

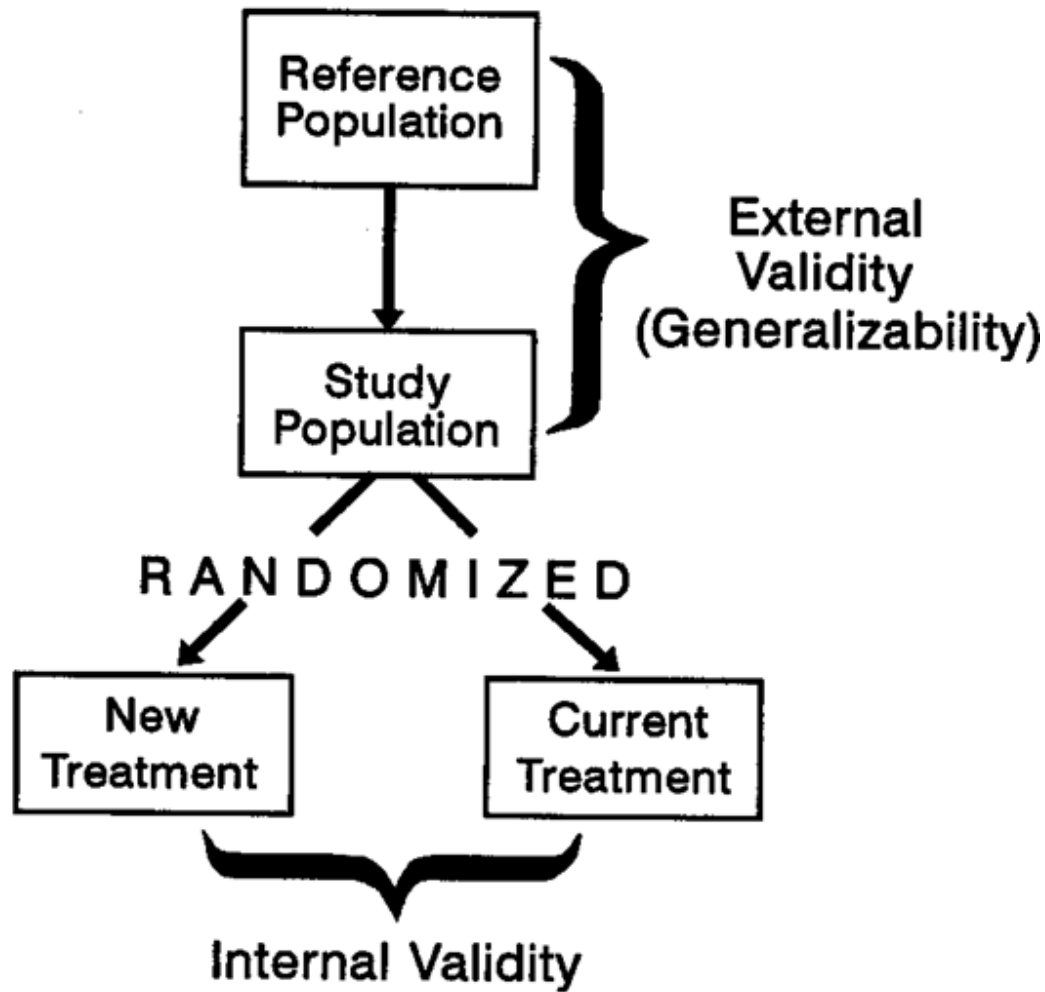
Elementi critici RCT



SAPIENZA
UNIVERSITÀ DI ROMA

annarita.vestri@uniroma1.it

Elementi critici degli studi randomizzati



Bad statistics is bad ethics

Non è etico condurre ricerca mal pianificata o mal eseguita

La ricerca deve soddisfare predefiniti standard scientifici

garantiti da un Comitato Etico

Frequenti caratteristiche che identificano studi di bassa qualità:

- **Disegno incoerente rispetto agli obiettivi**
- **Presenza di distorsioni (nella selezione dei soggetti, nella definizione dei trattamenti di confronto, nella valutazione della risposta sperimentale, nella individuazione degli insiemi di soggetti da sottoporre ad analisi, ecc.)**
- **Pochi soggetti**
- **Nessuna pubblicazione dei risultati**

Elementi critici degli studi randomizzati

- **Metodi di randomizzazione (generazione e realizzazione)**
- **Modalità valutazione intervento: “cieco” / “aperto”**
- **Scelta degli indicatori di esito (“end points”)**
- **Analisi dei dati (per trattamento ricevuto - “per protocol” - o trattamento assegnato - “intention to treat”)**
- **Numerosità del campione**
- **Tipo di popolazione studiata (generalizzabilità risultati)**
- **Scelta del trattamento/i di controllo**
- **Interpretazione dei risultati (significatività statistica vs rilevanza clinica)**

RANDOMIZZAZIONE

- Rappresenta, assieme al controllo, la caratteristica fondamentale dei trial clinici
- I pazienti, reclutati sulla base dei criteri di inclusione ed esclusione stabiliti a priori, vengono assegnati al trattamento sperimentale o a quello di controllo mediante una forma più o meno sofisticata di sorteggio (ad esempio generazione da parte di un computer di sequenze numeriche casuali)
- La procedura di randomizzazione deve essere tale da rendere imprevedibile a quale trattamento verrà assegnato il paziente successivo

RANDOMIZZAZIONE

- Con la randomizzazione si riescono ad ottenere (almeno teoricamente) gruppi omogenei tra di loro per tutte le caratteristiche note ed ignote (es. età, sesso, gravità patologia, abitudini di vita, ecc.)
- L'omogeneità tra i gruppi sarà maggiore anche in relazione alla numerosità del campione

MODALITA' DI VALUTAZIONE DELL'INTERVENTO

Singolo cieco/doppio cieco

Nascondere al paziente, allo sperimentatore (medico) o ad entrambi (doppio cieco) la conoscenza del trattamento allocato

Razionale: il paziente potrebbe essere influenzato dal sapere quale trattamento ha ricevuto

... ma anche i medici potrebbero valutare diversamente la condizione del paziente, sapendo quale trattamento ha ricevuto

... oppure evitare al paziente un trattamento da cui ritengano che il paziente stesso non possa trarre benefici

La possibilità di mantenere la “cecità” dipende dal tipo di trattamento

SCELTA INDICATORI DI ESITO

- esiti clinicamente importanti (mortalità - totale o causa-specifica, eventi non fatali)
- esiti secondari (end-point “surrogati”): pressione arteriosa, colesterolemia, etc
- qualità della vita (problema della validazione degli strumenti/questionari)
- Esiti “indiretti” Es. utilizzo dei servizi sanitari

Popolazione di analisi

- **Si intende l'insieme dei pazienti su cui effettuare le analisi statistiche**

In genere sono

- **Intention to treat**
- **Per protocol**

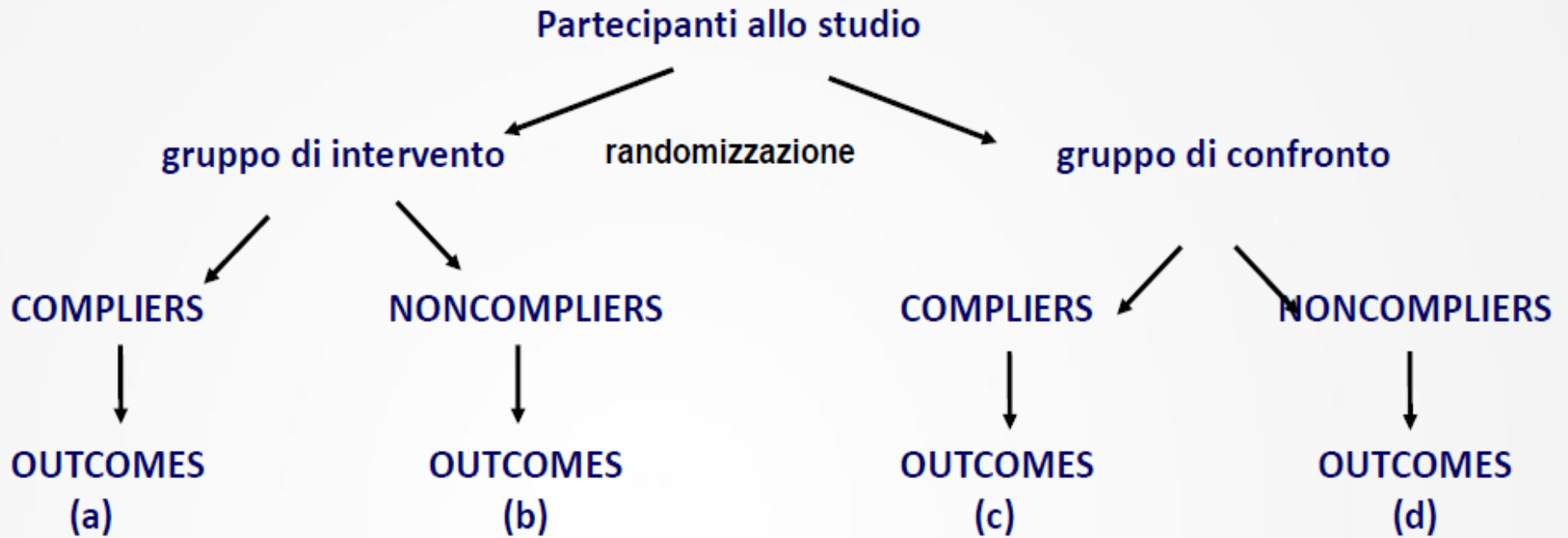
Popolazione di analisi

- **Intention to treat:** i pazienti vengono valutati in base ai gruppi ai quali erano stati originariamente assegnati (randomizzazione)
- L'analisi effettuata su questa popolazione deve rispondere alla domanda: **qual è l'effetto del trattamento in condizioni simili a quelle reali?**

Popolazione di analisi

- La popolazione “**per protocol**” considera solo i pazienti che hanno rispettato i principali criteri di selezione e le principali procedure del protocollo
- L’analisi su questa popolazione risponde alla domanda: qual è l’effetto del trattamento nelle migliori condizioni sperimentali?

ANALISI



ANALISI PRIMARIA: Intention-to-treat analysis: "una volta randomizzato sempre analizzato". (a+b) versus (c+d)

ANALISI SECONDARIA: Per Protocol analisi dei soggetti che arrivano al termine della sperimentazione: (a) versus (c).

Attenzione: non è un gruppo randomizzato

NUMEROSITA' DEL CAMPIONE

Il corretto dimensionamento del campione è punto fondamentale e richiede uno sforzo da parte del clinico nel definire esplicitamente in termini quantitativi gli obiettivi che lo studio si prefigge

➤ QUANTO miglioramento ci si aspetta dal nuovo trattamento?

NUMEROSITA' DEL CAMPIONE

L'idea che sta alla base dei calcoli delle dimensioni del campione è quella di avere una

“high chance of detecting, as statistically significant, a worthwhile effect if it exists”

Quindi di essere ragionevolmente sicuri che quel beneficio non esiste se non lo troviamo nel trial

NUMEROSITA' DEL CAMPIONE

I soggetti arruolati in uno studio devono essere “abbastanza numerosi” in modo che un effetto di una dimensione sia tale da essere “**scientificamente significativo**” ed anche “**statisticamente significativo**”

Ma è anche importante che gli arruolati non siano “troppi”, tanti da rendere statisticamente significativo un risultato di scarsa rilevanza scientifica

La numerosità campionaria è importante anche per motivi economici

Uno studio sottodimensionato può portare ad uno spreco di risorse, perché non riuscirà a produrre risultati utili

Uno studio sovradimensionato usa comunque più risorse del necessario

NUMEROSITA' DEL CAMPIONE

- **Potenza**
- **Livello di significatività**
- **Tasso atteso dell'evento nella popolazione in studio**
- **Effetto atteso del trattamento**
- **Variabilità della differenza**

POTENZA STATISTICA

- **La potenza di uno studio è la capacità di individuare una vera differenza nell'outcome tra i gruppi a confronto**
- **Una potenza dell'80% significa che si accetta di non individuare una reale differenza nel 20% dei casi**
- **La potenza in genere si fissa al 90% per ridurre al 10% la possibilità di un risultato falso negativo**

POTENZA STATISTICA

La POTENZA di un test dipende:

- 1) dalla numerosità del campione**
- 2) dalla variabilità del fenomeno in studio**
- 3) dalla differenza minima che si vuole mettere in evidenza**
- 4) dal livello di significatività adottato**

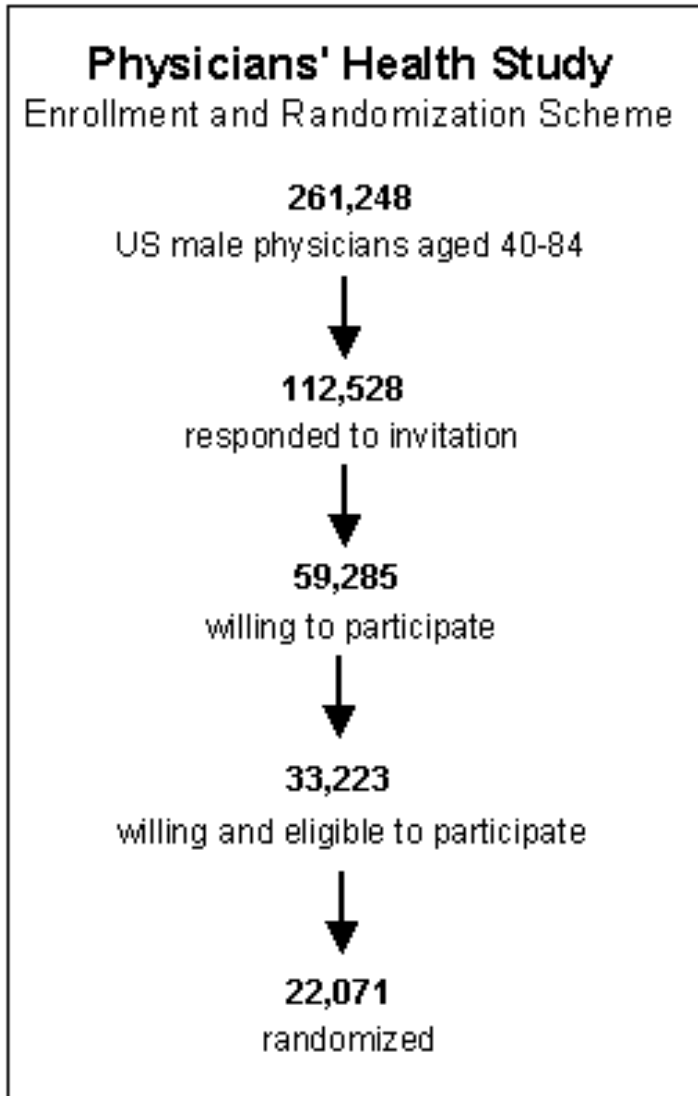
Il modo principale per raggiungere un'adeguata potenza è pianificare un'adeguata numerosità campionaria nel protocollo dello studio.

LIVELLO DI SIGNIFICATIVITA'

Indica il rischio di individuare l'effetto del trattamento quando non esiste (risultato falso-positivo) e definisce il valore soglia di P

Risultati con un valore di P maggiore del livello prefissato conducono alla conclusione che una differenza osservata può essere dovuta all'effetto del caso

TIPO DI POPOLAZIONE STUDIATA



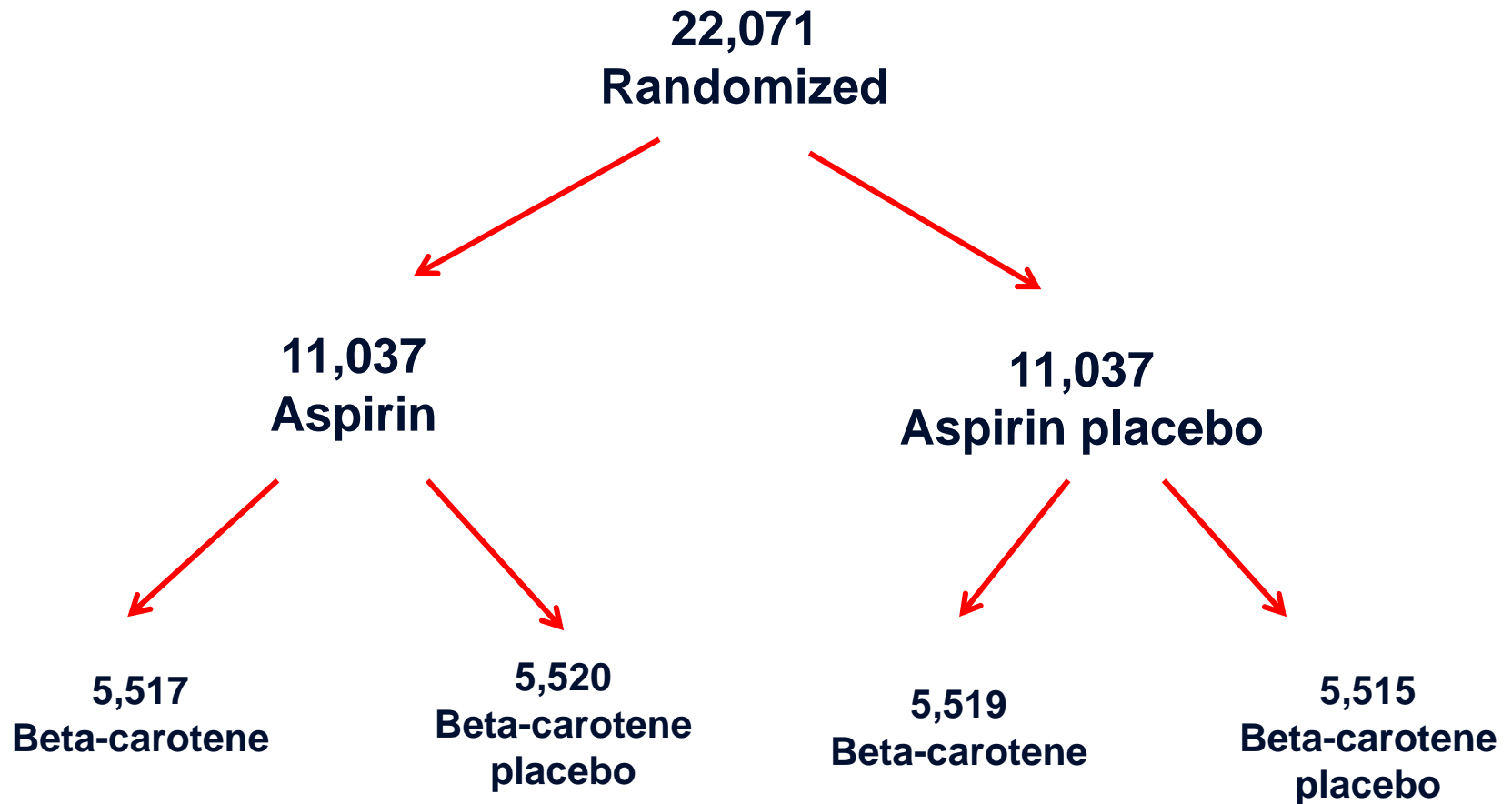
[N Engl J Med. 1989 Dec 28;321\(26\):1825-8.](#)

[N Engl J Med. 1989 Jul 20;321\(3\):183-5.](#)

Final report on the aspirin component of the ongoing Physicians' Health Study. Steering Committee of the Physicians' Health Study Research Group.

PHYSICIANS' HEALTH STUDY

Randomization Scheme



PHYSICIANS' HEALTH STUDY I

Assignment	ASA + BC (n = 5,517)	ASA + PL (n = 5,520)	PL + BC (n = 5,519)	PL + PL (n = 5,515)
Demographic				
Mean age (years)	53.2±9.5	53.3±9.6	53.2±9.5	53.2±9.5
Geographic region				
Northeast	27.8	27.9	27.3	28.8
Midwest	22.4	21.9	22.1	20.3
South or Southwest	26.8	26.1	27.0	27.9
Mountain or Far West	22.0	23.3	22.8	22.1
Other	1.0	0.9	0.7	0.9

PHYSICIANS' HEALTH STUDY II

Assignment	ASA + BC	ASA + PL	PL + BC	PL + PL
Medical history				
Reported hypertension	11.8	11.8	12.4	11.5
Mean systolic BP	126.0±11.9	126.3±12.1	126.1±11.7	126.1±11.7
Mean diastolic BP	78.8±7.6	78.9±7.5	78.9±7.5	78.7±7.5
High cholesterol	7.9	7.5	8.0	8.0
Mean cholesterol (mg/dl)	211.7±43.2	212.3±45.0	212.7±46.5	211.4±44.8
Diabetes mellitus	2.1	2.5	2.2	2.2
Mean BMI	24.9±3.1	24.9±3.0	24.9±3.0	24.9±3.0
Overweight	13.5	14.0	13.6	13.6
History of angina pectoris	1.2	1.4	1.1	1.2
Parental history of MI	12.7	13.3	12.7	13.6

PHYSICIANS' HEALTH STUDY III

Assignment	ASA + BC	ASA + PL	PL + BC	PL + PL
Health habits				
Cigarette smoking				
Never	49.8	48.8	49.6	50.1
Past only	39.3	40.1	39.1	39.0
Current	10.9	11.1	11.4	10.9
Mean cigarettes/day	21.7 ± 13.0	22.1 ± 12.9	22.4 ± 13.7	22.8 ± 13.5
<15 (of current smokers)	28.8	27.4	27.3	24.6
15-24 (of current smoke	32.5	33.7	34.7	35.4
25+ (of current smokers)	38.7	38.8	37.9	40.0
Current cigar smoking	4.7	4.2	4.3	4.3
Current pipe smoking	8.8	7.7	8.2	8.0

PHYSICIANS' HEALTH STUDY IV

Assignment	ASA + BC	ASA + PL	PL + BC	PL + PL
Health habits				
Alcohol use				
Mean drinks/week	3.5±3.2	3.5±3.2	3.5±3.2	3.5±3.3
Daily	24.8	24.9	24.7	25.2
Rarely or never	14.8	14.6	14.7	15.4
Exercise frequency				
Mean times/month	10.2±8.5	9.8±8.3	10.0±8.6	10.0±8.6
Rarely or never	12.8	14.2	14.0	13.5

PHYSICIANS' HEALTH STUDY

This trial of aspirin for the primary prevention of cardiovascular disease demonstrates a conclusive reduction in the risk of myocardial infarction, but the evidence concerning stroke and total cardiovascular deaths remains inconclusive because of the inadequate numbers of physicians with these end points.

A che tipo di popolazione vanno applicati quei risultati?

- **Medici**
- **Maschi**
- **Giovani con poche patologie**
- **Bassa prevalenza di FR (peso, ore di attività fisica, fumo, alcool)**
- **Attenti alla prevenzione**
- **Meticolosi nell'assumere il farmaco (per 8 anni)**
- **Al termine dello studio si ottenne meno di ¼ delle morti attese nel gruppo di controllo. La vera conclusione quindi sarebbe: condurre una vita sana (non fumare, non bere, fare attività motoria, ecc.) è il miglior intervento di prevenzione cardiovascolare e neoplastica**

I risultati degli studi nella pratica clinica quotidiana

EFFICACIA CLINICA

(CONDIZIONI “IDEALI” *EFFICACY*)

VS

EFFICACIA NELLA PRATICA CLINICA (EFFECTIVENESS)

- paziente “reale”, non selezionato
- co-morbidità
- problemi di compliance
- continuità del trattamento

Interpretazione dei risultati

Un risultato statisticamente significativo è sempre rilevante dal punto di vista clinico?

Valutare in modo critico la rilevanza pratica dei risultati degli studi

Quando non è possibile fare un RCT

- **Verifica di tossicità di una esposizione/trattamento**
- **gli esiti da valutare sono troppo lontani nel tempo**
- **gli esiti da valutare sono rari (servirebbe un campione troppo ampio)**
- **selezione dei pazienti: preferenza per uno dei trattamenti, con conseguente rifiuto alla randomizzazione**
- **fondi non sufficienti**