

PILLOLE DI METODOLOGIA DELLA RICERCA

Quanti soggetti arruolare in un trial? (II)

Come ridurre il campione stimato se le risorse non sono sufficienti

Nel numero precedente abbiamo esaminato le componenti necessarie per stimare la dimensione del campione di un trial controllato randomizzato (RCT). Frequentemente, rispetto all'ipotesi iniziale, il numero di pazienti stimato è troppo elevato in relazione alle risorse disponibili. Al fine di avviare la sperimentazione clinica con un numero inferiore di pazienti, senza sottodimensionare lo studio, i ricercatori possono utilizzare diverse strategie:

- Arruolare una popolazione con un *Control Event Rate* (CER) più elevato.
- Espandere l'outcome primario.
- Estendere il periodo di follow-up.
- Prevedere una fase di run-in.

1. Arruolare una popolazione con un CER più elevato

A parità di efficacia dell'intervento terapeutico, una popolazione con un CER più elevato richiede un campione meno numeroso. In altre parole, arruolare pazienti "più gravi", nei quali la frequenza attesa dell'outcome primario è più elevata, permette di ridurre la dimensione del campione. Ad esempio, ipotizzando che la terapia sperimentale riduca del 25% l'incidenza dell'outcome primario, nel trial A (CER 20%) è necessario arruolare 2000 soggetti, mentre nel trial B (CER 40%) ne sono sufficienti la metà.

	Trial A		Trial B	
	Trattati	Controlli	Trattati	Controlli
N° di soggetti	1000	1000	500	500
N° di eventi	150	200	150	200
CER	15%	20%	30%	40%

Per tale ragione, in qualunque condizione clinica, nei trial di prevenzione secondaria (CER medio-alto) la dimensione del campione richiesto è nettamente inferiore rispetto a quelli di prevenzione primaria (CER basso). Ovviamente i vantaggi di ridurre la dimensione del campione arruolando una popolazione a rischio più elevato, sono controbilanciati dal tempo e dalle difficoltà richiesti per identificare soggetti ad alto rischio, specialmente quando sono necessari test diagnostici.

2. Espandere l'outcome primario

Costituisce una variante della strategia precedente: piuttosto che arruolare una popolazione con un CER più elevato viene scelto un outcome dalla frequenza attesa più elevata, ottenendo un analogo risultato: il CER aumenta e si riduce la dimensione del campione.

Per espandere l'outcome primario esistono due strategie. La prima utilizza un solo outcome dalla frequenza attesa più elevata: ad esempio, per dimostrare la riduzione della mortalità totale a 5 anni con le statine in prevenzione secondaria sono necessari circa 12.000 pazienti; se invece viene scelto come outcome primario la mortalità coronarica, la stima è di circa 8.000 pazienti. Analogamente, per dimostrare l'efficacia della profilassi della gastropatia da FANS il campione stimato si riduce se invece dell'ulcera peptica complicata (emorragia, perforazione), viene scelto come outcome primario l'ulcera sintomatica, molto più frequente o addirittura l'ulcera endoscopica, presente in un numero elevatissimo di soggetti che assumono FANS. Questo esempio dimostra che, generalmente, la rilevanza clinica dell'outcome è inversamente correlata alla sua frequenza.

La seconda metodologia prevede l'uso dei cosiddetti end-point (EP) combinati, al fine di aumentare l'efficienza statistica: in questi trial l'outcome primario è costituito da outcome diversi che, sommati, permettono di aumentare il CER. Riprendendo l'esempio delle statine, un trial che definisce come outcome primario la combinazione di morte coronarica e infarto non fatale richiede circa 4.000 pazienti che si riducono ulteriormente se l'outcome include tutti gli eventi vascolari (morti cardiovascolari, infarti e stroke non fatali, procedure di rivascolarizzazione). Ovviamente, l'uso esasperato degli EP combinati espone al rischio di includere nella "sommatoria" eventi di differente rilevanza clinica - una procedura di rivascolarizzazione non ha lo stesso "peso" di una morte cardiovascolare - sino addirittura ad EP surrogati. Ad esempio, nei trial che hanno documentato la presunta superiorità del fondaparinux rispetto all'enoaparina per la prevenzione degli eventi tromboembolici in chirurgia ortopedica maggiore, l'EP primario dello studio trombosi venosa profonda (TVP) combinava TVP sintomatica, TVP venografica prossimale e TVP venografica distale. Effettuando un'analisi per singoli EP il nuovo trattamento era efficace solo nel ridurre le TVP venografiche distali, un EP di nessuna rilevanza clinica che, tuttavia, garantiva un "effetto traino" sugli altri outcome. Pertanto, l'uso degli EP combinati richiede una certa cautela per almeno tre ragioni: la rilevanza clinica dei singoli EP che costituiscono l'EP combinato può essere molto diversa; l'efficacia del trattamento sull'EP combinato non può essere estesa ai singoli EP; gli EP più severi, generalmente sono meno frequenti e forn-

scono un “minore contributo” all’efficacia del trattamento, influenzato soprattutto dagli EP più frequenti, quasi sempre meno severi.

3. Estendere la durata del follow-up

Nella popolazione arruolata, la frequenza attesa dell’outcome primario potrebbe essere inferiore a quella prevista. Nel caso in cui questo sia conseguente a un follow-up troppo breve, per ridurre il rischio di risultati falsamente negativi è utile definire nel protocollo le *stopping rules* (criteri di interruzione del trial) in relazione al numero di eventi osservati, piuttosto che secondo una durata predefinita del trial, per consentire di estendere la durata del follow-up in relazione alla frequenza dell’outcome primario progressivamente rilevata.

4. Prevedere una fase di run-in

Nella fase di run-in a tutti i soggetti arruolati viene somministrato prima della randomizzazione, per un periodo definito, trattamento attivo o placebo, di cui solo i ricercatori sono a conoscenza. Considerato che la *non-compliance* influenza la stima della dimensione del campione, la fase di run-in viene utilizzata nei trial con follow-up di durata medio-lunga dove è maggiore la probabilità che i pazienti sospendano il trattamento (*withdrawals*) e/o abbandonino lo studio (*lost to follow-up*). Pertanto, la fase di run-in previene la potenziale *non-compliance* dei pazienti e permette di arruolare un campione meno numeroso. Questo obiettivo si raggiunge solitamente utilizzando il trattamento attivo che identifica e permette di escludere i pazienti intolleranti e/o con effetti avversi. Il run-in con il placebo, invece, permette di identificare sia falsi effetti avversi, sia le risposte terapeutiche al placebo: ovviamente, l’esclusione dal trial di questi soggetti sovrastima l’efficacia del trattamento.

In definitiva, la fase di run-in permette di ridurre il numero di soggetti da arruolare; tuttavia un campione della popolazione a elevata probabilità di *compliance* e/o a bassa prevalenza di risposta al placebo potrebbe essere estremamente “artificioso” e ridurre, sino a compromettere, l’applicabilità dei risultati dello studio.

5. Un case study: il FIELD

Il protocollo originale del *Fenofibrate Intervention and Event Lowering in Diabetes* (FIELD) - RCT di confronto tra fenofibrato e placebo in pazienti con diabete mellito di tipo 2 - definiva come outcome primario la mortalità coronarica a 5 anni, stimando un CER annuale di 0.8%. Poiché la frequenza di tale outcome risultò nettamente inferiore (0.31%), i ricercatori modificarono l’outcome primario in outcome combinato (mortalità coronarica più infarto non fatale) e aumentarono il campione da 8.000 a oltre 9.000 pazienti al fine di preservare la potenza statistica. Curiosamente, il fenofibrato non risultò efficace sull’outcome combinato, ma solo su una delle sue componenti: l’infarto non fatale.

KEY POINTS

CHECKLIST PER RIDURRE IL SAMPLE SIZE

- **Identificare nella popolazione di interesse un campione a rischio più elevato**
- **Assicurarsi che il profilo di rischio dei soggetti venga monitorato “in cieco” sia durante il reclutamento, sia durante il follow-up**
- **Se il CER dell’outcome primario è molto basso, identificare outcome alternativi più frequenti o utilizzare un end-point combinato**
- **Evitare di combinare end-point surrogati con end-point di provata rilevanza clinica**
- **In una eventuale fase di run-in, definire criteri rigorosi per randomizzare i soggetti con una compliance potenzialmente bassa**
- **“Aggiustare” il campione stimato in relazione al livello atteso di non-compliance**
- **Definire le stopping rules in relazione al numero di eventi osservati e non secondo una durata predefinita del trial**

Per saperne di più

Lecture introduttive

- Keech AC, GebSKI V. Managing the resource demands of a large sample size in clinical trials: can you succeed with fewer subjects? *Med J Aust* 2002;177:445-7.
- Freemantle N, Calvert M, Wood J, et al. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA* 2003;289:2554-9.
- Pablos-Méndez A, Barr RG, Shea S. Run-in periods in randomized trials: implications for the application of results in clinical practice. *JAMA* 1998;279:222-5.

Approfondimenti

- Lim E, Brown A, Helmy A, et al. Composite outcomes in cardiovascular research: a survey of randomized trials. *Ann Intern Med* 2008;149:612-7.
- Ulmer M, Robinaugh D, Friedberg JP, et al. Usefulness of a run-in period to reduce drop-outs in a randomized controlled trial of a behavioral intervention. *Contemp Clin Trials* 2008;29:705-10.
- Lee S, Walker JR, Jakul L, et al. Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? A meta-analytic evaluation. *Depress Anxiety* 2004;19:10-9.
- Keech A, Simes RJ, Barter P, et al. Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study): randomised controlled trial. *Lancet* 2005;366:1849-61.

Corso avanzato

- **GIMBE®**. Metodologia della ricerca clinica. Bologna, ottobre-dicembre 2009.