



SAMPLE SIZE: PERCHÉ È IMPORTANTE E COME SI CALCOLA

Perché parlare di sample size?

Ogni studio clinico nasce per rispondere ad una domanda. La risposta a questa domanda dipende dai dati raccolti, ma anche dal numero di partecipanti inclusi. Questo numero prende il nome di sample size (dimensione del campione) e rappresenta uno degli elementi più importanti nella progettazione di uno studio.

Un campione troppo piccolo rischia di non individuare un effetto realmente esistente (errore di tipo II), mentre un campione eccessivamente grande può identificare come statisticamente significative differenze prive di reale importanza clinica, oltre a comportare costi maggiori, tempi più lunghi ed esposizione non necessaria dei partecipanti.

In altre parole, il sample size serve a trovare il giusto equilibrio tra affidabilità scientifica, fattibilità e aspetti etici.

Perché non basta arruolare "più pazienti possibile"?

Un errore frequente è pensare che aumentare il numero dei partecipanti renda automaticamente uno studio migliore. In realtà, studi troppo piccoli possono portare a risultati incerti, al contrario studi troppo grandi comportano spreco di risorse e possibile sovrastima della rilevanza pratica di differenze minime. Lo scopo non è avere il campione più grande possibile, ma quello adeguato alla domanda di ricerca.

Gli errori che uno studio può commettere

L'obiettivo di uno studio clinico è stabilire se una differenza osservata tra due gruppi rappresenti un effetto reale oppure sia semplicemente dovuta al caso. Tuttavia, nessuno studio è esente da errori e, nell'interpretazione dei risultati, possono verificarsi due situazioni.

Il primo è l'errore di tipo I, che consiste nel concludere che esiste una differenza quando in realtà non esiste (falso positivo). È come condannare un innocente in un processo.

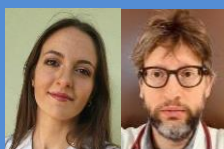
Il secondo è l'errore di tipo II, che consiste nel non rilevare una differenza che invece è realmente presente (falso negativo). È come assolvere un colpevole.

Entrambi gli errori possono avere importanti conseguenze nella ricerca clinica. Un falso positivo può portare all'introduzione di trattamenti inefficaci, mentre un falso negativo può impedire di riconoscere un intervento realmente utile. La progettazione dello studio, e in particolare il calcolo della dimensione del campione, ha proprio lo scopo di limitare il rischio di questi errori.

Come si controllano questi errori?

Per limitare il rischio di questi errori, prima dell'inizio dello studio è necessario definire alcuni parametri statistici che guideranno il calcolo della dimensione del campione:

- **Livello di significatività (α):** rappresenta la probabilità di commettere un errore di tipo I, cioè dichiarare significativa una differenza che in realtà non esiste. Nella maggior parte degli studi clinici α viene fissato a 0,05, corrispondente a un rischio del 5% di falso positivo.
- **Errore di tipo II (β):** rappresenta la probabilità di commettere un errore di tipo II, cioè non individuare una differenza realmente esistente. Ridurre β richiede generalmente un campione più numeroso.
- **Potenza statistica (Power):** corrisponde a $1 - \beta$ ed esprime la probabilità che lo studio riesca a individuare una differenza reale quando essa esiste. Nella pratica clinica viene generalmente scelta una potenza dell'80% o del 90%.



Parametro	Cosa rappresenta	Valore più utilizzato
α	Probabilità di falso positivo (errore di tipo I)	0,05
β	Probabilità di falso negativo (errore di tipo II)	0,20 - 0,10
Power (1- β)	Probabilità di rilevare una vera differenza	80% - 90%

Che cosa determina il sample size?

Il calcolo deriva da quattro elementi principali.

- **Effect size (d):** corrisponde all'effetto che vogliamo dimostrare, ovvero la differenza minima considerata clinicamente importante (es. riduzione del colesterolo LDL di 20 mg/dL, riduzione della pressione arteriosa di 5 mmHg, riduzione della glicata del 5%). Non si tratta della differenza che ci si aspetta necessariamente di osservare, ma della più piccola differenza che, se dimostrata, avrebbe un reale significato clinico. Più piccola è la differenza che si desidera individuare, più difficile sarà distinguerla dalla normale variabilità dei dati e maggiore dovrà quindi essere il numero di partecipanti. Corrisponde Δ (differenza clinicamente rilevante) / SD (deviazione standard attesa).
- **Variabilità dei dati:** i valori dei pazienti sono molto diversi tra loro (alta deviazione standard), distinguere un vero effetto dalla normale variabilità diventa più difficile; di conseguenza saranno necessari più partecipanti. Al contrario, quando i dati sono più omogenei, può essere sufficiente un campione più piccolo.
- **Livello di significatività (α):** visto in precedenza, più si vuole essere rigorosi (ad esempio $\alpha=0,01$), maggiore sarà il campione richiesto.
- **Potenza statistica (Power):** visto in precedenza, una maggiore potenza richiede un numero più elevato di partecipanti.

Come si calcola?

Nella pratica, non si esegue il calcolo manualmente: si inseriscono i parametri in un software dedicato, che restituisce il numero minimo di partecipanti necessario. Tra i software si annoverano G*Power e R (free) o PASS e nQuery (free trial).

Serve conoscere il tipo di studio, l'outcome principale, la differenza clinicamente rilevante, la variabilità attesa, il livello di significatività e la potenza desiderata. Non esiste un unico metodo di calcolo: la procedura varia in base al disegno dello studio e all'outcome considerato.

Parametro	Da dove si ricava?
Differenza clinicamente rilevante	Studi precedenti, linee guida, consenso di esperti o studio pilota
Variabilità attesa (SD)	Studi precedenti, studio pilota, database locali o revisioni sistematiche

Quando non si conosce la variabilità della popolazione o l'effetto atteso, può essere utile uno studio pilota.

Conclusioni

Il calcolo del sample size rappresenta uno dei passaggi fondamentali nella progettazione di uno studio clinico. Un campione adeguato permette di ottenere risultati affidabili, riducendo il rischio di falsi positivi e falsi negativi e garantendo che le risorse disponibili siano utilizzate in modo efficiente ed eticamente corretto.

Esempio pratico

Immaginiamo di voler verificare se un nuovo trattamento determini una riduzione del colesterolo LDL di almeno 20 mg/dL rispetto alla terapia standard. Tale valore viene definito a priori come la più piccola differenza considerata clinicamente rilevante. Sulla base dei dati disponibili in letteratura, si assume una deviazione standard di 40 mg/dL (doi: 10.1056/NEJMoa1912387). Fissando $\alpha = 0,05$ e una potenza dell'80%, sono necessari almeno 64 partecipanti per gruppo.

