



Quel pasticciaccio della metanalisi canadese sui betabloccanti

Data 04 luglio 2006
Categoria cardiovascolare

Molti sono i problemi metodologici della metanalisi canadese che mette in guardia circa l'uso dei betabloccanti come antipertensivi di prima scelta negli anziani.

La metanalisi canadese (vedi <http://www.pillole.org/public/aspnuke/news.asp?id=2549>) recentemente pubblicata (CMAJ 2006 Jun 6; 174: 1737-1742) da cui emergerebbe che i beta-bloccanti non dovrebbero essere considerati farmaci di prima scelta negli anziani ipertesi a meno che non esista un'altra indicazione mentre nei giovani sarebbero associati ad una significativa riduzione della mortalità e della morbilità, presenta gravi manchevolezze metodologiche che sono di sotto evidenziate:

1) gli intervalli di confidenza di alcuni risultati che gli autori hanno enfatizzato per la loro 'significatività' sfiorano in realtà la 'linea del non effetto' (che per un Rischio Relativo corrisponde al valore di uno). Per i confronti betabloccanti vs placebo che riguardano i pazienti 'più giovani' ad esempio il Rischio Relativo dell' outcome composito (morte CV+ infarto + stroke) è 0,86 con intervalli di confidenza da 0.74 a 0.99. La 'significatività statistica' in questo caso coincide con la 'significatività clinica'? Ci serve in altri termini un intervallo di confidenza con un estremo vicino al valore di uno? Con un RR pari a 0.99 il Number Needed To Treat (NNT) sfiora il valore <infinito>, rendendo praticamente inesistente l' utilità clinica dell' intervento. Le stesse osservazioni si possono fare per i risultati ottenuti nei confronti betabloccanti versus farmaci attivi nei pazienti 'anziani', dove il Rischio Relativo è pari a 1.06 e dove gli intervalli di confidenza vanno da 1.01 a 1.10. Questo risultato 'significativo' sembra dover scoraggiare la scelta dei betabloccanti ogni volta sia possibile utilizzare altri farmaci nei pazienti 'anziani' ma anche con un RR pari a 1.01 il Number Needed To Harm (NNH) presenta valori troppo elevati per avere un significato pratico.

2) Ma indipendentemente dal loro (scarso) significato pratico: le osservazioni stimolate da questa metanalisi (ossia che l' utilizzo dei betabloccanti sembra produrre vantaggi assai scarsi rispetto al non far niente nei pazienti giovani e sembra comportare addirittura danni quando i betabloccanti sono scelti al posto di altri farmaci nei pazienti anziani) sono in realtà degne di fiducia?

Ecco qualche dubbio sulla solidità metodologica del lavoro:

3) Gli autori hanno analizzato separatamente in due grandi gruppi gli RCT inclusi nella revisione: trial che hanno reclutato pazienti 'giovani' e rispettivamente 'trial che hanno reclutato pazienti anziani'. L' età di 60 anni (= valore medio delle età dei pazienti reclutati) è stata utilizzata come 'cut off' per assegnare all' una o all' altra categoria di analisi gli studi sottoposti a revisione. La metanalisi effettivamente riporta i valori dell' età 'media' riscontrata entro ciascun campione. Chi ci assicura però che non esista un preoccupante 'overlap' di età tra le casistiche che caratterizzano ciascun trial? Simile sovrapposizione renderebbe piuttosto discutibili i risultati rilevati entro i gruppi 'giovani' e 'anziani' (su cui la metanalisi basa le proprie conclusioni e le proprie raccomandazioni!). L' ampiezza degli intervalli di confidenza citati al punto 1 –sempre che si consideri verosimile l' ipotesi di un effetto differenziale nei pazienti giovani e nei pazienti anziani- potrebbe essere giustificata proprio da 'contaminazioni anagrafiche' entro la casistica sottoposta ad analisi. Sarebbe stato molto più corretto (sembra impossibile che gli autori non ci abbiano pensato..semplice dimenticanza?) riportare nell' articolo anche le deviazioni standard intorno alle medie delle età dei pazienti reclutati in ciascun trial. La deviazione standard infatti ci dice molto sulla distribuzione dei valori di un parametro all' interno di un campione, e ci consente di fare calcoli molto interessanti..

Proviamo? Questi calcoli partono dal presupposto che la distribuzione delle età entro i bracci di ogni trial sia di tipo gaussiano. E' evidente che questo assunto è assai grossolano.

Infatti la 'population study' di ogni trial non è stata selezionata dalla popolazione generale in modo da essere 'rappresentativa' (ossia attraverso un campionamento random) ma è piuttosto rappresentata da una serie consecutiva di pazienti selezionati perché obbediscono ai criteri di inclusione della ricerca, che vengono successivamente distribuiti dalla randomizzazione entro i due bracci dello studio. La randomizzazione garantisce comunque che la distribuzione dei valori del parametro all' interno di ciascun braccio sia quella osservabile nel campione in toto.

La nostra approssimazione è probabilmente ragionevole per i trial di dimensioni particolarmente considerevoli, in cui le caratteristiche delle distribuzioni dei valori dell' età entro la 'population study' possono avvicinarsi a quelle osservabili all' interno della popolazione che ha generato il campione anche in assenza di un campionamento 'per forza' randomizzato. Si sottolinea che la forza di questo presupposto è fortemente condizionata anche dalle dimensioni dei trial. Negli RCT (vedi NORDIL o ASCOT/BPLA) dove sono state reclutate parecchie migliaia di persone la distribuzione delle età entro il campione segue un andamento probabilmente molto prossimo a quello gaussiano. Per gli studi di dimensioni più piccole esiste invece la possibilità di una distribuzione non perfettamente gaussiana (la varianza di qualsiasi parametro è inversamente proporzionale alla numerosità dei soggetti reclutati). Questo inconveniente –di ostacolo per i nostri calcoli- potrebbe essere degno di nota per i trial più 'datati', ossia per i confronti betabloccanti versus placebo, le cui dimensioni campionarie risultano inferiori a quelle dei più moderni confronti 'versus molecole attive'. Infatti nei 12 confronti 'betabloccanti versus altri farmaci' recensiti dalla metanalisi solo un trial (UKPDS n= 758) presenta una casistica inferiore a duemila soggetti. Dunque almeno per il gruppo dei 12 trial più 'moderni' l' assunto di una distribuzione gaussiana delle età a livello dei singoli bracci è da considerare, con ogni probabilità, abbastanza ragionevole.

Ora: se conosciamo media e deviazione standard dei valori dell' età dei vari campioni e se la distribuzione delle età segue la curva gaussiana, con un calcolo statistico relativamente semplice possiamo calcolare in modo fedele la



percentuale di soggetti che in ogni campione presenta una età inferiore o rispettivamente superiore al 'cut off' di 60 anni. Per esempio nel trial UKPDS (dove l'età media dei pazienti è 56 anni nel gruppo atenololo e 56,3 anni nel gruppo captopril) le deviazioni standard intorno alla media -non riportate nella metanalisi e rispettivamente corrispondenti a 8,2 e a 8,1 anni entro ciascun braccio- ci permettono di calcolare che la percentuale di ultrasessantenni reclutata nel braccio di intervento e nel braccio di controllo corrisponde rispettivamente al 31% per il gruppo atenololo e al 32% per il gruppo captopril. Se all'opposto prendiamo in considerazione un trial inserito nel gruppo 'anziani' come il NORDIL (dove l'età media nel braccio betabloccanti/diuretici è pari a 60,3 e l'età media nel braccio diltiazem è pari a 60,5) le deviazioni standard intorno a queste medie -non riportate dagli autori e corrispondenti a 6,5 anni per ciascun braccio) ci permettono di calcolare che la percentuale di soggetti con età inferiore a 60 anni è pari al 48% nel braccio di intervento e al 47% nel braccio di controllo!. A buon intenditor, poche parole...

4) Infine, come 'ciliegina', un'altra considerazione. Per fare un confronto metanalitico riguardo l'incidenza dell'outcome della metanalisi (si tratta dell'end-point composito: "morte cardiovascolare + "stroke" + "infarto miocardico") gli autori avrebbero dovuto avere a disposizione i dati relativi alla 'first occurrence' di ciascuno di questi end-point, per non contare due volte lo stesso paziente per lo stesso outcome. E' molto difficile -in realtà- disporre di questi dati. In primo luogo non tutti i trial recensiti dalla metanalisi hanno incluso l'outcome scelto dagli autori entro i propri end-point; alcuni studi per esempio riportano gli eventi "morte CV" o "infarto" o "stroke" come singleton end-point secondari. In questi casi raramente è possibile risalire dalla semplice lettura di un articolo alla percentuale di soggetti che hanno subito più di un end-point: queste informazioni richiedono il più delle volte un contatto diretto con i singoli ricercatori, cosa alquanto difficile per l'analisi di trial condotti qualche decennio fa. In alternativa - senza pretendere di costruire a tavolino un nuovo end-point- gli autori avrebbero dovuto utilizzare criteri di inclusione più attinenti gli obiettivi dichiaratamente perseguiti: per esempio avrebbero potuto scegliere di sottoporre a metanalisi solo i trial che avessero considerato in modo specifico quell'outcome composito tra i loro end-point primari o secondari (in una metanalisi questa differenza non presenta le criticità tipiche della interpretazione dei risultati di un singolo trial - vedi www.evidenzaqualitametodo.it). Ad una prima lettura di questo lavoro canadese pare invece che gli autori abbiano molto grossolamente l'incidenza del loro outcome composito attraverso un disordinato assemblaggio di trial caratterizzati da end-point a volte anche molto differenti senza assolutamente tener conto degli aspetti metodologici sopra riportati. Per fare un esempio, per il trial UKPDS le frequenze riportate nella metanalisi per l'end-point composito "morte CV + infarto + stroke" (34/358 e rispettivamente 48/400 nei due bracci) rappresentano in realtà i dati riportati dall'articolo originale per l'outcome 'morti dovute al diabete'. Altro esempio: per il trial ASCOT-BPLA le frequenze riportate per l'outcome composito "morte CV + infarto + stroke" nei due bracci (474/9618 e rispettivamente 429/9639) si riferiscono in realtà ai dati riportati nell'articolo originale per l'end-point primario ASCOT-BPLA (= 'infarto miocardico silente e non silente' + 'morte da causa coronarica'). Questa osservazione appare ancor più sconcertante se si considera che nella tabella originale dello studio ASCOT-BPLA i dati relativi all'outcome "morte cardiovascolare + infarto + stroke" erano stati comunque riportati (937/9618 e rispettivamente 796/9639 nei due bracci!). Per altri trial (esempio: ELSA, NORDIL, CONVINCE) l'outcome "morte cardiovascolare + infarto + stroke" era stato considerato anche nella ricerca originale e le frequenze vengono correttamente riportate nella metanalisi.

5) Insomma: un bel pasticcio. L'unica soluzione per sfuggire alle insidie dei messaggi a volte contrastanti della letteratura sia non tanto il nichilismo (cui prodest? sicuramente non ai pazienti!) quanto la selezione dei lavori più seri e metodologicamente più solidi attraverso l'ausilio di esperti o di pubblicazioni secondarie sicuramente affidabili (per esempio: la Cochrane Library, che pubblica sia revisioni sistematiche originali di elevatissima qualità sia recensioni di revisioni sistematiche prodotte altrove). Questo compito evidentemente non viene più assolto da chi dovrebbe 'di per se' rappresentare una garanzia di qualità per il lettore-utente: le grandi testate editoriali che pubblicano i lavori originali. Questo problema appare ancor più importante quando i messaggi rivolti al lettore sono particolarmente 'forti', come ad esempio per la metanalisi oggetto di questa valutazione critica.

La medicina generale però sta affilando le armi della metodologia e del critical appraisal.. aiutate ad affilare queste armi e a sostenere chi le utilizza.

Alessandro Battaglia

EQM - Evidenza, qualità e metodo in Medicina Generale

il testo completo del critical appraisal:

http://www.evidenzaqualitametodo.it/files/METANALISI_BETABLOCCANTI_GIOVANI_ANZIANI_PER_IL_SITO.pdf

Lancet 2000;356:359-65.

Lancet 2005;366:895-906.

BMJ 1998;317:713-20.

Circulation 2002;106:2422-7.

JAMA 2003;289:2073-82.