

Real world evidence: uno sguardo oltre le sperimentazioni cliniche controllate

Antonio Nicolucci

Center for Outcomes Research and Clinical Epidemiology, Pescara

LE SPERIMENTAZIONI CONTROLLATE RANDOMIZZATE

Nella valutazione comparativa di efficacia fra due o più farmaci o interventi, le sperimentazioni cliniche randomizzate (randomized clinical trials, RCT), quando possibile in doppio cieco con placebo, rappresentano il gold standard per la produzione di nuove evidenze. Questo approccio metodologico garantisce infatti la maggiore validità scientifica, grazie alla possibilità di minimizzare il rischio di errori sistematici (bias) che potrebbero rendere problematica l'interpretazione dei risultati ottenuti. Fra i tanti tipi di bias descritti, alcuni sono particolarmente rilevanti e pericolosi. Come primo punto, se lo sperimentatore fosse in grado di scegliere liberamente quali pazienti assegnare all'intervento sperimentale e quali al gruppo di controllo, si incorrerebbe nel bias di selezione. In altre parole, lo sperimentatore potrebbe, in maniera più o meno conscia, assegnare tutti i pazienti più giovani o più sani al nuovo o al vecchio trattamento in base alle aspettative di efficacia e tollerabilità. Si creerebbero quindi due gruppi di pazienti con caratteristiche differenti, rendendo difficile l'interpretazione del confronto di efficacia.

Una volta creati, sarà necessario seguire i gruppi di trattamento nel tempo per verificare a distanza i risultati ottenuti. Anche in questo caso, se i due gruppi venissero seguiti con modalità o frequenze diverse (performance bias), potrebbero verificarsi differenze di risultato dipendenti dal diverso livello di attenzione posto, piuttosto che da reali differenze di efficacia dei trattamenti messi a confronto. Un altro aspetto che può minare in modo sostanziale la validità scientifica di uno studio è rappresentato dall'exclusion bias, vale a dire la perdita di pazienti durante lo studio, con impossibilità di recuperarne i dati per l'analisi di efficacia/sicurezza, o l'esclusione di parte dei pazienti dall'analisi per via di violazioni al protocollo o di insufficiente aderenza alle terapie. Infatti, i pazienti persi durante lo studio potrebbero essere proprio quelli che hanno avuto l'evento di interesse, o che hanno risposto meno al trattamento, o che hanno avuto esperienza di tossicità o effetti collaterali. L'esclusione di questi pazienti dall'analisi potrebbe di conseguenza portare a conclusioni del tutto distorte. Infine, se chi è chiamato a valutare l'efficacia dei trattamenti messi a confronto fosse a conoscenza della terapia assegnata al singolo paziente, egli potrebbe, sulla base delle proprie aspettative, sovrastimare o sottostimare il beneficio attribuibile al trattamento sperimentale (detection bias). Questi bias possono essere minimizzati grazie al processo della randomizzazione (selection bias), dell'uso del doppio cieco e del placebo (performance bias, detection bias) e dell'analisi per intention to treat (exclusion bias). In particolare, la randomizzazione consente di ottenere gruppi di pazienti ben bilanciati per qualsiasi fattore di rischio noto e non noto, consentendo di attribuire alla diversa efficacia dei trattamenti messi a confronto qualsiasi differenza di risultato dovesse emergere dallo studio.

Tabella 1 ♦ **Punti di forza e limiti delle sperimentazioni cliniche randomizzate**

PUNTI DI FORZA	LIMITI
Popolazione in studio ben definita	Esclusione di molti pazienti che necessitano di trattamento
Il disegno massimizza la validità interna	Generalmente condotte in centri specializzati
Trattamento somministrato in condizioni strettamente controllate	Difficoltà a generalizzare i risultati a popolazioni più ampie e meno selezionate
Compliance massimizzata da uno stretto protocollo e da frequenti visite di follow-up	La durata e le dimensioni del campione generalmente limitate rendono difficile l'identificazione di eventi avversi rari o a lungo termine
	"trial effect"

Sebbene altamente validi dal punto di vista scientifico, gli RCT possono porre problemi di trasferibilità dei risultati alla normale pratica clinica. Gli RCT abitualmente forniscono evidenza di cosa può essere ottenuto in condizioni strettamente controllate, in gruppi di pazienti selezionati, trattati per un definito periodo di tempo. La tabella 1 riassume i principali vantaggi e svantaggi degli RCT. Gli svantaggi sono specificamente legati alla rappresentatività dei pazienti inclusi. Infatti, negli RCT sono spesso esclusi i pazienti più anziani e quelli con multimorbidità o, come capita nei più recenti trials di safety cardiovascolare (CVOT), i pazienti a rischio cardiovascolare medio-basso. Inoltre, molte sperimentazioni vengono condotte in centri altamente specializzati e in condizioni strettamente controllate, con una frequenza di visite molto superiore alla normale pratica clinica e con particolari attenzioni a mantenere una elevata compliance. Tutto questo determina risultati positivi, anche a prescindere dal tipo di trattamento in studio. La presenza di un "trial effect" è facilmente desumibile dal miglioramento, ad esempio, dei livelli di HbA_{1c} anche nei pazienti assegnati al placebo. Va infine sottolineato come gli RCT abbiano generalmente una durata limitata che non consente una valutazione degli effetti a lungo termine, e che le dimensioni dello studio sono sufficienti a garantire una potenza statistica adeguata per evidenziare eventuali differenze di efficacia sull'endpoint primario, ma non su possibili eventi avversi rari ma importanti. Per tutte queste ragioni, a fianco delle prove di efficacia in condizioni ideali (efficacy) è necessario produrre evidenze di efficacia e sicurezza in condizioni di normale pratica clinica (effectiveness).

LA REAL WORLD EVIDENCE

La necessità di affiancare ai risultati degli RCT dati che derivino dalla pratica assistenziale è sempre più largamente riconosciuta, non solo dal mondo clinico, ma anche dagli enti regolatori. Ad esempio, la Linea Guida AIFA (Determinazione 20 marzo 2008. Linea Guida per la classificazione e conduzione degli studi osservazionali sui farmaci – pubblicata il 31 marzo 2008 sulla Gazzetta Ufficiale) sottolinea come gli studi osservazionali sui farmaci siano di particolare importanza per la valutazione del profilo di sicurezza nelle normali condizioni di uso e su grandi numeri di pazienti, per approfondimenti sull'efficacia nella pratica clinica, per la verifica dell'appropriatezza prescrittiva e per valutazioni di tipo farmacoeconomico.

I dati del mondo reale (real world data) sono dati sullo stato di salute e sull'assistenza raccolti routinariamente da una varietà di possibili fonti, quali cartelle cliniche informatizzate, dati amministrativi, dati di prescrizione, registri di patologia, o dati generati dai pazienti stessi. Questi dati possono essere utilizzati nell'ambito di studi con disegno diverso, che spaziano dai clinical trials pragmatici agli studi osservazionali, retrospettivi o prospettici. Per real world evidence (RWE) si intende una evidenza di efficacia e sicurezza che derivi dall'analisi di real world data.

La RWE può essere prodotta, con finalità diverse, nei vari momenti dello sviluppo di un farmaco. In fase di pre-lancio, può essere utile disegnare studi che permettano di descrivere i pattern prescrittivi correnti e di valutare la storia naturale della patologia, con la finalità di capire quale potrebbe essere il posizionamento nel mercato del nuovo farmaco.

Subito dopo l'immissione in commercio, può essere importante verificare quale sia il livello di adozione del nuovo trattamento, quali siano le caratteristiche dei pazienti trattati e quale sia il profilo di efficacia e sicurezza a breve e lungo termine in condizioni di normale pratica clinica. In questa fase può anche essere documentata l'accettabilità del trattamento da parte dei pazienti, l'impatto sull'utilizzo di risorse e l'esistenza di pazienti non trattati nonostante le indicazioni.

Quando l'impiego del nuovo trattamento è ormai consolidato, la RWE può aiutare a capire se il farmaco è utilizzato in accordo con le linee guida, a descrivere l'appropriatezza prescrittiva, o a identificare sottogruppi di pazienti che abbiano una probabilità più elevata di beneficiare del trattamento, fornendo eventualmente elementi a supporto dell'estensione dei criteri di rimborsabilità.

Le finalità delle analisi di real world data sono riassunte in tabella 2.

Uno degli scopi più importanti si riferisce sicuramente alla possibilità di confermare il profilo di efficacia e sicurezza di un trattamento in popolazioni più ampie e meno selezionate rispetto a quelle incluse negli RCT.

Un esempio paradigmatico è rappresentato dallo studio CVD-REAL, che ha valutato su un'ampia, non selezionata popolazione di oltre 300.000 soggetti, l'efficacia degli SGLT2 inibitori (1). La tabella 3 mette a confronto le caratteristiche dello studio EMPAREG, il trial che ha dimostrato l'efficacia di empagliflozin vs placebo nel ridurre il rischio di eventi cardiovascolari maggiori (2), con le caratteristiche del CVD-REAL. Appare evidente come in condizioni di normale pratica clinica questa classe di farmaci sia stata utilizzata in una popolazione con un rischio cardiovascolare sostanzialmente più basso rispetto al trial randomizzato. Nonostante le marcate differenze nelle due popolazioni, viene confermato l'effetto significativo di riduzione del rischio di ospedalizzazione per scompenso cardiaco e di mortalità per tutte le cause. Inoltre, gli stessi benefici sono stati documentati in Paesi diversi, caratterizzati dall'uso di diverse molecole appartenenti alla classe degli SGLT2 inibitori, suggerendo l'esistenza di un effetto di classe (1).

Un altro vantaggio delle RWE, derivante dallo studio di grandi numeri di soggetti, consiste nella possibilità di identificare sottogruppi di pazienti che presentino una maggiore probabilità di beneficiare del trattamento, consentendone un uso più appropriato. Ad esempio lo studio REAL, condotto in Italia sui pazienti che iniziavano un trattamento con liraglutide, ha documentato come il beneficio maggiore in termini di riduzione dei livelli di HbA1c fosse riscontrabile nei soggetti con valori di HbA1c al baseline al di sopra di 9.0% (riduzione media di oltre il 2%) (3). Paradossalmente, a causa delle limitazioni di rimborsabilità imposte da AIFA, a questi pazienti non può essere prescritto un agonista recettoriale del GLP-1.

Un altro ambito di grande interesse nell'utilizzo della RWE è rappresentato dalla valutazione della sicurezza di un trattamento, quando somministrato a pazienti non selezionati, spesso politrattati e quindi a rischio di sommazione di effetti o di interazioni fra farmaci. Le numerosità necessariamente limitate degli RCT si prestano male a valutare eventi avversi rari ma potenzialmente pericolosi. Inoltre, sbilanciamenti di eventi rari fra i gruppi in studio possono essere effetto del caso. Ad esempio, un eccesso di ricoveri per scompenso cardiaco è stato documentato nei pazienti trattati con sitagliptin nel trial Savor TIMI-53 (4). Tuttavia, una successiva analisi di dati italiani derivanti da database amministrativi ha permesso di evidenziare una riduzione del rischio di ospedalizzazione per scompenso cardiaco associato all'uso degli inibitori del DPP-4 rispetto alle sulfaniluree o al pioglitazone (5).

Analogamente, dall'analisi dei dati relativi ad oltre 180.000 persone con diabete di tipo 2 trattate con basse dosi di aspirina è emerso come, contrariamente alle persone senza diabete, la terapia antiaggregante non si associa ad un eccesso di rischio di sanguinamenti intracranici o gastrointestinali (6). Questi dati suggeriscono un minore effetto antiaggregante dell'aspirina nei soggetti affetti da diabete.

Come già discusso in precedenza, gli RCT tendono a massimizzare la compliance al trattamento. Di converso, nella pratica clinica una bassa aderenza può essere responsabile di una efficacia subottimale di una terapia. A parità di efficacia, un miglior profilo di accettabilità ed una maggiore persistenza in terapia potrebbero indirizzare la scelta del farmaco da prescrivere. Ad esempio, dall'analisi delle prescrizioni dei vari agonisti recettoriali del GLP-1 è emerso come la persistenza in terapia dopo 300 giorni oscillasse fra circa il 60% con dulaglutide a somministrazione settimanale e circa il 30% con lixisenatide o con exenatide in somministrazione duplice giornaliera (7).

Tabella 2 ♦ Finalità degli studi di real world evidence

VALUTARE:
Se i risultati ottenuti negli RCT sono replicabili in condizioni di normale pratica clinica
L'impatto di una terapia in un più ampio spettro di popolazioni e setting assistenziali
L'effetto del trattamento sul lungo periodo e su outcomes non considerati negli RCT
L'impatto della terapia in diversi sottogruppi di pazienti, molti dei quali potrebbero essere stati esclusi o sottorappresentati negli RCT
Gli eventi avversi rari
Il diverso profilo di tossicità/tollerabilità del farmaco in relazione ai diversi sottogruppi di pazienti
La persistenza in terapia e la compliance
Il consumo di risorse, i costi e il rapporto costo-efficacia
La resa di modelli assistenziali diversi
I bisogni assistenziali di popolazioni o di aree geografiche specifiche

I LIMITI DELLA RWE

A fianco degli innegabili vantaggi della RWE, è tuttavia necessario sottolinearne i limiti.

Innanzitutto, in assenza del processo di randomizzazione, il rischio di selection bias è reale, rendendo difficile la valutazione del nesso di causalità. Mentre il processo della randomizzazione garantisce l'assegnazione casuale di uno dei trattamenti messi a confronto ed offre sufficienti garanzie che i gruppi in studio siano ben bilanciati anche per caratteristiche non note, nel mondo reale la decisione di privilegiare un trattamento rispetto ad un altro non può prescindere da una complessa valutazione di caratteristiche cliniche, socio-culturali e comportamentali. Nella RWE, la possibilità di ottenere informazioni dettagliate è limitata dalla disponibilità dei dati, soprattutto in caso di analisi di database amministrativi o di studi retrospettivi. La possibilità di dati mancanti (sia sugli eventi che sui fattori di rischio) o di eterogeneità nella definizione/classificazione di patologie o di eventi è inoltre sempre presente. L'applicazione di tecniche di analisi sofisticate, come ad esempio la propensity score analysis, può ridurre il rischio di bias favorendo il bilanciamento fra i gruppi in studio per una serie più o meno ampia di fattori noti. Il propensity score di

Tabella 3

CARATTERISTICHE AL BASELINE	EMPAREG ²	CVD-REAL ¹
Età (anni)	63.1±8.6	56.9±10
Maschi (%)	71.2%	55.7%
Pregresso infarto del miocardio	46.7%	2.5%
Angina instabile	-	1.6%
Pregresso ictus	23.1%	4.1%
Vasculopatia periferica	21.0%	3.4%
Scompenso cardiaco	9.9%	3.1%
Malattia renale cronica	25.9%	2.5%

un individuo rappresenta la probabilità di ricevere il trattamento in oggetto data una serie di caratteristiche osservate. Tale probabilità viene stimata attraverso una regressione logistica nella quale la variabile dipendente è rappresentata dall'essere trattati o meno con uno specifico farmaco, mentre le covariate sono rappresentate da tutti i fattori noti che possono influenzare sia l'outcome che la prescrizione dello specifico trattamento. Pertanto, è possibile identificare per ciascun soggetto che riceve uno specifico trattamento un altro soggetto che, pur presentando una probabilità molto simile di ricevere quel trattamento (cioè analogo propensity score), è stato in realtà trattato con un altro farmaco (matching).

Tuttavia, raramente la ricchezza di informazioni è tale da consentire un matching su un numero molto vasto di caratteristiche. Ad esempio, informazioni sullo stato socio economico non sono abitualmente disponibili, sebbene rappresentino un importante fattore di rischio di outcomes sfavorevoli.

In conclusione, la RWE non si pone in antitesi o come sostituto degli RCT, ma rappresenta un utile complemento delle sperimentazioni cliniche controllate, consentendo di confermare su popolazioni più ampie e non selezionate e su periodi di tempo più lunghi i risultati dei trials.

In aggiunta, la RWE può fornire importanti informazioni sulla sicurezza e la tollerabilità a lungo termine dei farmaci e, grazie alla valutazione più estesa di rischi e benefici, può aiutare nell'identificazione di sottogruppi di pazienti con più elevata probabilità di beneficio. Infine, l'osservazione di dati dal mondo reale può rappresentare un importante strumento per la generazione di nuove ipotesi e fornire lo spunto per nuovi RCT. Tuttavia, a causa dell'impossibilità di escludere l'esistenza di fattore di confondimento, gli studi osservazionali dovrebbero essere utilizzati con grande cautela per valutazioni comparative di efficacia. Da questo punto di vista, i più recenti sviluppi metodologico-statistici hanno offerto un importante contributo alla riduzione dei rischi di bias, senza tuttavia consentire la loro completa eliminazione. Cionondimeno, il rigore nella conduzione di studi di RWE e la coerenza di risultati fra ricerca osservazionale e sperimentale permettono di guardare con fiducia ai risultati ottenuti e di utilizzare in modo integrato tutte le fonti di evidenza per guidare il processo decisionale medico e regolatorio.

Quale ruolo possa e debba avere la real world evidence nel supportare sia gli aspetti regolatori che le linee guida per la pratica clinica è ancora oggetto di dibattito nei vari Paesi. Nella "piramide delle evidenze", gli studi di coorte di qualità metodologica adeguata vengono subito dopo le sperimentazioni cliniche controllate randomizzate. Tuttavia, soprattutto se le sperimentazioni sono condotte su numeri limitati di pazienti o hanno breve durata, le informazioni aggiuntive derivanti dai dati del mondo reale riguardo l'efficacia, la sicurezza e il rapporto costo efficacia possono essere di grande utilità per la valutazione complessiva di una tecnologia. Un approccio più uniforme e standardizzato nel disegno e la conduzione di studi di real world sarà di grande importanza per accrescere la credibilità e la fiducia nella ricerca osservazionale.

BIBLIOGRAFIA

1. Kosiborod M, Cavender MA, Fu AZ, Wilding JP, Khunti K, Holl RW, Norhammar A, Birkeland KI, Jørgensen ME, Thureson M, Arya N, Bodegård J, Hammar N, Fenici P; CVD-REAL Investigators and Study Group. Lower Risk of Heart Failure and Death in Patients Initiated on Sodium-Glucose Cotransporter-2 Inhibitors Versus Other Glucose-Lowering Drugs: The CVD-REAL Study (Comparative Effectiveness of Cardiovascular Outcomes in New Users of Sodium-Glucose Cotransporter-2 Inhibitors). *Circulation* 136: 249-259, 2007.
2. Zinman B, Wanner C, Lachin JM, Fitchett D, Bluhmki E, Hantel S, Mattheus M, Devins T, Johansen OE, Woerle HJ, Broedl UC, Inzucchi SE; EMPA-REG OUTCOME Investigators. Empagliflozin, Cardiovascular Outcomes, and Mortality in Type 2 Diabetes. *N Engl J Med* 373: 2117-28, 2015.
3. Simioni N, Berra C, Boemi M, Bossi AC, Candido R, Di Cianni G, Frontoni S, Genovese S, Ponzani P, Provenzano V, Russo GT, Sciangula L, Lapolla A, Bette C, Rossi MC; ReaL (NN2211-4118) Study Group. Predictors of treatment response to liraglutide in type 2 diabetes in a real-world setting. *Acta Diabetol* 55: 557-568, 2018.
4. Scirica BM, Bhatt DL, Braunwald E, Steg PG, Davidson J, Hirshberg B, Ohman P, Frederick R, Wiviott SD, Hoffman EB, Cavender MA, Udell JA, Desai NR, Mosenzon O, McGuire DK, Ray KK, Leiter LA, Raz I; SAVOR-TIMI 53 Steering Commit-

- tee and Investigators. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus. *N Engl J Med* 369: 1317-26, 2013.
5. Fadini GP, Avogaro A, Degli Esposti L, Russo P, Saragoni S, Buda S, Rosano G, Pecorelli S, Pani L; OsMed Health-DB Network. Risk of hospitalization for heart failure in patients with type 2 diabetes newly treated with DPP-4 inhibitors or other oral glucose-lowering medications: a retrospective registry study on 127,555 patients from the Nationwide OsMed Health-DB Database. *Eur Heart J* 36: 2454-62, 2015.
 6. De Berardis G, Lucisano G, D'Ettore A, Pellegrini F, Lepore V, Tognoni G, Nicolucci A. Association of aspirin use with major bleeding in patients with and without diabetes. *JAMA* 307: 2286-94, 2012.
 7. Orsini Federici M, McQuillan J, Biricolti G, Losi S, Lebrech J, Richards C, Miglio C, Norrbacka K. Utilization Patterns of Glucagon-Like Peptide-1 Receptor Agonists in Patients with Type 2 Diabetes Mellitus in Italy: A Retrospective Cohort Study. *Diabetes Ther* 9: 789-801, 2018.