



Pillole di EBM - Capitolo 11

Data 11 marzo 2007
Categoria scienze_varie

In questo capitolo ci occuperemo di un aspetto particolare degli studi clinici: le analisi per sottogruppi.

Quando si disegna uno studio di solito si definiscono degli end-point che poi si andranno a misurare: quanti infarti fatali e non fatali si avranno alla fine dello studio, quanti ictus, eccetera.

Abbiamo visto nel capitolo sugli end-point che la potenza statistica dello studio viene tarata sull' end-point primario, ma possono essere previsti anche degli end-point secondari.

Spesso i ricercatori però non si limitano ad analizzare i dati sull'intero campione arruolato ma scompongono quest'ultimo in vari sottogruppi. Per esempio si decide di andare a vedere se tra tutti i pazienti arruolati il trattamento si è dimostrato più utile in particolari tipi di pazienti (nelle donne piuttosto che negli uomini, nei diabetici piuttosto che negli obesi, ecc.) oppure se il trattamento porta a risultati diversi nei soggetti che fumano rispetto a chi non fuma o ancora in coloro che assumono regolarmente il farmaco rispetto a chi ha una cattiva compliance farmacologica e così via.

Questo modo di procedere è utile perché permette di ricavare molte informazioni, ma esse vanno sempre prese con cautela in quanto i risultati potrebbero essere dovuti al caso. E' buona norma quindi considerare i risultati derivanti da una analisi per sottogruppi come un'ipotesi che dovrebbe essere convalidata da studi successivi. Comunque le analisi per sottogruppi vengono considerate più affidabili se erano originariamente previste nel protocollo del trial, lo sono meno se vengono effettuate a posteriori e senza essere state originariamente predefinite.

L'ideale sarebbe non solo che l'analisi fosse già stata definita prima, ma che venisse usata la cosiddetta "randomizzazione stratificata" che permette una eguale distribuzione nei due bracci dei vari sottogruppi. Nello studio Val-heFT pazienti con scompenso cardiaco vennero trattati con valsartan o placebo. Siccome è noto che i betabloccanti incidono sulla prognosi dello scompenso si decise addirittura di randomizzare i pazienti che assumevano e non assumevano betabloccanti ai due gruppi così che la loro presenza nel gruppo trattamento e nel gruppo controllo fosse omogenea.

Ma perché diciamo che le analisi per sottogruppi possono dare risultati inaffidabili?

I moderni metodi di analisi statistica effettuati con potenti elaboratori elettronici permettono di valutare decine e decine di dati e qualche volta salta fuori qualche risultato che apparentemente ha una significatività statistica, ma in realtà è dovuto al semplice gioco del caso.

Per esempio se viene fatto uno studio su 20.000 pazienti ipertesi e si decide di fare un'analisi per sottogruppi dividendo i pazienti per età (maggiori o minori di 65 anni), per la presenza o meno di diabete, per lo stato di fumatore (fumo sì/no), per il sesso (maschi/femmine), per l'uso o meno di aspirina (asa sì/no), per la presenza o meno di colesterolo alto (colesterolo > 200 mg/ml sì/no) eccetera, si ottiene un numero incredibile di combinazioni. L'analisi computerizzata potrebbe allora mostrare che il farmaco, che nello studio nel suo insieme non ha evidenziato risultati favorevoli, sia invece efficace in una certa sottopopolazione (per esempio nelle donne con più di 65 anni, con colesterolo > 200 mg/ml senza diabete e che non fumano). E' probabile che questo risultato sia semplicemente dovuto ad una combinazione casuale. E' stato dimostrato che tante più sono le analisi per sottogruppi effettuate in un determinato campione e quindi tante più sono le combinazioni, tanto più aumenta la probabilità di trovare un risultato apparentemente significativo dal punto di vista statistico ma in realtà dovuto solamente al gioco del caso.

Chi si occupa di valutare criticamente gli studi clinici considera una serie di aspetti: come si è già accennato, per esempio si controlla se l'analisi per sottogruppi era stata pianificata o meno nel protocollo dello studio, oppure viene valutata l'entità dell'effetto trovato nei vari sottogruppi rispetto al campione totale, o ancora la plausibilità o meno dei risultati, ecc.

Si tratta di aspetti notevolmente complessi, per quanto ci riguarda basti ricordare il seguente principio: considerare in genere preliminari e quindi meritevoli di ulteriori studi i risultati di una analisi per sottogruppi, ancorché ben progettata e condotta.

Restano da esaminare le analisi a posteriori (o post-hoc analysis). Di che cosa si tratta?

Con questo termine si intendono quelle analisi non contemplate nel protocollo di ricerca che vengono effettuate dopo la conclusione della raccolta dati. E' un fenomeno molto frequente in letteratura: per esempio in uno studio è stato calcolato che circa metà dei lavori cardiologici considerati erano analisi a posteriori effettuate su vari sottogruppi. Durante uno studio, infatti, vengono raccolti molti dati sui pazienti (per esempio il tipo di dieta, la loro attività fisica, il tipo di farmaci assunti, le loro abitudini sessuali, voluttuarie, lo stato socio-economico, i loro viaggi, i vari passatempi, ecc). Succede così che dopo la fine dello studio venga la curiosità di andare a vedere per esempio quello che succede in certi tipi di pazienti rispetto ad altri.

Alcuni ricercatori, per dimostrare quanto possa essere ingannevole questo tipo di analisi, ne fece una sui dati di uno studio che aveva valutato i beta-bloccanti nel post-infarto: risultò che il farmaco era molto più efficace nei pazienti nati sotto un determinato segno zodiacale rispetto ad altri!. Ora, per quanta fede si possa avere nei segni zodiacali e nell'oroscopo, è un po' difficile ammettere che questo risultato abbia validità scientifica.

Vale quindi l'avvertenza di sempre: le analisi a posteriori permettono di esplorare molteplici aspetti ma è opportuno prendere con le molle questi dati e considerarli più che altro alla stregua di ipotesi.

Qualcuno è arrivato a definire le analisi a posteriori una specie di "tortura dei dati".