Qualche ulteriore considerazione sul significato dei risultati quando i test statistici applicati sono inadeguati

Giuseppe Cardillo
MeriGen S.r.l., Napoli

Caro Editore,

ho letto con molto interesse il recente articolo di Vidali e Bellomo pubblicato su *Biochimica Clinica* (1). Dalla sua lettura si ricava che non tutta la nostra comunità scientifica ha familiarità con il corretto uso e interpretazione della statistica. In questo si deve evidentemente includere anche una quota non irrilevante di Revisori, che evidentemente non sempre sono in grado di esaminare la metodologia statistica utilizzata con la stessa cura con la quale valutano il protocollo sperimentale.

Basta andare qualche pagina oltre l'articolo citato, per imbattersi nell'articolo di Sorio et al. (2), che rappresenta un buon esempio di questa situazione. Nel suddetto lavoro, gli Autori affermano di aver utilizzato il test di Kruskal-Wallis per le variabili continue con distribuzione asimmetrica e il t-test di Student per le variabili con distribuzione gaussiana. Nelle tre Tabelle del lavoro, tuttavia, tutte le variabili continue sono presentate come media ± DS, senza far capire al Lettore quali siano distribuite in maniera gaussiana e quali no. Come anche ricordato da Vidali e Bellomo (1), in una distribuzione gaussiana il 68%, 95% e 99% dei dati si distribuisce entro 1, 2 o 3 DS intorno alla media. Di conseguenza, ad esempio, per quanto riguarda l'età dei soggetti analizzati il calcolo dell'intervallo di età sarebbe compreso tra 41 e 103 anni, cosa molto improbabile. Lo stesso varrebbe per l'indice di massa corporea (BMI), con un intervallo calcolato compreso tra 12,4 e 43,0 kg/m^2: ciò farebbe pensare all'arruolamento di soggetti gravemente anoressici da un lato e gravemente obesi dall'altro, cosa che probabilmente non corrisponde alla verità. E così di seguito (vedi il parametro frazione d'eiezione ventricolare sinistra con un intervallo calcolato tra 18% e 78%: ancora una volta è inverosimile che i pazienti della coda sinistra, con uno scompenso cardiaco così pronunciato, siano stati sottoposti ad angiografia).

Per quanto riguarda le variabili categoriche (ce ne sono ben 20!), l'unica che appare associata allo sviluppo di nefropatia da mezzo di contrasto (NMC) è un volume di mezzo di contrasto superiore a 300 mL. Sarebbe allora stato interessante sapere con quale criterio si decide quanto mezzo di contrasto iniettare, anche se in questa situazione, chiamandosi appunto NMC, è abbastanza ovvio che più mezzo di contrasto si inietta più aumentano le probabilità che un soggetto possa sviluppare questa complicanza. Per quanto riguarda i farmaci assunti da questi pazienti, in entrambi i gruppi ognuno sembrerebbe assumerne almeno 4. Tuttavia, gli Autori non sembra abbiano tenuto in considerazione l'eventuale interazione tra i vari tipi di farmaci nel causare lo sviluppo della NMC.

Riguardo alle metodologie statistiche utilizzate, il fatto di impiegare un "software" statistico molto diffuso, come ad esempio SPSS, non mette al riparo dall'utilizzo sbagliato di un particolare approccio di analisi. Ad esempio, appare evidente che il t-test non possa essere applicato ai dati presentati nella Tabella 1. D'altra parte, riguardo al test di Kruskal-Wallis, non solo non è chiaro per quali dati esso è stato utilizzato, ma il suo impiego in questo contesto di dati appare decisamente sbagliato: tale test, infatti, si utilizza in caso di esperimenti con tre o più gruppi, nei quali ciascun gruppo è composto da individui diversi. Visto che il "software" SPSS non ne consentirebbe l'uso per confrontare solo 2 gruppi, rappresentati da quello con presenza di NMC e da quello senza NMC, si può ipotizzare che il test sia stato utilizzato per confrontare la creatininemia [o la velocità di filtrazione glomerulare (GFR)] basale, a 12 ore, a 24 ore e a 48 ore nei due sottogruppi. In tal caso, il test di Kruskal-Wallis non è comunque applicabile, perché i dati sono "appaiati" (provengono cioè dagli stessi individui e non da individui differenti). In questo caso, andava utilizzata l'ANOVA per misure ripetute (se la distribuzione dei dati fosse stata gaussiana) o il test di Friedman.

Criticabile appare anche il commento ai risultati della Tabella 3: "Nessuno degli scostamenti calcolati delle variazioni biochimiche relative alla funzionalità renale presentava differenze statisticamente significative nei pazienti con NMC e con NMC persistente. Tuttavia, sia il valore assoluto che la variazione percentuale di creatinina plasmatica e di GFR tendevano ad essere più elevati a 12 ore ... nei pazienti che avrebbero sviluppato NMC persistente rispetto

Corrispondenza a: Giuseppe Cardillo, MeriGen S.r.l., Traversa Michele Pietravalle 11, 80131 Napoli. Tel. 0815465026, Fax 0812203571, giuseppe.cardillo@merigen.it

Ricevuto: 14.01.2011 Revisionato: 24.01.2011 Accettato: 25.01.2011

agli altri pazienti con NMC... " (2). A mio parere, questo potrebbe fornire la sensazione di voler forzare i dati relativamente alla validazione delle proprie ipotesi, anche quando non ve ne sia in realtà alcuna evidenza.

Vorrei concludere citando un grande cardiologo, diventato un grande divulgatore di statistica medica, Stanton Arnold Glantz: "Come può contribuire il Lettore a migliorare la situazione? Non transigendo su un'analisi statistica negligente più di quanto transigerebbe su una metodologia clinica o scientifica negligente. Scrivendo Lettere al Direttore di una rivista. Facendo domande in aula e durante riunioni e congressi. Quando qualcuno risponde che non sa da dove viene il valore di P o come sia stato ottenuto, occorre chiedergli come si può essere certi che i risultati significhino quanto dichiara. Può darsi che la risposta sia che non si può esserlo" (3).

BIBLIOGRAFIA

1. Vidali M, Bellomo G. Uso della statistica nei contributi scientifici pubblicati su *Biochimica Clinica*: problemi e proposte. Biochim Clin 2010;34:600-6.
2. Sorio A, Pighi M, Pesarini G, et al. Modificazioni precoci nella concentrazione della creatininemia sono in grado di predire la nefropatia da mezzo di contrasto e il danno renale persistente dopo angiografia. Biochim Clin 2010;34:612-7.
3. Glantz SA. Statistica per discipline biomediche. V^ ed. Milano: McGraw-Hill, 2003:436.