

Opinione

Trucchi, artifici, espedienti e stratagemmi nella conduzione e interpretazione delle sperimentazioni cliniche in cardiologia

Simone Frea, Marco Bobbio

S.C.U. Cardiologia I, A.O. San Giovanni Battista, Torino

(G Ital Cardiol 2006; 7 (11): 761-767)

© 2006 CEPI Srl

Le opinioni espresse in questo articolo non riflettono necessariamente quelle dell'Editor del Giornale Italiano di Cardiologia.

Ricevuto il 12 aprile 2006; nuova stesura il 26 luglio 2006; accettato il 28 luglio 2006.

Per la corrispondenza:

Dr. Marco Bobbio

S.C. di Cardiologia
A.O. Santa Croce e Carle
Via M. Coppino, 26
12100 Cuneo
E-mail: bobbio.m@ospedale.cuneo.it

Da una ventina di anni assistiamo alla pubblicazione di un numero crescente di sperimentazioni cliniche. Con pochi pazienti o megatrial, ben condotti o raffazzonati, ripetitivi o innovativi, condotti per poche settimane o per lunghi anni. Noi cardiologi abbiamo imparato molto: curiamo meglio l'infarto acuto, lo scompenso cardiaco, l'ipertensione. Ma a leggere queste sperimentazioni, abbiamo anche imparato a riconoscere trucchi, artifici, espedienti, stratagemmi adottati per "far venire" i risultati. Nulla di illecito, per carità. Solo qualche peccatuccio, che però in qualche modo modifica il messaggio che si dovrebbe trarre dai risultati della ricerca. Senza l'ambizione di essere completi e sistematici, vediamo alcuni.

L'analisi per sottogruppi non sempre dice il vero

L'analisi per sottogruppi nasce dal desiderio dei ricercatori di scoprire, all'interno della mole di dati raccolti, informazioni "nascoste" da cui trarre nuove ipotesi di lavoro. Molto spesso questo lavoro viene svolto *a posteriori*, ovvero una volta che lo studio si è concluso, illuminati dai risultati ottenuti dall'iniziale analisi dei dati. In pratica si mettono a confronto numerose variabili nella speranza di trovare relazioni statistiche significative e imprevedute. Queste relazioni possono trarre in inganno il lettore perché talvolta risulta difficile discriminare quando un'associazione sia dettata dal caso o da una relazione clinicamente rilevante. Recentemente, nell'analisi

si per sottogruppi dello studio Val-HeFT¹, risultava che i pazienti in trattamento con inibitori dell'enzima di conversione dell'angiotensina (ACE-inibitori) e betabloccanti ai quali veniva aggiunto un antagonista dell'angiotensina II avessero un'incidenza significativamente maggiore di eventi cardiaci rispetto a coloro a cui tale antagonista non era stato aggiunto. Questo dato, che ha frenato per alcuni anni la prescrizione della triplice associazione, è stato successivamente smentito dai risultati dello studio CHARM².

Quando l'analisi per sottogruppi rispecchia in maniera omogenea il risultato ottenuto sull'intera popolazione studiata viene rafforzato il messaggio, ma non viene fornita alcuna informazione aggiuntiva alla ricerca. Questa situazione può essere esemplificata dallo studio COPERNICUS³, nel quale è stata valutata l'efficacia del carvedilolo nei pazienti con scompenso cardiaco severo (frazione di eiezione [FE] < 25%); il rischio di morte e scompenso risultava ridotto nell'insieme dei pazienti dello studio e anche in modo omogeneo in tutti i sottogruppi esaminati (sesso, età, etnia, FE, eziologia, recenti ospedalizzazioni).

Quando l'analisi per sottogruppi evidenzia una maggiore efficacia di una terapia in un determinato sottogruppo o, viceversa, un risultato opposto a quello che si verifica nella popolazione globale, l'analisi per sottogruppi può diventare fonte di errore, ma anche fonte di conoscenza. Nel MADIT II⁴, nel quale è stata valutata l'efficacia dell'impianto di un defibrillatore rispetto alla terapia convenzionale in un

gruppo di pazienti con recente infarto miocardico e FE < 30%, è stata considerata una serie di sottogruppi suddivisi per età, sesso, FE, classe funzionale NYHA, larghezza del QRS (Figura 1). In base all'analisi per sottogruppi il defibrillatore risulterebbe utile nei pazienti < 60 o > 70 anni; applicando acriticamente il risultato di questa analisi, si verrebbe a creare una situazione a dir poco paradossale, escludendo dall'impianto i sessantenni. Il defibrillatore, inoltre, risulta progressivamente più utile man mano che si allunga la durata del QRS: potrebbe trattarsi di un risultato dovuto al caso oppure, dal momento che si profila un trend progressivo, di una relazione clinicamente rilevante (maggiore lunghezza del QRS → peggiore funzione ventricolare sinistra → maggior rischio di aritmie → maggiore utilità del defibrillatore).

Quando, al contrario, una sperimentazione clinica giunge a un risultato neutro, la dimostrazione di efficacia solo in un sottogruppo di pazienti non può essere accettata. Nello studio PRAISE I⁵, per esempio, l'amlodipina non si è dimostrata migliore del placebo nell'intero gruppo di pazienti con scompenso cardiaco avanzato, mostrando una riduzione significativa della mortalità soltanto nel sottogruppo di pazienti con cardiomiopatia su base non ischemica (28 vs 37%; $p < 0.05$). Basandosi su questi risultati positivi è stato condotto un secondo studio solo su pazienti non ischemici (PRAISE II, 49th Scientific Session of the American College of Cardiology, Anaheim, CA, USA, 2000, Packer M., dati non pubblicati) che non ha confermato gli effetti favorevoli dell'amlodipina su tale gruppo di pazienti (33.7 vs 31.7%; $p = NS$).

Consiglio: diffidare dei risultati "trovati" da un'analisi *a posteriori* dei dati e di quelli che risultano significativi solo per sottogruppi di pazienti.

Gli endpoint surrogati talvolta portano fuori strada

Medici e pazienti si aspettano da un nuovo trattamento la riduzione della morte o dell'incidenza di un evento grave. Lo studio di questi endpoint veri risulta difficile, perché richiede di solito una popolazione di studio assai numerosa e un follow-up molto lungo. Pertanto, nelle sperimentazioni cliniche, vengono spesso utilizzati endpoint surrogati, alla luce dei fondamenti fisiopatologici disponibili in letteratura. Invece di studiare l'incidenza di infarto miocardico, si può studiare la riduzione del colesterolo, dal momento che livelli elevati di colesterolo si associano a un maggior rischio di infarto, oppure, al posto di valutare la riduzione della necessità di dialisi, si può esaminare l'aumento della creatinemia.

Gli endpoint surrogati si fondano, pertanto, su un sillogismo fisiopatologico. Facciamo alcuni esempi. Tra i pazienti che hanno avuto un infarto, coloro che presentano aritmie ventricolari hanno un rischio maggiore di morte. La flecainide è in grado di ridurre le aritmie ventricolari nei pazienti con infarto miocardico; si potrebbe dedurre che essa riduca il rischio di morte nei pazienti con infarto miocardico e aritmie ventricolari. Lo studio CAST⁶, nel quale è stata valutata l'efficacia della flecainide nei pazienti con disfunzione ventricolare sinistra postinfartuale e aritmie ventricolari asintomatiche, ha confermato l'efficacia del farmaco nel ridurre le aritmie ventricolari, rivelando però un eccesso significativo di mortalità a 10 mesi nei pazienti trattati con il farmaco antiaritmico (4.5 vs 1.2%; rischio relativo [RR] 3.6; intervallo di confidenza [IC] 95% 1.7-8.5; $p < 0.01$). Una bella e istruttiva valutazione del significato degli endpoint surrogati si trova in una completa rassegna di qualche anno fa⁷.

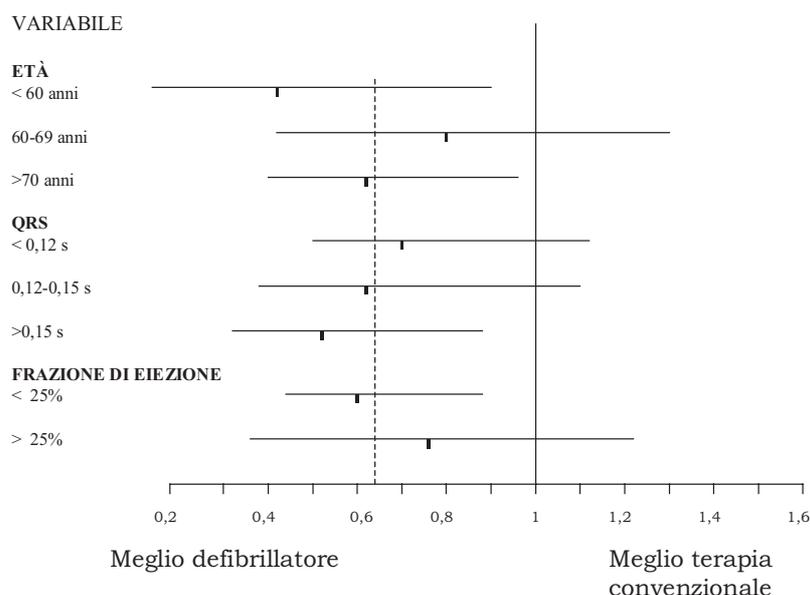


Figura 1. Analisi per alcuni sottogruppi di pazienti arruolati nello studio MADIT II.

Consiglio: fare attenzione che un farmaco, in grado di ridurre un endpoint surrogato, non sempre si rivela utile nella pratica clinica; anzi, potrebbe addirittura risultare dannoso.

L'endpoint combinato può trarre in inganno

L'efficacia di un trattamento può essere valutato, invece che su un singolo evento (il decesso, l'incidenza di infarto o di un intervento di rivascolarizzazione), su un insieme di eventi (endpoint combinato), in tal modo aumenta la probabilità che si osservi un evento e si può ridurre la numerosità del campione, la durata dell'osservazione e ottenere una misura complessiva di tutti gli eventi avversi che è possibile evitare con un determinato trattamento. Negli studi sulla cardiopatia ischemica vengono spesso adottati due gruppi di endpoint combinati: gli eventi cardiaci avversi maggiori che comprendono morte, infarto cardiaco, intervento di rivascolarizzazione (chirurgica o percutanea) e gli eventi cardiaci avversi maggiori e cerebrovascolari che aggiungono anche gli eventi cerebrovascolari agli endpoint degli eventi cardiaci. Talvolta, però, l'endpoint viene combinato mettendo insieme endpoint rilevanti sul piano clinico con altri più deboli (surrogati o intermedi) che si verificano più frequentemente. In tal modo si induce nel lettore affrettato l'impressione che l'efficacia riguardi gli endpoint importanti, mentre il risultato è spesso a carico solo di quelli meno importanti. Nello studio AVERT⁸, per esempio, è stata confrontata l'efficacia dell'atorvastatina rispetto all'angioplastica nei pazienti con cardiopatia ischemica cronica; i pazienti trattati con atorvastatina hanno avuto una minore incidenza di eventi ischemici (endpoint combinato di morte, ictus, infarto, necessità di rivascolarizzazione e aggravamento dell'angina) rispetto al gruppo trattato con angioplastica (13 vs 21%; $p = 0.048$). Valutando, però, l'incidenza dei singoli eventi, si realizza che non sono state osservate differenze fra i due gruppi per quanto riguarda gli endpoint maggiori (morte, infarto, rivascolarizzazione), mentre gran parte del beneficio si è osservato per il sintomo anginoso (Tabella 1). Si tenga

Tabella 1. Numero di eventi osservati nell'AVERT Study nel gruppo di pazienti trattati con atorvastatina e quelli trattati con angioplastica.

	Atorvastatina	Angioplastica
Endpoint combinato (morte, infarto miocardico acuto, angioplastica coronarica, bypass aortocoronarico, peggioramento dell'angina)	22	37
Morte	1	1
Infarto miocardico acuto	4	4
Angioplastica coronarica	18	21
Bypass aortocoronarico	2	9
Peggioramento dell'angina	11	25

conto che la ricerca non era stata condotta in doppio cieco; pazienti e ricercatori sapevano quale trattamento era stato assegnato e quindi la valutazione dell'aggravamento dell'angina non può essere considerata oggettiva, perché influenzata dalla conoscenza del trattamento assunto.

Consiglio: non fidarsi di un trattamento che riduce l'incidenza dell'endpoint combinato; si vada a vedere da quali endpoint è composto e quale è stata l'efficacia negli endpoint reali.

L'endpoint primario è difficile da trovare

Talvolta la scelta di un singolo endpoint forte, come la morte, per esempio, può sembrare l'opzione migliore, più semplice e quella capace di dare il maggior "peso" ai risultati di uno studio. Questo è vero. È anche vero, però, che per portare a termine uno studio, l'evento si deve verificare in un numero sufficiente di pazienti e possibilmente entro tempi non "storici". Al contrario, la scelta di endpoint combinati, come già visto, permette di osservare un maggior numero di eventi, con la necessità di un follow-up più breve, ma con il rischio di indebolire se non "annacquare" i risultati.

Nello studio CAPRICORN⁹ è stata valutata l'efficacia a lungo termine del carvedilolo nei pazienti con FE < 40% conseguente a infarto miocardico recente, trattati secondo le norme correnti dell'*evidence-based medicine* (trombolisi o angioplastica primaria e ACE-inibitori da almeno 48 h). L'endpoint primario originale era la mortalità per ogni causa. Le analisi *ad interim*, registrando una mortalità più bassa del previsto (popolazione a basso rischio soprattutto alla luce delle nuove terapie), hanno però indotto a correggere in corsa il protocollo, definendo come un endpoint primario la combinazione di mortalità totale e ricoveri per cause cardiovascolari. Morale: non si è evidenziata nessuna differenza tra i due gruppi di trattamento in termini di endpoint primario combinato (carvedilolo 35% vs placebo 37%; hazard ratio 0.92; $p = 0.3$), mentre l'endpoint singolo originario di mortalità per tutte le cause ha mostrato una significativa riduzione dell'evento nel gruppo trattato con carvedilolo (12 vs 15%; hazard ratio 0.77; $p = 0.031$).

Consiglio: si traggano le conclusioni sulla base del risultato ottenuto sull'endpoint primario, ma si vada anche a vedere cosa è successo negli endpoint secondari importanti.

Il modo in cui si presentano i dati conta (e molto)

Come ci facciamo influenzare dalla pubblicità così capita di essere inconsciamente "manipolati" dal modo in cui vengono presentati i risultati di un medesimo stu-

dio. Ragioniamo sui numeri. Nell'Helsinki Heart Study¹⁰ è stata testata l'efficacia del gemfibrozil vs placebo, nella prevenzione primaria di eventi di origine cardiaca (infarto miocardico e morte cardiaca) randomizzando circa 4000 uomini dislipidemici fra 40 e 55 anni. In termini assoluti, nel gruppo trattato 56 persone hanno avuto un evento contro 84 nel gruppo placebo. Questo risultato può essere descritto in modi diversi. Si può dire che l'incidenza di eventi nel gruppo trattato è stata del 2.73% contro il 4.14% nel gruppo placebo e che pertanto il rischio assoluto di avere un evento cardiaco è stato ridotto dal farmaco dell'1.41%; oppure si può dire che la riduzione relativa del rischio nel gruppo che ha assunto il gemfibrozil è stata del 34% (4.14-2.73/4.14); oppure ancora che bisogna trattare con gemfibrozil 71 uomini per 5 anni per evitare che un paziente abbia un evento cardiaco [$1/(4.14-2.73) \times 100$]. È tutto vero, ma se si presentano i risultati di uno studio in un modo o nell'altro, la percentuale di medici che somministreranno il farmaco ai propri pazienti potrebbe cambiare. A riguardo è stato condotto uno studio¹¹ che ha confermato questo sospetto. Se si presentano i risultati dello studio sul gemfibrozil in termini di riduzione assoluta del rischio (-1.41%) solo il 25% circa dei medici intervistati si dichiara favorevole alla terapia, mentre se si ricorre alla riduzione del RR (-34%) il favore dei medici intervistati triplica (75%).

Consiglio: occhio ai numeri. Se guardati velocemente i risultati possono ingannare. È sempre opportuno calcolarsi la differenza assoluta tra i risultati, per valutarne l'importanza.

Il conflitto di interessi condiziona gli autori

Un caso che ha interessato la comunità cardiologica è quello che riguarda la presunta pericolosità dei calcioantagonisti. Nel 1995 furono riportati i dati di una ricerca di tipo caso-controllo¹² su un gruppo di pazienti ipertesi, nella quale gli autori, dopo aver preso in considerazione numerose variabili, verificarono che il RR di infarto era 1.62 (ovvero il 62% in più) tra i pazienti trattati con nifedipina e che aumentava con l'aumentare della dose del farmaco. Poco dopo fu pubblicata una metanalisi¹³, nella quale vennero presi in considerazione 16 trial clinici randomizzati riguardanti la nifedipina. Il rischio cumulativo di mortalità saliva da 1.06 a 1.18 a 2.83 per dosi rispettivamente di 30-50, 60 e 80 mg/die. Nei mesi successivi vennero pubblicati numerosi articoli di commento alla vicenda.

Quattro ricercatori canadesi¹⁴ provarono a verificare se esisteva un rapporto tra la posizione assunta da coloro che erano intervenuti pubblicamente nel dibattito sulla sicurezza dei calcioantagonisti e i loro legami finanziari con l'industria farmaceutica. Presero in considerazione tutti gli articoli pubblicati tra il marzo 1995 e il settembre 1996 sull'argomento e ne individuarono 70.

Di questi, 30 vennero considerati a favore del farmaco, 17 neutrali e 23 critici. Chiesero agli autori degli articoli di esplicitare i loro rapporti economici con l'industria farmaceutica. Coloro che avevano dato un giudizio favorevole nei confronti della nifedipina avevano maggiori legami con le industrie che producono la nifedipina rispetto a coloro che avevano espresso pareri neutrali o contrari (96 vs 60 e 37%; $p < 0.001$). In ogni caso i favorevoli avevano maggiori legami economici con industrie farmaceutiche, indipendentemente dai farmaci prodotti (100 vs 67 e 43%; $p < 0.001$). Gli autori sostengono, quindi, che in quella circostanza si sia verificata una forte associazione tra la posizione assunta dagli autori degli articoli a favore della sicurezza dei calcioantagonisti e i loro legami con l'industria farmaceutica.

Consiglio: prima di prendere per buono il giudizio di un autore, valutare quali legami ha con l'industria che produce il farmaco, il test o il dispositivo di cui parla. Purché abbia dichiarato pubblicamente gli eventuali conflitti di interesse.

È meglio diffondere subito i risultati positivi e ritardare quelli negativi

Che fare quando uno studio non ottiene i risultati sperati? Si può tentare di non pubblicarli così da non intralciarne la commercializzazione, o ritardarne la pubblicazione. Nel numero del *New England Journal of Medicine* del 31 maggio 2001 troviamo pubblicati due articoli sull'uso dei betabloccanti nello scompenso cardiaco. Nello studio COPERNICUS, i pazienti con FE < 25% trattati con carvedilolo hanno avuto una mortalità nettamente inferiore rispetto ai pazienti trattati con placebo (11.4 vs 18.5%; $p = 0.0014$). Lo studio è stato pertanto pubblicato 14 mesi dopo la conclusione avvenuta il 20 marzo 2000. Nello studio BEST¹⁵, invece, i pazienti con FE < 35% trattati con bucindololo non hanno avuto una riduzione significativa della mortalità rispetto ai pazienti trattati con placebo (30.2 vs 33%; $p = 0.11$). Per la pubblicazione di questa ricerca, nella quale si dimostrava la non efficacia del bucindololo, sono trascorsi 22 mesi, essendo stata interrotta il 26 luglio 1999.

Consiglio: valutare quanti mesi sono stati impiegati per pubblicare i risultati negativi di uno studio. Può essere istruttivo verificare per quanto tempo la comunità scientifica è stata tenuta all'oscuro di un'informazione essenziale, ma scomoda.

Basta studiare tanti pazienti che la differenza tra i trattamenti risulta statisticamente significativa

Talvolta gli endpoint combinati non bastano, è anche necessario arruolare molti pazienti per ottenere dei ri-

sultati significativi. Nello studio CAPRIE¹⁶ è stata valutata l'efficacia del clopidogrel vs l'aspirina in 19 185 pazienti affetti da malattia cardiovascolare (ictus ischemico e infarto miocardico recenti, arteriopatia periferica sintomatica) nel ridurre l'incidenza di eventi cardiovascolari. Durante un follow-up di 2 anni circa, i pazienti trattati con clopidogrel hanno avuto una incidenza significativamente inferiore di eventi rispetto a quelli trattati con aspirina ($p = 0.043$). La riduzione del rischio assoluto di eventi è stata dello 0.5%, ovvero occorre trattare 196 pazienti per 2 anni per evitare un evento cardiovascolare.

Questo risultato, seppur statisticamente significativo, che peso ha nella pratica clinica? Saremmo disposti a prescrivere un farmaco sapendo *a priori* che l'entità del beneficio in termini assoluti è modesta? Sempre meglio di niente, se ci si ferma agli endpoint combinati. Per quanto riguarda l'infarto miocardico, l'incidenza è stata del 5.03% nei pazienti trattati con clopidogrel e del 4.84% in quelli trattati con aspirina, ovvero trattando 526 pazienti con clopidogrel si potrebbe procurare un infarto miocardico a un paziente (il dato però non è statisticamente significativo).

Consiglio: i grandi numeri rendono significativa una differenza che potrebbe nascondere un debole impatto nella pratica clinica. Oltre alla significatività statistica valutare anche la rilevanza clinica del risultato.

La riduzione della mortalità allunga di poco la sopravvivenza

Non sempre è chiaro (anzi di solito non viene detto) di quanto allungano la vita le terapie universalmente riconosciute come "salva-vita". Lo studio SOLVD¹⁷ ha rivoluzionato la terapia dello scompenso cardiaco, dimostrando l'efficacia di un ACE-inibitore, l'enalapril, nel ridurre la mortalità. In tale studio sono stati randomizzati a enalapril e placebo 2500 pazienti sintomatici per scompenso con $FE \leq 35\%$ ed è stata dimostrata una riduzione di mortalità del 16% nel gruppo trattato ($p = 0.0036$). Nello studio SOLVD Prevenzione¹⁸ sono stati arruolati pazienti con $FE \leq 35\%$ asintomatici e non è stata dimostrata una differenza significativa di mortalità a 1 anno (-8%, NS). Proseguendo il follow-up è stata però evidenziata una riduzione significativa della mortalità a 5 e a 12 anni dall'arruolamento nei pazienti trattati con enalapril (23 vs 27% e 53 vs 59%; $p = 0.001$). Ragionando in incremento medio dell'aspettativa di vita, la situazione appare meno rosea. Nei pazienti sintomatici per scompenso assumere tutti i giorni l'enalapril per oltre 10 anni allunga la vita media di 8.6 mesi (IC 95% 1-17.3 mesi; $p = 0.03$), mentre nei pazienti ancora asintomatici prolunga la vita media di 9.2 mesi (IC 95% 0-19.2 mesi; $p = 0.05$). Sarebbe interessante verificare il consenso al trattamento informando un paziente asintomatico che prendendo tutti i giorni 20

mg di enalapril per 12 anni potrebbe vivere mediamente 9 mesi in più. Potremmo anche aggiungere che, se fosse sfortunato, il trattamento potrebbe essere del tutto inutile e se fosse fortunato potrebbe allungargli la vita di 19 mesi.

Consiglio: quando si affronta una patologia cronica, non ci si deve accontentare di valutare l'efficacia di un farmaco in termini di riduzione della mortalità, ma anche rispetto alla modificazione di altri parametri quali il prolungamento della vita e il miglioramento della qualità della vita.

Si vince il confronto se si sceglie l'avversario

Se gareggiassi nella finale dei 100 m con un paziente in classe NYHA III vincerei e anche con un buon margine di vantaggio. Questo però non dimostra che io sia un buon centometrista e nemmeno un discreto sportivo. Una situazione simile la possiamo ritrovare in alcuni studi quando, per dimostrare la maggiore efficacia di un farmaco, viene scelto un concorrente debole. Non esistono norme e criteri per stabilire quale trattamento debba essere utilizzato come confronto (il più usato, il più economico, il più efficace, il più simile).

Per ogni scelta può essere trovata una giustificazione, anche se è opinabile e quindi discutibile. Per esempio nello studio di confronto di un nuovo inibitore dell'angiotensina II con un calcioantagonista nella protezione della nefropatia diabetica¹⁹, i pazienti trattati con irbesartan hanno avuto un'incidenza significativamente inferiore di peggioramento della funzione renale rispetto ai pazienti trattati con amlodipina (raddoppio della creatinina 16.9 vs 25.4%; nefropatia terminale 14.2 vs 18.3%). D'altro canto i pazienti del gruppo amlodipina hanno avuto un'incidenza maggiore di peggioramento della funzione renale rispetto al placebo (raddoppio creatinina 25.4 vs 23.7%; nefropatia terminale 18.3 vs 17.8%). È stata scelta l'amlodipina, pur sapendo che non ci sono prove di un suo effetto protettivo sulla nefropatia diabetica. Uno degli altri 6 antagonisti dell'angiotensina II o uno dei 14 ACE-inibitori in commercio, scelti come farmaci di confronto, non avrebbe potuto dimostrare l'efficacia dell'irbesartan?

Consiglio: quando viene dimostrato che un farmaco è più efficace di un altro, valutare se il confronto è stato equo ed è stato scelto il concorrente più idoneo.

Si vince il confronto se si indebolisce l'avversario

Per fare una figura migliore si potrebbe gareggiare nei 100 m contro un coetaneo di pari massa muscolare e pari allenamento, non prima però di avergli messo fuori uso una gamba. Nello studio US Carvedilol²⁰ è stato dimostrato che il carvedilolo alla dose di 25 mg \times 2/die

(dose media 45 ± 27 mg) riduce significativamente la mortalità nei pazienti in classe NYHA II-IV con FE < 35% (3.2 vs 7.8%; $p < 0.001$). Parallelamente nello studio MERIT-HF²¹ è stato dimostrato che il metoprololo succinato (a rilascio prolungato) alla dose di 200 mg/die (dose media 159 mg) riduce significativamente la mortalità nei pazienti in classe NYHA III-IV con FE $\leq 40\%$ (mortalità totale ridotta del 35%; $p < 0.0001$). I ricercatori dello studio COMET²² si sono proposti di confrontare l'efficacia in termini di mortalità di questi due betabloccanti, in un'ampia popolazione di pazienti con scompenso cardiaco ad eziologia ischemica e non ischemica.

Il risultato finale dimostra la maggior efficacia del carvedilolo: infatti la mortalità per tutte le cause è risultata del 34% (512 su 1511) nel gruppo carvedilolo e del 40% (600 su 1518) nel gruppo metoprololo (differenza statisticamente significativa: RR 0.83; IC 95% 0.74-0.93; $p = 0.0017$). Peccato che la dose target e media di carvedilolo (25 mg \times 2/die, media 42 ± 15 mg) siano state simili a quelle del US Carvedilol mentre per il metoprololo (di cui è stato usato il sale tartrato) la dose target è stata dimezzata (50 mg \times 2, media 85 ± 29 mg). Lo studio COMET può soltanto dimostrare che il carvedilolo a dosaggio massimale è più efficace del metoprololo a dosaggio dimezzato nel ridurre la mortalità dei pazienti con scompenso cardiaco.

Consiglio: prima di stabilire che un farmaco sia migliore di un altro, accertarsi che le dosi utilizzate nella ricerca di confronto siano quelle utilizzate nelle ricerche precedenti.

Non bisogna applicare pedestremente i risultati delle ricerche cliniche

Le popolazioni degli studi e quelle reali non sempre coincidono: i pazienti nella pratica clinica spesso sono più complessi, hanno un maggior numero di comorbidità e assumono un maggior numero di medicine che possono creare interferenze imprevedibili. Inoltre i controlli non sempre possono essere così stretti e meticolosi come nelle popolazioni seguite nel corso delle sperimentazioni cliniche. Perciò non sempre un trattamento che ha dimostrato di ridurre la mortalità in un grande studio si rivela universalmente utile nella pratica.

Nello studio RALES²³ è stato dimostrato che nello scompenso cardiaco di classe NYHA III-IV con FE $\leq 35\%$ l'aggiunta alla terapia standard di basse dosi di spironolattone (25 mg/die) riduce significativamente la mortalità (35 vs 46%; RR 0.70; IC 95% 0.60-0.82; $p < 0.001$), con una trascurabile incidenza di iperpotassiemia. Dal momento della pubblicazione di questo studio è stato però evidenziato un aumento significativo sia dei tassi di mortalità per iperpotassiemia (dallo 0.3 per 1000 nel 1994 al 2 per 1000 nel 2001; $p < 0.001$) fra

i pazienti ricoverati per scompenso cardiaco in terapia con ACE-inibitori sia dei tassi di riospedalizzazione per iperpotassiemia²⁴ (dal 2.4 per 1000 nel 1994 all'11 per 1000 nel 2001; $p < 0.001$). Nella pratica clinica sembrerebbe pertanto che gli effetti collaterali della terapia con spironolattone siano maggiori rispetto a quelli registrati nello studio, tanto da controbilanciare il dimostrato effetto positivo nel ridurre la mortalità. Probabilmente i risultati positivi ottenuti nel corso della sperimentazione hanno indotto in un primo tempo i clinici ad un uso estensivo del risparmiatore di potassio, senza tenere conto del dosaggio o dei limiti di creatinina e potassio utilizzati nello studio, né delle raccomandazioni riguardo al monitoraggio degli elettroliti plasmatici e alla sospensione del farmaco.

Consiglio: prima di introdurre con entusiasmo un nuovo trattamento nella propria pratica clinica, accertare quali pazienti erano stati arruolati, quali condizioni escludevano i pazienti dal trattamento e quali accorgimenti erano stati adottati per seguire i pazienti nel corso della ricerca.

Se le cose si mettono male si interrompe la ricerca

Dal settembre 1996 al dicembre 1998, nello studio CONVINCES²⁵, sono stati arruolati 16 608 pazienti ipertesi con uno o più fattori di rischio, randomizzati a verapamil (180 mg), atenololo (50 mg) o idroclorotiazide (12.5 mg). Lo sponsor ha chiuso, per ragioni commerciali, lo studio 2 anni prima di quanto pianificato. Al momento dell'interruzione non si erano evidenziate differenze significative fra i tre gruppi di trattamento, con un trend di riduzione della mortalità nei gruppi trattati con betabloccante o diuretico (eventi totali: verapamil 4.5% vs atenololo/idroclorotiazide 4.4%; decessi: verapamil 4.1% vs atenololo/idroclorotiazide 3.8%). Estrapolando i dati ottenuti fino al momento dell'interruzione per i successivi 2 anni, si sarebbe dimostrato che il verapamil è leggermente migliore dell'atenololo e il diuretico è migliore di entrambi²⁶. Quali ripercussioni ci sarebbero state, nella vendita del verapamil, se lo studio fosse stato condotto fino alla fine, come originariamente programmato?

Anche le ricerche interrotte precocemente per un "evidente" beneficio di un trattamento spesso mostrano un effetto non plausibile soprattutto quando il numero di eventi analizzato è ancora modesto^{26,27}.

Consiglio: diffidare dai risultati delle ricerche interrotte non per motivi etici (ovvero, quando si mostra che un trattamento è migliore di un altro), ma per motivi economici. C'è sempre il sospetto che la scelta sia stata dettata dalla prevedibile *defaillance* del trattamento prodotto dall'industria che ha sponsorizzato la ricerca.

Bibliografia

1. Cohn JN, Tognoni G, for the Valsartan Heart Failure Trial Investigators. A randomized trial of the angiotensin-receptor blocker valsartan in chronic heart failure. *N Engl J Med* 2001; 345: 1667-75.
2. McMurray JJ, Ostergren J, Swedberg K, et al, for the CHARM Investigators and Committees. Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function taking angiotensin-converting-enzyme inhibitors: the CHARM-Added trial. *Lancet* 2003; 362: 767-71.
3. Packer M, Coats AJ, Fowler MB, et al, for the Carvedilol Prospective Randomized Cumulative Survival Study Group. Effect of carvedilol on survival in severe chronic heart failure. *N Engl J Med* 2001; 344: 1651-8.
4. Moss AJ, Zareba JW, Hall WJ, et al, for the Multicenter Automatic Defibrillator Implantation Trial II Investigators. Prophylactic implantation of a defibrillator in patients with myocardial infarction and reduced ejection fraction. *N Engl J Med* 2002; 346: 877-83.
5. Packer M, O'Connor CM, Ghali JK, et al. Effect of amlodipine on morbidity and mortality in severe chronic heart failure. Prospective Randomized Amlodipine Survival Evaluation Study Group. *N Engl J Med* 1996; 335: 1107-14.
6. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. *N Engl J Med* 1989; 321: 406-12.
7. Fleming TR, DeMets DL. Surrogate end points in clinical trial: are we being misled? *Ann Intern Med* 1996; 125: 605-13.
8. Pitt B, Waters D, Brown WV, et al. Aggressive lipid-lowering therapy compared with angioplasty in stable coronary artery disease. Atorvastatin versus Revascularization Treatment Investigators. *N Engl J Med* 1999; 341: 70-6.
9. Dargie HJ. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: the CAPRICORN randomised trial. *Lancet* 2001; 357: 1385-90.
10. Huttunen JK, Manninen V, Manttari M, et al. The Helsinki Heart Study: central findings and clinical implications. *Ann Med* 1991; 23: 155-9.
11. Bobbio M, Demichelis B, Giustetto G. Completeness of reporting trial results: effect on physicians willingness to prescribe. *Lancet* 1994; 343: 1209-11.
12. Psaty BM, Heckbert SR, Koepsell TD, et al. The risk of myocardial infarction associated with antihypertensive drug therapies. *JAMA* 1995; 274: 620-5.
13. Furberg CD, Psaty BM, Meyer JV, et al. Nifedipine dose-related increase in mortality in patients with coronary heart disease. *Circulation* 1995; 92: 1326-31.
14. Stelfox HT, Chua G, O'Rourke K, Detsky AS. Conflict of interest in the debate over calcium-channel antagonists. *N Engl J Med* 1998; 338: 101-6.
15. The Beta-Blocker Evaluation of Survival Trial Investigators. A trial of the beta-blocker bucindolol in patients with advanced chronic heart failure. *N Engl J Med* 2001; 344: 1659-67.
16. The CAPRIE Steering Committee. A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischemic events (CAPRIE). *Lancet* 1996; 348: 1329-39.
17. The SOLVD Investigators. Effects of enalapril on survival in patients with reduced left ventricular ejection fraction and congestive heart failure. *N Engl J Med* 1991; 325: 293-302.
18. The SOLVD Investigators. Effects of enalapril on mortality and the development of heart failure in asymptomatic patients with reduced left ventricular ejection fractions. *N Engl J Med* 1992; 327: 685-91.
19. Lewis EJ, Hunsicker LG, Clarke WR, et al. Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. *N Engl J Med* 2001; 345: 851-60.
20. Packer M, Bristow MR, Cohn JN, et al, for the US Carvedilol Heart Failure Study Group. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. *N Engl J Med* 1996; 334: 1349-55.
21. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF). *Lancet* 1999; 353: 2001-7.
22. Poole-Wilson PA, Swedberg K, Cleland JG, et al. Comparison of carvedilol and metoprolol on clinical outcomes in patients with chronic heart failure in the Carvedilol or Metoprolol European Trial (COMET): randomised controlled trial. *Lancet* 2003; 362: 7-13.
23. Pitt B, Zannad F, Remme WJ, et al. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. Randomized Aldactone Evaluation Study Investigators. *N Engl J Med* 1999; 341: 709-17.
24. Juurlink DN, Mamdani MH, Lee DS, et al. Rates of hyperkalemia after publication of the Randomized Aldactone Evaluation Study. *N Engl J Med* 2004; 351: 543-51.
25. Black HR, Elliot WJ, Grandits G, et al, for the CONVINCENCE Research Group. Principal results of the Controlled Onset Verapamil Investigation of Cardiovascular End Points (CONVINCE) trial. *JAMA* 2003; 289: 2073-83.
26. Psaty BM, Rennie D. Stopping medical research to save money: a broken pact with researchers and patients. *JAMA* 2003; 289: 2128-31.
27. Pocock SJ. When (not) to stop a clinical trial for benefit. *JAMA* 2005; 294: 2228-30.