



Il Manuale di Clinica Pratica

Titolo La statistica in medicina
Data 25 gennaio 2006 alle 21:53:00
Autore R. Rossi

Rossi:
Perché parlare di statistica in un manuale di Clinica Pratica? Che c'azzecca?

Ressa:
Perché è incredibile come un medico, anche di ottimo livello, possa per anni dispensare ai suoi assistiti consigli e terapie senza chiedersi quanto, statisticamente, questo incida sul decorso naturale delle patologie. Ovviamente parliamo principalmente di quelle a carattere cronico (ad es. ipertensione, diabete), perché nelle acute il problema non si pone, nell'immediato, in quanto, generalmente, si passa subito all'azione, senza frapporre indugi. Quando poi l'ignaro medico decide di mettere il naso nel campo della statistica viene travolto da un'infinità di formule che gli causano, di solito, un'immediata ripulsa e lo ricacciano immediatamente nella condizione precedente di ignoranza. Ci limiteremo, quindi, a trattare solo l'essenziale, l'irrinunciabile, di cui ogni medico dovrebbe essere padrone.

Rossi:
Mi hai quasi convinto. Da dove cominciamo?

Ressa:
Di solito, per verificare l'effetto del nostro operare terapeutico su pazienti affetti da una determinata patologia, usiamo degli studi clinici in cui dividiamo i soggetti arruolati in due gruppi: il gruppo dei trattati e quello dei non trattati (di controllo). Ci si pone degli obiettivi da raggiungere in un X tempo e si verificano, alla fine, i risultati con le differenze tra i due gruppi. E' chiaro che il discorso è inficiato, fin dall'inizio, da un vizio di fondo, noi mettiamo insieme soggetti che sono affetti da medesima patologia ma che sono diversissimi dal punto di vista genetico, psicologico, ambientale; dall'altra parte queste variabili non sono facilmente standardizzabili.

Rossi:
Allora, che possiamo fare? Ressa:
In qualche modo a questo limite cerca di porre rimedio la randomizzazione: i soggetti partecipanti allo studio vengono assegnati ai due gruppi in modo del tutto casuale affinché fattori di rischio noti e non noti siano distribuiti equamente (vedi capitolo sull'EBM). Se non si ricorresse all'artificio della randomizzazione può per esempio succedere che la maggior parte dei soggetti diabetici, infartuati o con altra patologia finisca in uno solo dei due gruppi, creando uno squilibrio che indebolirebbe i risultati dello studio.

Rossi:
E se mi facessi un esempio?

Ressi:
Prendiamo uno studio che si propone, ad esempio, di testare l'efficacia dell'aspirina, nella prevenzione dell'infarto, in pazienti affetti da cardiopatia ischemica. Li dividiamo in due gruppi randomizzati e quantitativamente sovrapponibili (il gruppo dei trattati lo chiamiamo T, il gruppo di controllo lo chiamiamo C). Alla fine del tempo di osservazione che ci siamo posti, ad esempio cinque anni, confrontiamo i risultati tramite tre indicatori principali di valutazione: RIDUZIONE DEL RISCHIO ASSOLUTO (ARR: absolute risk reduction), RIDUZIONE DEL RISCHIO RELATIVO (RRR: relative risk reduction), NNT (number needed to treat, cioè il numero di pazienti da trattare con l'aspirina affinché almeno uno di loro ottenga il risultato sperato, cioè la prevenzione dell'infarto). Noi li abbiamo chiamati indicatori per usare un termine facilmente comprensibile ma in gergo tecnico ARR, RRR ed NNT vengono definiti misure di efficacia.

Rossi:
Comincia a spiegare meglio, ho le idee confuse al riguardo... Ressa:
Il rischio assoluto (AR = Absolute Risk) indica quanti pazienti (in numero assoluto) hanno avuto l'evento clinico che stiamo studiando (infarto) sia nel gruppo dei trattati con aspirina sia in quello di controllo (non trattati); supponiamo che nel primo siano 8 su 100 (AR trattati = 8) e nel secondo 10 su 100 (AR controlli = 10), avremo quindi rispettivamente, nel primo gruppo, un rischio assoluto dell'8%, nel secondo del 10%. La RIDUZIONE DEL RISCHIO ASSOLUTO (ARR: Absolute Risk Reduction) sarà la differenza dei due valori (10 meno 8) quindi: 2% (ARR = 2%)
Lo scarto sembra piccolo e, di conseguenza, l'efficacia del trattamento bassa.

Rossi:
Direi sconsolante...

Ressa:



Se, però, sottoponiamo gli stessi dati alla verifica dell'indicatore della riduzione del rischio relativo le cose sembrano cambiare; questo parametro, infatti, ci dice di quanto si è ridotto, in percentuale, un certo evento clinico (in questo caso l'infarto) nel gruppo dei pazienti trattati rispetto al gruppo di controllo (non trattati con aspirina). I risultati vanno espressi in termini di RIDUZIONE DEL RISCHIO RELATIVO tra i due gruppi usando una semplicissima formula; non scoraggiamoci è veramente elementare, si tratta di fare in sequenza: una divisione ed una moltiplicazione.

RRR (riduzione del rischio relativo) = ARR / AR controlli $\times 100$

Nel nostro esempio $ARR = 2$, mentre AR controlli = 10; inserendo questi dati nella formula otteniamo $2/10 = 0.2 \times 100 = 20$.

Abbiamo cioè ottenuto una riduzione del rischio relativo del 20%; questo ci appare un grande risultato, contrariamente alla riduzione del rischio assoluto, precedentemente calcolata, che era un misero 2%. In effetti con il trattamento siamo passati da 10 infarti a 8, quindi abbiamo avuto una riduzione del 20% in termini relativi ma del 2% in termini assoluti. Gli stessi dati, quindi, visti da angolazioni diverse, hanno un impatto emozionale opposto, nel caso della riduzione del rischio assoluto la delusione è grande, in quella del rischio relativo si ha, invece, l'impressione che si possa ottenere molto mettendo in opera una certa strategia terapeutica e l'entusiasmo vola.

Non a caso, i risultati degli studi, specialmente dalle ditte produttrici dei farmaci, vengono riferiti in termini di riduzione del rischio relativo. Rossi:

Per la miseria! Parlare solo di rischio relativo può farci credere di essere troppo bravi...

Ressa:

Per semplificare tutto il discorso interviene l'ultimo indicatore, forse il più importante: l'NNT.

Esso indica quanti pazienti, affetti da una determinata patologia, devo trattare con un determinato farmaco, affinché abbia la probabilità che almeno UNO di essi riceva i benefici che mi sono posto come obiettivo.

Altra formuletta di calcolo, ancora più semplice:

$NNT = 100 / ARR$ (riduzione del rischio assoluto, che abbiamo già calcolato precedentemente)

Nell'esempio citato abbiamo: $100/2$ che fa 50.

Questo $NNT = 50$ vuol dire che devo trattare con aspirina 50 PAZIENTI per 5 anni affinché UNO eviti l'infarto, ovviamente non posso prevedere QUALE sia il fortunato tra i 50.

Questo vuole anche dire che ne tratterò inutilmente 49, magari esponendoli ad effetti collaterali anche pericolosi. Detto in altre parole l'intervento sarà utile in un paziente e inutile in 49.

Rossi:

Grazie Ressa, le tue spiegazioni sono state esaurienti e nello stesso tempo molto chiare.

Ressa:

Se siete sopravvissuti finora, diamo questo CONSIGLIO FINALE: di fronte a chi ci mostra risultati espressi in termini di riduzione del rischio relativo è bene non avere timori reverenziali (si tratti di informatori scientifici o colleghi specialisti) e domandiamo sempre qual è l'NNT e quanto deve durare il trattamento per ottenerlo. Come esemplificazione diamo alcuni NNT ricavati dalla letteratura:

- la somministrazione di una statina per 5 anni in soggetti con precedente infarto evita un decesso circa ogni 35 pazienti trattati
- la somministrazione di una statina per 5 anni a soggetti senza precedenti cardiovascolari (prevenzione primaria) evita un decesso ogni 110-160 pazienti trattati
- la somministrazione di un trattamento antipertensivo per 5 anni a soggetti a basso rischio evita un decesso cardiovascolare ogni 800-950 trattati
- la somministrazione di un trattamento antipertensivo per 5 anni a soggetti a medio rischio evita un decesso cardiovascolare ogni 160-190 trattati
- la somministrazione di un trattamento antipertensivo per 5 anni a soggetti a rischio medio-elevato evita un decesso cardiovascolare ogni 25-50 trattati
- la somministrazione di aspirina per 5 anni a soggetti a basso rischio cardiovascolare evita un decesso ogni 2000-10000 trattati
- la somministrazione di aspirina per 5 anni a soggetti a rischio medio-elevato evita un decesso cardiovascolare ogni 300-600 trattati

Rossi:

A leggere questi dati, espressi dai rispettivi NNT, l'ignaro medico si accorge improvvisamente di quanta sia diversa la realtà dalla "medicina dagli effetti speciali", molto in voga nei giorni nostri e dalla quale, forse, lui stesso era stato plagiato.

Questo ci dovrebbe portare ad un atteggiamento umile nel nostro operare, ci consola il fatto che i risultati su esposti siano il massimo ottenibile in questo momento storico.

Ressa:

Un altro campo in cui la statistica entra prepotentemente in giuoco è quella degli ACCERTAMENTI DIAGNOSTICI, come si spiegherà anche nel capitolo sugli screening. Ogni esame ha 2 parametri di valutazione: sensibilità e specificità.

Se dico che un esame diagnostico ha una sensibilità del 90% nell'individuare una certa malattia vuol dire che su 100 soggetti malati di quella patologia, sottoposti al test, 90 saranno positivi e 10 saranno negativi (falsi negativi). Se dico che un esame ha una specificità del 90% vuol dire che su 100 soggetti non affetti da quella patologia, sottoposti al test, 90 saranno negativi e 10 saranno positivi (falsi positivi).



Mnemonicamente viene più facile pensarla così: se mi dicono che un test ha una sensibilità del 90% vuol dire che solo 10 malati (100-90) sono FALSI NEGATIVI cioè risultano erroneamente non affetti dalla patologia in esame, se il test ha una specificità del 90% vuol dire che solo 10 soggetti non affetti dalla malattia (100-90) sono FALSI POSITIVI, cioè risultano erroneamente malati.

Conoscere la sensibilità e la specificità di un test è importante perché permette di avere un'idea dell'efficacia del test stesso.

Dato che difficilmente, nella pratica clinica, avremo test che hanno contemporaneamente sensibilità e specificità elevate, come nell'esempio su scritto, la lezione da imparare è che gli esami diagnostici vanno richiesti su ben specifiche indicazioni cliniche, per non trovarsi poi in labirinti ancora più complicati.

Appendice (per gli amanti dei numeri)

Tuttavia se ci limitassimo a quanto detto sopra, potremmo essere tratti in inganno. Se ho un test che ha una sensibilità e una specificità del 90% potrei anche dirmi soddisfatto perché ci saranno solo un 10% di falsi positivi e di falsi negativi. Purtroppo la statistica non è così semplice.

Bisogna tener conto infatti anche della "prevalenza" della malattia che si vuol diagnosticare. Prendiamo per esempio un esame di laboratorio e supponiamo che abbia una sensibilità ed una specificità del 90% per la diagnosi della patologia X. Supponiamo ancora che tale patologia abbia una prevalenza, nella popolazione generale del 2% (cioè ogni 100 persone prese a caso dalla popolazione generale 2 hanno la malattia X).

Questo vuol dire che ogni 1.000 persone, 20 sono malate e 980 sono sane.

Applichiamo il test alle 20 persone malate: siccome il test ha una sensibilità del 90% vuol dire che di esse, 18 (cioè il 90% di 20) avranno il test positivo e 2 avranno il test negativo.

Applichiamo adesso il test alle 980 persone sane: siccome il test ha una specificità del 90% vuol dire il 10%, cioè 98 avranno un test positivo (falsi negativi) e 882 avranno il test negativo.

In totale quindi se applico il test alle 1.000 persone prese dalla popolazione generale avremo 116 (18 + 98) test positivi, ma di questi solo 18 saranno veri malati (vale a dire solo il 15%). In altri termini se applico il test indiscriminatamente a tutta la popolazione avrò che ogni 100 test positivi solo 15 sono veri malati, gli altri sono falsi positivi.

Questa probabilità del test di individuare i veri malati si chiama "valore predittivo positivo" del test (VPP) ed è quello che realmente bisognerebbe conoscere.

E' intuitivo che se applico il test non alla popolazione generale ma solo a soggetti con sintomatologia indicativa della malattia X la prevalenza aumenta e di conseguenza aumenta anche il VPP. Se poniamo per esempio che la prevalenza della malattia in soggetti che hanno la triade congiuntivite + cefalea + febbre sia del 50% otteniamo un VPP del 90%.

Ovviamente è impensabile che il medico conosca sensibilità, specificità e VPP di ogni test, ma la lezione è sempre la stessa: gli esami vanno richiesti su sospetto clinico e non a caso!