Vecchi e nuovi antiperstensivi pari sono

Data 04 dicembre 2005
Autore admin

L' ipotesi testata nello studio ASCOT-BPLA (8) è:
"L' amlodipina (eventualmente associata a perindopril) è più efficace dell' atenololo (eventualmente associato a bendrofumethiazide-potassio) nel ridurre l' incidenza di morti coronariche e infarti non fatali?".
Si tratta pertanto di uno studio di superiorità.
Gli autori hanno reclutato 19342 pazienti ipertesi a 'moderato' rischio cardiovascolare. L' end-point primario era rappresentato dall' outcome combinato . End-point secondari erano: mortalità da tutte le cause, stroke, outcome combinato , tutti gli eventi e le procedure CHD, morti cardiovascolari, scompenso cardiaco fatale e non fatale. End-point terziari erano: infarti miocardici silenti, angina instabile, angina stabile, arteriopatia periferica, aritmie pericolose, diabete , insufficienza renale. Erano considerati end-point terziari anche le analisi in alcuni sottogruppi prespecificati dei seguenti outcome: end-point primario, tutti gli eventi e le procedure cardiovascolari. I sottogruppi erano categorizzati per assenza /presenza di : diabete, abitudine al fumo, obesità, età>60 anni, sesso maschile, IVSX, precedenti vascolari, disfunzione renale, sindrome metabolica. Lo studio è un RCT in aperto, vale a dire che sia il medico caregiver che il paziente erano a conoscenza del braccio a cui il paziente era randomizzato. La rilevazione degli outcome è avvenuta però in cieco da parte di operatori esterni. La numerosità del campione è stata calcolata prevedendo una potenza statistica dell' 80% nel rilevare per l' outcome primario un Hazard Ratio pari a 0.84 con livelli di errore alfa pari a 0.05. Gli autori hanno considerato le violazioni del protocollo in un modello di analisi 'Intention to treat '. Nell' ottobre 2004 lo 'steering Committee' ha deciso la interruzione precoce della ricerca in seguito al riscontro di differenze significative nella incidenza di outcome secondari a sfavore del braccio assegnato all' atenololo (mortalità totale e altri end-points secondari). La durata media del follow-up è stata di 5.5 anni per un totale di 106153 anni-paziente. Sono stati persi al follow-up 292 pazienti (1.51% dell' intera casistica).
Alla fine dello studio era conosciuto lo stato vitale di 18965 pazienti . Hanno effettivamente assunto il trattamento assegnato dalla randomizzazione (pazienti compliant) 8119 pazienti (84,2%) assegnati al gruppo di intervento e 7115 pazienti (73,9%) assegnati al gruppo di controllo. Hanno assunto il trattamento assegnato all' altro braccio (pazienti cross-overs) 1520 pazienti assegnati inizialmente all' amlodipina (15.7%) e 2503 pazienti assegnati inizialmente all' atenololo (26.0%). Gli autori riportano che il 25% dei pazienti in toto (n= 4790) ha abbandonato lo studio per effetti collaterali del trattamento, senza differenze significative tra i due bracci. La randomizzazione (tabella 1 del lavoro originale) ha prodotto due bracci perfettamente confrontabili nelle condizioni basali. I pazienti assegnati all' amlodipina hanno ottenuto sulla PA risultati migliori rispetto ai pazienti assegnati all' atenololo; questo risultato era particolarmente evidente a tre mesi dall' inizio del trial. L' incidenza dell' end-point primario (Morte coronarica + Infarto non fatale) non era significativamente diversa tra i due bracci (HR 0.90 IC 0.79-1.02) (figura 4 del lavoro originale). L' incidenza degli outcome secondari : scompenso cardiaco, infarto miocardico silente, aritmie minacciose non era significativamente diversa tra i due bracci. L' incidenza degli outcome secondari : , tutti gli eventi CHD, tutti gli eventi e le procedure CVD, mortalità da tutte le cause, mortalità CVD , stroke era significativamente minore nel gruppo assegnato alla amlodipina. L' incidenza degli outcome terziari : angina instabile, arteriopatia periferica, diabete, insufficienza renale era significativamente minore nel braccio assegnato ad amlodipina. L' analisi per sottogruppi eseguita solo per l' outcome ha dimostrato una incidenza di questo end-point significativamente diversa tra i due bracci a favore del braccio di intervento (amlodipina) (Figura 7 del lavoro originale). Il testo non riporta però per gli stessi sottogruppi i risultati rilevati nell' analisi dell' outcome primario . Le due analisi di end-point secondari eseguite ex post "end-point primario+procedure di rivascolarizzazione" e "Morti CVD+Infarto Miocardico+ Stroke" hanno dimostrato differenze significative a favore del braccio trattato con amlodipina (fig. 4 del lavoro originale). Gli autori affermano che lo studio ASCOT-BPLA ha dimostrato che 'in pazienti ipertesi a moderato rischio CVD il trattamento con amlodipina (eventualmente integrato con perindopril) è vantaggioso rispetto al trattamento con atenololo (eventualmente integrato con diuretico tiazidico) in quanto "riduce l' incidenza di tutti i tipi di eventi cardiovascolari" e "la mortalità generale" nonchè il Gli autori ammettono di non aver riscontrato differenze significative tra i due gruppi in termini di incidenza di outcome primario "Morte CHD + infarto non fatale" ma affermano che ciò può essere attribuito a sottocampionamento in quanto l' interruzione precoce dello studio ha comportato la rilevazione di 903 casi di outcome primario contro i 1150 su cui era stata calcolata la potenza statistica. Gli autori concludono definendo 'rassicurante' la conferma dell' abbassamento dell' incidenza dell' outcome composito "tutti gli eventi e tutte le procedure CVD" in tutti i sottogruppi considerati.
Per ragioni di spazio non viene commentata la trasferibilità dei risultati del trial a popolazioni diverse da quella esaminata (Validità esterna). In ogni caso i revisori ritengono che essi siano discretamente trasferibili alla popolazione europea in generale, tenendo conto però che la casistica era composta prevalentemente da maschi a rischio vascolare medio-alto (con almeno tre fattori di rischio). L' ASCO- BPLA è uno "studio di superiorità" basato su una ipotesi chiara, vale a dire dimostrare l' esistenza tra il braccio assegnato ad amlodipina e a perindopril e il braccio assegnato ad atenololo e a bendrofumethiazide-potassio una differenza precisa in termini di incidenza dell' outcome primario. Il disegno Randomizzato e Controllato è stato ben condotto, nonostante l' allocation e la conduzione dello studio siano state eseguite in 'aperto'; gli autori hanno però eseguito in 'blinding' la rilevazione degli outcome. Per la randomizzazione sono state utilizzate sequenze generate dal computer. L' efficacia della randomizzazione è confermata dall' eccellente distribuzione delle caratteristiche 'baseline' dei due gruppi a confronto (tabella 1 del lavoro originale) . La qualità del confronto è stata garantita anche dalla analisi multivariata (modello di Cox) per fattori prognostici noti utilizzata dagli autori. Il sample size è stato determinato con metodi statisticamente corretti prevedendo di rilevare ex ante una precisa differenza tra i due bracci in termini di incidenza dell' outcome primario per livelli prestabiliti di errore alfa e beta. La qualità del follow-up è stata notevole, con un numero irrilevante di drop-outers. Le percentuali di cross-overs e di non

compliant per varie cause sono simili a quelle riscontrate in altri grandi trial di lunga durata (3). L' analisi degli outcome è stata condotta, correttamente, secondo il modello 'Intention to treat'. Gli autori hanno previsto ex ante l' analisi di 18 sottogruppi e hanno specificato in maniera trasparente di aver eseguito ex-post l' analisi di due outcome secondari. La lunghezza del follow-up era adeguata alla rilevazione di tutti gli outcome. L' opinione dei revisori -in sintesi- è che lo studio sia un trial di alta qualità metodologica. Chi scrive però ha delle forti riserve sulle modalità con cui gli autori hanno interpretato i risultati della ricerca.
Come nella maggior parte delle sperimentazioni gli autori hanno utilizzato come outcome primario un outcome composito al fine di aumentare il rischio assoluto dell' evento osservabile in condizioni basali. Ciò aumenta l' efficienza statistica della ricerca in quanto arruolare pazienti ad alto rischio può consentire di rilevare una differenza maggiore tra un braccio e l' altro e rendere quindi possibile -a parità di potenza- il reclutamento di campioni meno numerosi rispetto a quelli necessari per analizzare l' incidenza dei singoli outcome presi da soli (5). Nella fattispecie nello studio ASCOT-BPLA era stato previsto di reclutare 18000 pazienti per dimostrare differenze del 16% tra i due bracci i termini di riduzione relativa di rischio. Il campione randomizzato (19342 pazienti) era ampiamente sufficiente a dimostrare questa differenza, che nei fatti non è stata rilevata. Gli autori si appellano al fatto che l' interruzione precoce della sperimentazione non ha concesso di rilevare un numero di 247 eventi che avevano invece ipotizzato di rilevare completando il follow-up secondo protocollo. Si può obiettare però che questo è un problema comune alla maggior parte dei i trial in cui per ragioni etiche o di altro genere la sperimentazione viene interrotta precocemente. Gli autori invece hanno basato le proprie conclusioni sull' analisi degli end-point secondari, enfatizzando come alternativa alla "mancata dimostrazione di significatività" dei risultati che riguardano l' outcome primario la "significatività" dei risultati rilevati per gli outcome secondari. L' analisi degli end-point secondari dovrebbe essere -più correttamente- utilizzata come valore aggiunto a supporto delle conclusioni ricavate dall' analisi dell' end-point primario (7) oppure - al limite- per generare ipotesi di lavoro (2). Le conclusioni degli autori rispecchiano quindi in modo paradigmatico quello che il dott. Moyè ha definito nel 1999 'nascondere la zia scema in soffitta', essendo ' la zia scema* rappresentata dai risultati riscontrati per l' outcome primario (1). Una lettura corretta del trial avrebbe dovuto invece rispondere a questa domanda: "l' ipotesi primaria dello studio è stata soddisfatta"? .Il trial era stato architettato per rispondere al preciso quesito: "il trattamento con amlodipina (ed eventualmente con pirendopril) porta ai pazienti ipertesi vantaggi in termini di riduzione della incidenza di morti coronariche e infarti non fatali rispetto al trattamento a base di atenololo (ed eventualmente con bendroflumetazide". In base ai risultati riportati la risposta è : no, in quanto i valori degli intervalli di confidenza della stima dell' Hazard Ratio variano per l' outcome primario da 0.79 a 1.02 (P=0.1052).Le analisi di molteplici outcome sui pazienti randomizzati ai due bracci (= analisi di endpoint secondari) come le analisi di singoli o più outcome in strati del campione originale (= analisi per sottogruppi) rappresentano situazioni in cui si eseguono confronti multipli entro lo stesso campione, quello inizialmente arruolato nella ricerca. In primo luogo occorre ricordare che la numerosità dei soggetti da reclutare in un trial viene tarata solo sull' outcome primario. In pratica quando si studia l' incidenza di 'eventi' gli autori ipotizzano ex ante che alla fine della ricerca possa essere rilevata una precisa differenza tra i due bracci nell' incidenza dell' end-point primario e affidano al calcolo statistico la definizione di un sample size sufficiente a poterla dimostrare (2). In questo calcolo viene accettato un certo margine di errore, che può andare in due direzioni opposte. Esiste infatti sempre il rischio di identificare -erroneamente- differenze tra i due bracci quando in realtà non esistono (errore alfa: il livello accettato è di solito 5%) e al contrario il rischio di non riuscire –erroneamente- a dimostrare differenze tra i due bracci quando queste in effetti esistono (errore beta: il livello accettato è di solito 10%-20%). Quando si analizza l' incidenza di molteplici "outcome secondari" entro il campione originale il mancato riscontro di 'significatività statistica' per la differenza rilevata tra i due bracci nell' incidenza di uno di questi deve essere interpretato con grande cautela. Infatti a volte ciò si verifica solo perchè -essendo la numerosità del campione originale tarata non su quell' endpoint ma sull' outcome primario- la differenza riscontrata tra i due bracci per quell' endpoint è troppo piccola per poter essere accompagnata da dimostrazione di "significatività". Il vero protagonista della ricerca è infatti l' end-point primario: una volta scelto questo obiettivo tutta la potenza statistica dello studio viene spesa per le analisi ad esso riferite. Ecco allora che la 'non dimostrazione di significatività' statistica' per un outcome secondario può coincidere semplicemente con un errore "di tipo II" (errore beta, o 'falso negativo') ossia col non rilevare differenze tra i due bracci quando queste in effetti esistono. L' errore beta è tanto più frequente quanto più i confronti tra i due bracci vengono eseguiti in presenza di numerosità campionarie inadeguate. Questa situazione si realizza ancora più facilmente nelle "analisi per sottogruppi". In questi casi infatti l' errore beta è in agguato non solo per i risultati che si riferiscono agli end-point secondari ma anche per quelli che riguardano l' end-point primario: quando i due bracci vengono confrontati entro il contesto di un "sottogruppo", ovviamente la numerosità di questo strato del campione è -per forza!- sempre minore di quella del campione originale. E' interessante notare che nello studio ASCOT-BPLA gli outcome secondari sono rappresentati per lo più da end-point compositi non tra di loro indipendenti e che le dimensioni del campione dello studio ASCOT-BPLA sono per il 62.5% (10/16) inferiori a quelle necessarie a dimostrare nei singoli casi le differenze rilevate tra un braccio e l' altro. Gli autori anzichè limitarsi a giustificare solo per l' outcome primario la possibilità di un sottocampionamento dovuto a precoce interruzione dello studio avrebbero dovuto pertanto esprimere incertezze anche per i risultati rilevati per gli outcome secondari.Un secondo ordine di problemi collegato a ogni tipo di confronti multipli è rappresentato dall' aumentato rischio di rilevare differenze tra i due bracci quando invece queste non esistono ("falsi positivi" o errori di tipo I o 'errori alfa'): i confronti multipli infatti aumentano (anche) notevolmente la probabilità di ottenere risultati legati solo ad un effetto del caso. Un esempio impressionante di tutto ciò è fornito dall' analisi dei 146 sottogruppi dello studio BATH (studio controllato dove si analizzava la mortalità dei pazienti trattati con con betabloccanti): le differenze tra braccio di intervento e braccio di controllo in termini di tassi di mortalità analizzate per ciascuno dei 146 sottogruppi si distribuivano secondo una perfetta curva gaussiana (4). In elaborazioni diverse da quelle previste per l' outcome primario (come appunto nelle analisi di outcome secondari o nelle analisi per sottogruppi) l' errore random può giocare quindi un ruolo molto importante nel condizionare la direzione dei risultati e ciò anche in presenza della desiderata 'significatività statistica' (2). Questo pericolo aumenta in modo proporzionale al numero dei confronti eseguiti all' interno del campione originale.
B) Gli autori hanno enfatizzato in 18 sottogruppi i risultati che riguardano un outcome composito
Questa analisi riguarda 18 sottogruppi determinati ex ante dal protocollo ed è raffigurata nella figura 7 del lavoro

originale, che riporta i risultati che si riferiscono all' outcome composito "total cardiovascular events and procedures". Questi sono perfettamente coerenti con quanto riportato-per lo stesso outcome nella analisi eseguita sulla casistica in toto e visibile nella figura 4 del lavoro originale. L' outcome composito "total cardiovascular events and procedures" è in ogni caso un outcome secondario, e questa particolare scelta rappresenta altresì a giudizio di chi scrive una ulteriore forzatura metodologica verso la direzione dei risultati che i ricercatori si proponevano di sostenere. Non dimentichiamo che "total cardiovascolar events and procedures" rappresenta la somma di tutto (ma proprio tutto) quello che poteva capitare alla casistica. E' per questo che il numero di eventi (1362 nel braccio di intervento, 1602 nel braccio di controllo -vedi figura 4 del lavoro originale) è il più alto riscontrato per tutti gli outcome dello studio. In altri termini con questa scelta quasi un paziente su cinque reclutato dallo studio ASCOT-BPLA era 'ad alto rischio' per questo outcome (Rischio Assoluto dei controlli = 16.6%). Come dimostrato da Sackett nel saggio pubblicato nel 2001 su CMAJ la scelta di pazienti 'ad alto rischio' rappresenta uno dei 4 metodi per forzare i risultati verso un aumento della dimensione dell' effetto (5). Inoltre la dimensione dell' effetto rappresenta uno dei tre principali determinanti della precisione della stima (5). In tal senso appare abbastanza logico che la differenza tra i due bracci nella frequenza di questo outcome secondario risultasse grande e con un intervallo di confidenza molto stretto (figura 4 del lavoro originale). Il riscontro di una grande coerenza tra i 18 sottogruppi dei risultati che si riferiscono all' outcome composito non deve quindi stupire più di tanto. Il lettore è invece colpito in modo negativo da fatto che nonostante gli autori alla voce 'Procedures' dichiarino di aver eseguito sui sottogruppi anche l' analisi dei risultati che si riferiscono all' outcome primario vengano riportati nel testo solo i risultati che si riferiscono all' outcome composito secondario . Certamente l' analisi dell' outcome primario in questi sottogruppi avrebbe confermato per lo più le conclusioni riportate nell' analisi di tutto il campione.Lo studio ASCOT-BPLA è un trial di elevata qualità metodologica i cui risultati sono forzati verso direzioni volute dagli autori. Chi scrive sostiene, analogamente a più illustri opinioni, che esistano molti argomenti a favore di un utilizzo più prudente e ponderato delle analisi degli outcome secondari e dei sottogruppi (2). Una volta scelto l' outcome primario tutta la potenza dello studio è stata spesa per le analisi ad esso riferite : l' effetto del caso può giocare quindi un ruolo molto forte nel condizionare i risultati di efficacia riscontrati in singoli sottogruppi o che riguardino outcome diversi da quello primario, anche in presenza della desiderata 'significatività statistica'(2). La US Food and Drug Administration, coerentemente, per l' autorizzazione all' immissione in commercio di nuove molecole richiede trial randomizzati ben condotti in cui la potenza statistica sia tarata su differenze relative ad un outcome primario (6). In un' epoca in cui qualsiasi Servizio Sanitario sempre più tende a considerare l' analisi costo-efifcacia come parte integrante dei processi decisionali di rimborsabilità dei nuovi farmaci diventa sempre più importante che le evidenze scientifiche siano interpretate in modo corretto, al fine di evitare di incorporare nella pratica clinica scelte terapeutiche inappropriate (2). Pertanto ispirandoci alle considerazioni cliniche che hanno giustificato l' organizzazione dello studio ASCOT-BPLA non siamo autorizzati ad utilizzare i suoi risultati per consigliare l' utilizzo dei 'nuovi' farmaci rispetto ai 'vecchi' rappresentati dal betabloccante e dal diuretico o, peggio, a sconsigliare (come hanno fatto gli autori) l' utilizzo di questi ultimi in alternativa ai primi.
Come l' ALLHAT "non ha dimostrato superiorità" del diuretico rispetto al calcioantagonista e dell' ACE inibitore testati in quel trial (3) allo stesso modo l' ASCOT "non ha dimostrato superiorità" dell' amlodipina eventualmente associata a perindopril rispetto al betabloccante eventualmente associato al diuretico. Non esistono dunque ancora prove sufficienti , se vogliamo applicare in modo molto esteso e molto liberamente un 'effetto classe' , che i 'nuovi' farmaci siano 'migliori' dei vecchi in termini di riduzione di morte coronarica e di infarti non fatali.

Alessandro Battaggia, Alberto Vaona , Giulio Rigon, Saffi Ettore Giustini E.Q.M. Evidenza, Qualità e Metodo in Medicina Generale S.I.M.G. Società Italiana di Medicina generale

(1) Moyè LA. End-point interpretation in clinical trials: the case for discipline. Control Clin trials 1999 20:40-9
(2) Freemantle N..Interpreting the results of secondaru end points and sobgroup analyses in clinical trials: should we lock the crazy aunt in the attic?.BMJ 322 2001 322:989-91
(3) The ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group Maior Outcomes in High-Risk Hypertensive Patients Randomized to Angiotensin-Converting Enzyme Inhibitor or Calcium Channel Blocker vs Diuretic - JAMA, 2002 - 288:2981
(4) Furberg CD , Byington RP What do subgroup analyses reveal about differential response to beta-blocker therapy? The Beta-Blocker Heart Attack Trial experience. Circulation. 1983 Jun;67(6 Pt 2):98-101.
(5) Sackett DL. Why randomized controlled trials fail but needn't: 2. Failure to employ physiological statistics, or the only forlula a clinician-trialist is ever lickely to need (or understand). JAMC 2001 165(9):1226-37
(6) Fischer LD. Carvedilol and the Food and drug Administration (FDA) approval process: the FDA paradigm and reflections on hypothesis testing. Contr Clin Trials 1999 20:16-39
(7) University of Washington. RCDRC Guidelines for Clinic Trial Designs http://depts.washington.edu/rcdrc/app.html
(8) Björn Dahlöf et al. Prevention of cardiovascular events with an antihypertensive regimen of amlodipine adding perindopril as required versus atenolol adding bendromethiazide as required, in the Anglo-Scandinavian Cardiac Outcomes Trial-Blood Pressure Lowering Arm (ASCOT-BPLA): a multicentre randomised controlled trial'. Lancet 2005; 366: 895–906