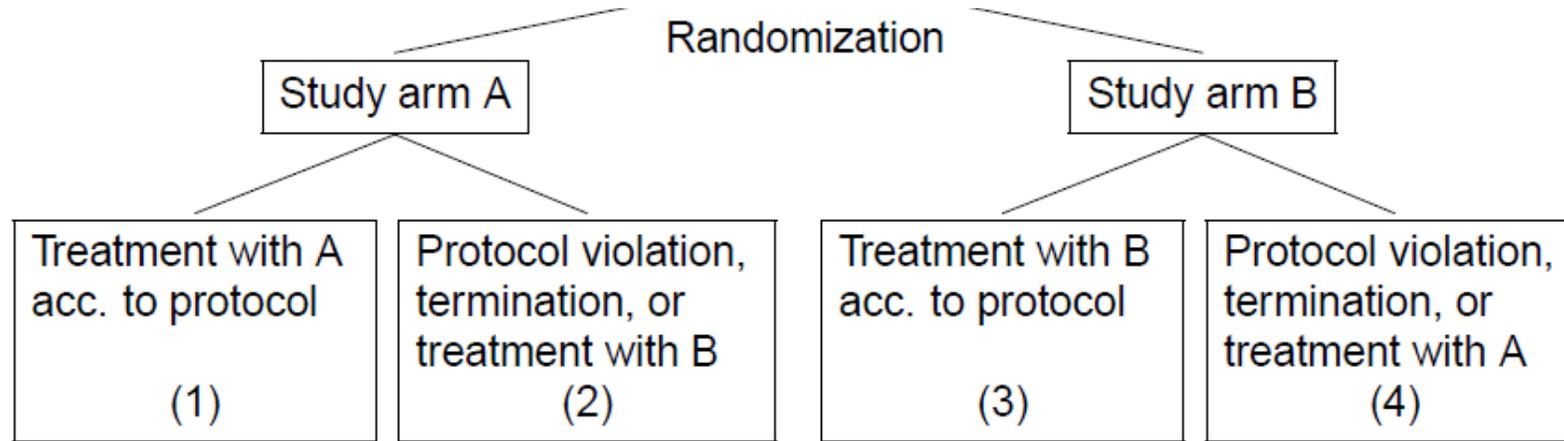
Se non viene effettuata: **OPEN TRIAL**

Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. Lancet. 2002 Feb 16;359(9306):614-8



Deciphering the allocation concealment scheme

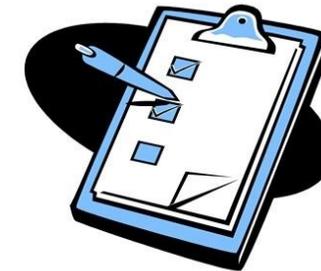
Studi Clinici Sperimentali (Clinical Trials): Analisi dei dati



ITT: intention to treat
(1)+(2) vs (3)+(4)

PP: per protocol (1) vs (3)

AT: as treated (1)+(4) vs (2)+(3)
(ma escludendo violazioni del protocollo o chiusure precoci)



Studi Clinici Sperimentali (Clinical Trials): limiti

- **Mis-classificazione/ 'information bias'** sull'outcome o sull'esposizione;
- **Co-interventi**: se un braccio riceve un intervento addizionale più frequentemente rispetto all'altro;
- **Contaminazione**: quando una parte di soggetti di «*controllo*» ricevono il trattamento al di fuori dello studio

Inoltre:

«**Volunteer bias**»: partecipano solo soggetti volontari (eligibilità allo studio, *compliance*, vicinanza geografica, stato socio-economico, stato di salute...)

Popolazione omogenea selezionata -> validità interna (bassa variabilità)
Ma: limite per la validazione esterna (generalizzabilità) dei risultati.

EFFICACY ≠ EFFECTIVENESS



Successo di un trattamento in un contesto
«artificiale»



Successo in un contesto
«**real world**»

STUDI OSSERVAZIONALI:

Possono essere suddivisi in due macro-categorie:

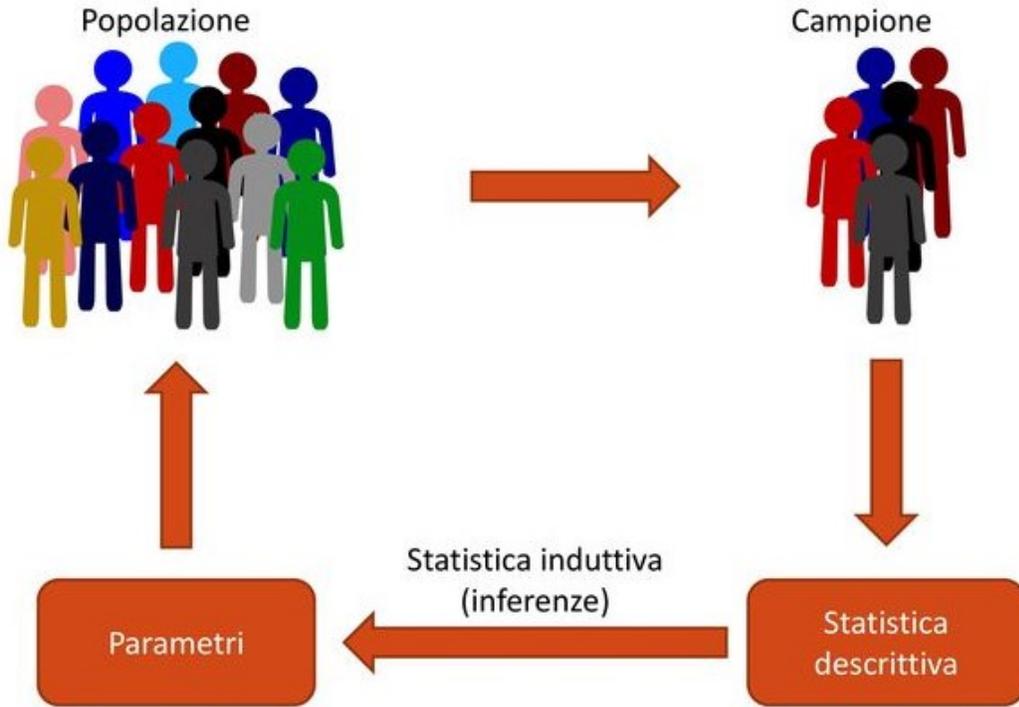
- (1) Studi **descrittivi**: ottenere delle stime (misure) di alcune caratteristiche della popolazione target;
- (2) Studi **analitici**: analizzare le relazioni tra fenomeni di interesse nella popolazione target.

Esempi:

- (1) Studi per stimare la **diffusione** di una certa malattia in una popolazione o la **frequenza** di effetti avversi ad un certo farmaco nella popolazione di utilizzatori;
- (2) Studio dell'**associazione** tra il consumo di un certa sostanza e l'insorgenza di patologie specifiche.

In (1) ci concentriamo sulla precisione del risultato,
-> stima attendibile del fenomeno sulla popolazione;

In (2) ci focalizziamo sulla relazione tra le variabili



Così come negli **studi clinici randomizzati** i pazienti vengono suddivisi in modo casuale nei gruppi da confrontare, così anche negli studi osservazionali un **campione casuale** viene (dovrebbe...) estratto dalla **popolazione di riferimento**.

<https://youtu.be/pTuj57uXWlk>

Types of Sampling Methods

Aspetti fondamentali del campionamento:

- Ogni soggetto della popolazione di riferimento deve avere la stessa probabilità di essere incluso nel campione
- La probabilità di essere incluso nel campione NON deve dipendere da particolari caratteristiche del soggetto.



Campionamento semplice casuale (alla base di tutti gli schemi, anche più complessi)

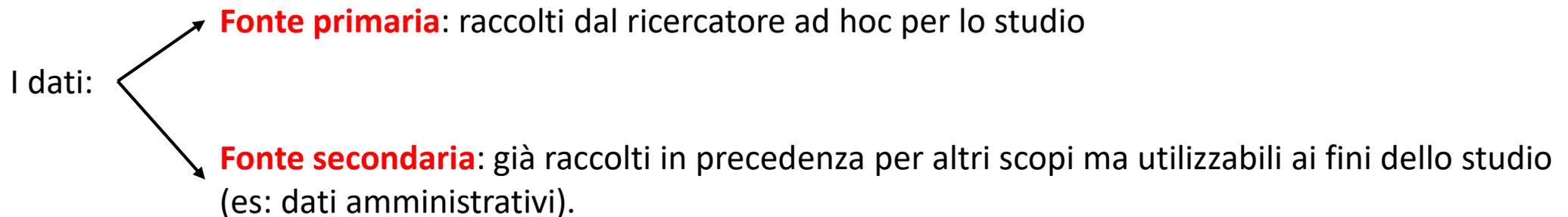
STUDI OSSERVAZIONALI

Gli studi osservazionali analizzano l'effetto di una «esposizione» o di un trattamento/intervento su dei soggetti.

Non c'è randomizzazione; non c'è manipolazione da parte del ricercatore.

Osservazione diretta del «**real world**»....

Confondimento: esiste una spiegazione alternativa ai risultati osservati?



Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ 1996;312:1215–1218.

Haynes B. *Can it work? Does it work? Is it worth it?* BMJ 1999;319:652.

Primary Data

Advantages

Investigator controls all aspects of the study including design, sampling techniques, data collection, and follow-up methods

All variables of interest can be measured

Disadvantages

Time consuming

Expensive

Secondary Data

Advantages

Relatively fast and inexpensive

Sample sizes tend to be large

Sample may cover a large geographic area and thus provide ability to assess national trends

Unobtrusive to subjects

Disadvantages

Data may not include all variables of interest

Maybe difficult to understand how and why data elements were collected

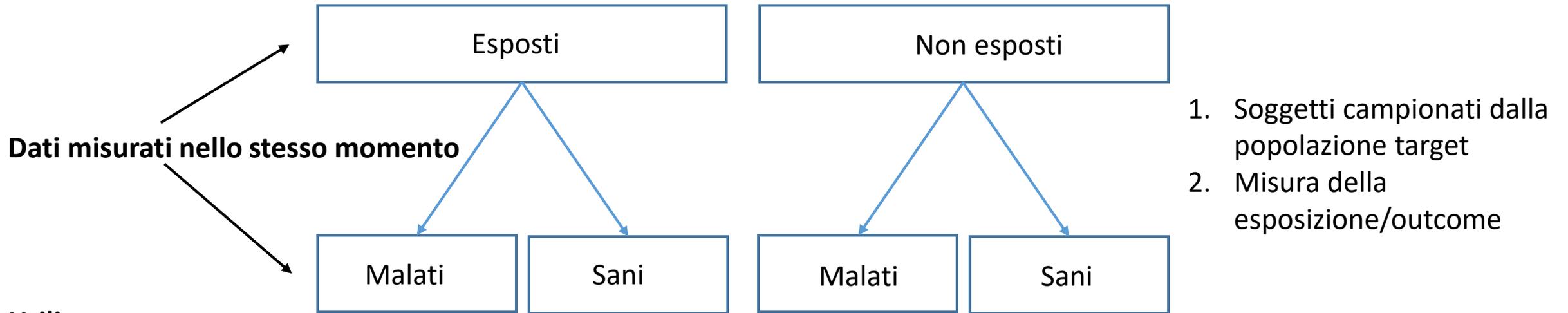
Gli studi osservazionali possono essere suddivisi in diverse **categorie** sulla base di :

- **Obiettivo** della ricerca;
- **Come** vengono **selezionati** i soggetti;
- **«Timeline»** della raccolta dei dati

- **CROSS-SECTIONAL**
- **COORTE**
CASO-COORTE
- **CASO-CONTROLLO**
CASE-CROSSOVER

Studi Cross-sectional

Studio osservazionale in cui l'esposizione e l'outcome sono determinati **simultaneamente** per ogni soggetto.



Utili per:

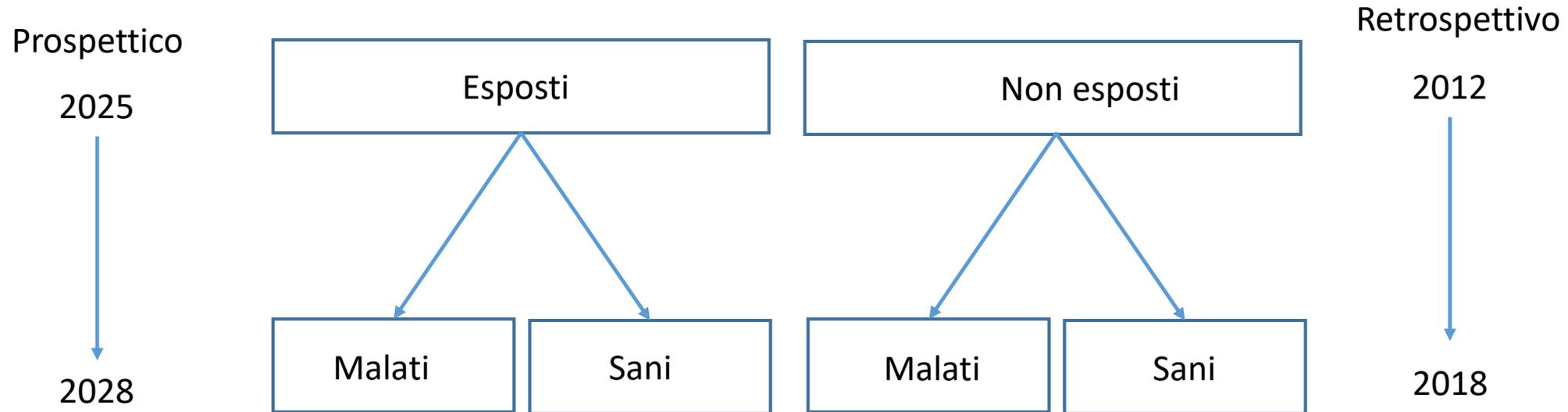
- screening delle ipotesi sul quesito in studio, rapidi da realizzare;
- valutazioni di prevalenza;
- valutazioni di fattori associati all'outcome di interesse.

Limiti:

- Non c'è evidenza temporale di causa-effetto tra l'esposizione e l'outcome;
- Outcomes **prevalenti** ma non incidenti, esclusi i soggetti deceduti prima dello studio (bias «long survivors»);
- A causa della presenza di possibili confondenti bisogna valutare spiegazioni alternative alle associazioni osservate

Studi di coorte

1. Soggetti campionati dalla popolazione di esposti e di non esposti
2. Misura dell'outcome



Utili quando:

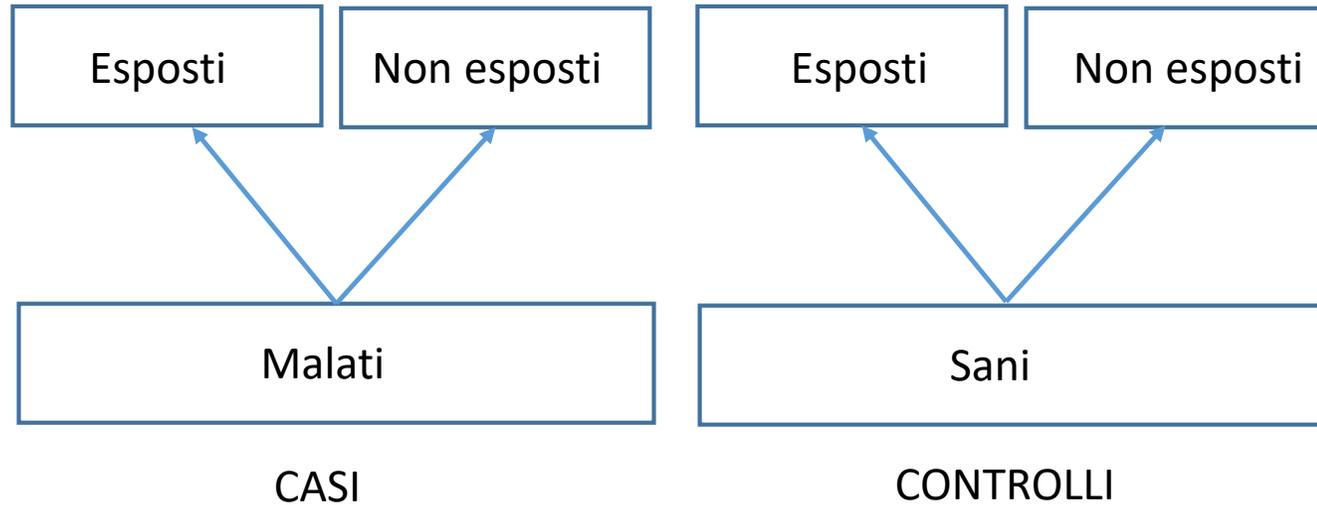
- C'è evidenza di una associazione temporale tra esposizione ed outcome;
- L'intervallo temporale tra esposizione e outcome non è troppo lungo;
- L'outcome non è troppo raro.

Limiti:

- Perdita al follow up;
- A causa della presenza di possibili confondenti bisogna valutare spiegazioni alternative
- Bias nella determinazione dell'outcome: non sono blinded!

Studi caso-controllo

1. Soggetti campionati dalla popolazione di casi e di controlli
2. Misura della esposizione



Utili quando:

- L'outcome è raro;
- C'è una ragionevole evidenza di una esposizione nel passato;

Limiti:

- «Recall/Interviewer bias»: i casi possono essere più precisi nel riferire sulla (passata) esposizione;
- Difficile la selezione di controlli adeguati;
- Se si utilizzano procedure di «**matching**» da valutare la generalizzabilità dei risultati;
- Confondenti : bisogna valutare spiegazioni alternative alle associazioni osservate

Stima di una **misura di rischio** (RCTs/studio osservazionale caso-controllo/coorte)

	MALATI	SANI	Totale
ESPOSTI	a	b	a+b
NON ESPOSTI	c	d	c+d
Totale	a+c	b+d	N

Come **quantificare** l'associazione fra il fattore di esposizione e l'outcome ?

Ipotetica Tabella di contingenza a conclusione del Trial:

	MALATI	SANI	Totale
TRATTATI	a	b	a+b
NON TRATTATI	c	d	c+d
Totale	a+c	b+d	N

Come **quantificare** l'associazione fra il trattamento e l'outcome ?
(ipotizzando l'assenza di fattori di confondimento, analisi «univariata»)