

Le frontiere dell'epidemiologia in continua evoluzione

Un excursus storico-concettuale sui limiti metodologici e l'incertezza dei risultati



Matteo Renzi

Dipartimento di epidemiologia
Servizio sanitario regionale del Lazio
Asl Roma 1

La "rarità" è un termine con il quale l'epidemiologia ha molto a che fare: si occupa spesso di patologie rare che colpiscono solo una piccolissima parte della popolazione generale, dove piccolo è il numero dei casi osservati di una determinata malattia e delle frazioni di esposti, come quando, per esempio, si valuta l'effetto di elevati livelli di un determinato inquinante. Ma come si approccia l'epidemiologia alla rarità? La risposta a questa domanda richiede un dovuto riferimento storico.

Negli anni cinquanta, l'epidemiologia si trovava in una fase ascendente e conquistava importanti traguardi, come la creazione dello storico Framingham heart study sulle patologie cardiovascolari, tuttora in attività. Una disciplina, dunque, in continua evoluzione che vedeva negli studi di coorte (un insieme di soggetti che condividono una caratteristica) la propria arma principale. Essenzialmente, una coorte veniva seguita nel tempo per poter "misurare" l'effetto di un'esposizione su vari esiti sanitari. Tuttavia questo procedimento richiedeva tempi lunghi e un numero abbastanza consistente di soggetti proporzionati all'esito da valutare. E allora, come approcciarsi a una malattia rara? La risposta a questa domanda è stata ottenuta rigirando il problema e cambiando punto di vista.

Invece che seguire un numero di soggetti nel tempo (con un approccio prospettico) e aspettare di poter valutare un numero sufficiente di casi, si è partiti dai casi stessi guardando indietro nel tempo. In questo modo, si sarebbe potuto aggirare il problema del tempo e della dimensione del campione. Così sono nati gli studi caso-controllo che valutano l'associazione tra un'esposizione ambientale e un esito raro confrontando i casi con dei controlli, ovvero dei soggetti paragonabili ai "casi" per diverse caratteristiche fisiche e/o sociali. Doveroso in questo contesto citare lo studio di Richard Doll e Bradford Hill, uscito sulle pagine del *British Medical Journal* nel 1950, in cui i due ricercatori inglesi hanno dimostrato un'associazione tra il fumo di sigaretta e il carcinoma del polmone tra i medici inglesi¹. Questa rimane tutt'ora una delle conquiste più importanti dell'epidemiologia moderna.



I piccoli numeri non aiutano le indagini epidemiologiche. Ma queste si rendono necessarie per l'importanza della domanda sanitaria e l'elevato allarme nella popolazione.

Il difficile caso dei cluster di leucemie infantili

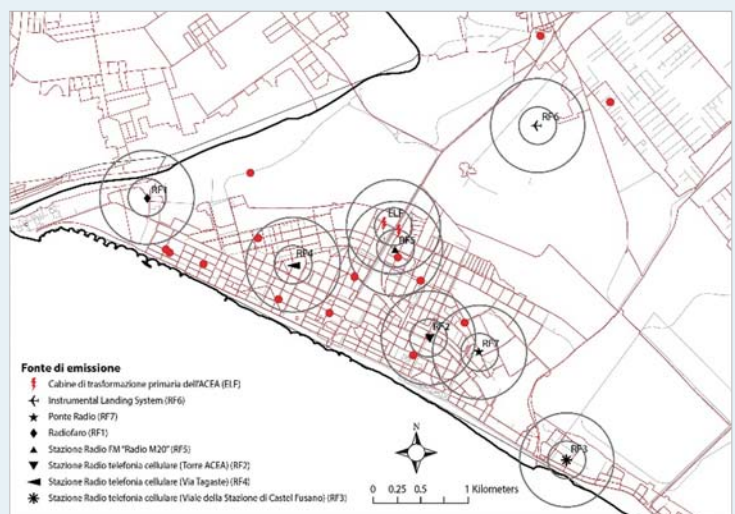
In epidemiologia ambientale il tema della rarità trova riscontro in numerosi ambiti. Per esempio, quando viene denunciato un cluster di leucemie infantili in una scuola, gli ingredienti ci sono tutti: la rarità della malattia, la rarità del numero di casi osservati, la rarità di alcuni fattori di rischio chiamati in causa. Dalla prima segnalazione di un possibile cluster di leucemia infantile agli inizi degli anni sessanta, molti studi hanno messo in evidenza la tendenza al "clustering" non solo spaziale ma anche temporale di questa patologia. Cluster di casi di leucemia infantile sono stati segnalati intorno a diverse sorgenti di inquinamento ambientale: industrie chimiche, centrali nucleari o antenne di trasmissione radiotelevisiva.

Alcuni ostacoli, come la rarità della malattia e il piccolo numero di casi in studio indagati negli studi di cluster, oltre all'assenza spesso di esposizioni comuni identificabili come causa della malattia, rendono questi studi inefficaci o dal difficile compimento. Già nel 1989 Kenneth Rothman², in un convegno ad Atlanta sul tema dei cluster (definito poi dallo stesso Rothman come il *cluster buster conference*), affermò la scarsa produttività dello studio dei cluster, sia per spiegare

un singolo evento sia per valutare la tendenza di alcune malattie ad aggregarsi nello spazio e nel tempo. Tuttavia, l'importanza della domanda sanitaria e l'elevato allarme nella popolazione rendono queste indagini spesso necessarie.

A queste criticità spesso si aggiunge anche la rarità della frazione esposta a un fattore di rischio ambientale. Per esempio, spesso si valuta il possibile ruolo dei campi elettromagnetici a bassa frequenza comunemente generati sia nell'ambiente esterno (centrali di produzione e stazioni di trasformazione dell'energia elettrica, impianti di illuminazione pubblica, cabine e quadri di distribuzione stradali) che in quello domestico (impianto elettrico, elettrodomestici e altri apparecchi a uso comune, macchine d'ufficio).

Alcuni studi, tra cui Interphone³ condotto in 13 paesi, Italia inclusa, hanno identificato un'associazione tra un'esposizione ad alti livelli di onde elettromagnetiche e sviluppo di tumori. Ragionando però sulla forza dell'associazione e su quelle che sono le armi a disposizione di un epidemiologo per avvalorare la propria ipotesi di ricerca, la rarità introduce anche in questo caso ulteriori problematiche. Infatti, aver a disposizione un numero molto piccolo di casi a p.20 →



Tumori infantili, fattori di rischio e cluster spazio-temporali.

Nello studio dell'eccesso di leucemie infantili identificato in un municipio di Roma è stata condotta un'analisi per valutare il rischio in prossimità di fonti di emissione di campi magnetici ad alta e bassa frequenza presenti nell'area. Le fonti di esposizione

identificate sono state otto: una sorgente di campi elettrici e magnetici a frequenza estremamente bassa (ELF) e sette fonti di radiofrequenze (RF). Le fonti di emissione sono state georeferenziate e sono state definite aree di 0,2 e 0,5 km da ciascuna fonte.

da p.19 → aumenta il grado di incertezza statistica e di conseguenza il p-value, cioè il valore di probabilità, che nella letteratura scientifica fornisce garanzie di robustezza: più piccolo è il valore calcolato per una serie di risultati, meno è probabile che i risultati siano dovuti esclusivamente al caso. Al di là dei piccoli numeri, il p-value è uno strumento da usare con cautela che è stato in realtà sovradimensionato. Recentemente, in un editoriale del JAMA, John Ioannidis lo ha definito come "misinterpreted, overtrusted, and misused", sottolineando come parte della comunità scientifica consigli lo spostamento del threshold per considerare i risultati "statisticamente significativi" da $p < 0,05$ a $p < 0,005^4$.

Quindi, considerata l'importanza delle ipotesi in studio e viste le problematiche che avvolgono l'utilizzo dei parametri standard per il reporting dei risultati, a volte è preferibile considerare il numero di casi attribuibili all'esposizione, ovvero tutti quei casi, tra quelli che abbiamo osservato, che possono essere imputabili all'essere esposti a un fattore di rischio ambientale. In questo modo si sarebbe in grado di dare una risposta concreta alla domanda di sanità pubblica.

Tuttavia, mentre la rarità della malattia non costituisce una criticità in questo esercizio, la rarità dell'esposizione introduce un nuovo problema. Per metterlo a fuoco torniamo all'analisi dei cluster di leucemia ed esposizione ai campi elettromagnetici a bassa frequenza: anche ammettendo la correttezza dei rischi relativi stimati nelle meta-analisi pubblicate⁵, la proporzione di casi attribuibili a questo fattore di rischio sarebbe inferiore al 5 per cento⁶, perché rarissima è la frazione

della popolazione generale esposta ai livelli elevati associati a un aumento di rischio.

Le risposte a un problema non "raro"

La maggior parte dei cluster di leucemia infantile rimane ad oggi senza una causa accertata. Se l'aggregazione spaziale chiama in causa esposizioni ambientali stabilmente presenti nella zona di residenza, l'aggregazione temporale suggerisce un'esposizione a un fattore di rischio transitoriamente presente.

Nel corso degli anni sono state ipotizzate numerose teorie per spiegare la formazione dei cluster di malattia, come quella dell'effetto di virus in aree dove si è verificato un recente mescolamento di popolazioni in precedenza isolate (teoria del population mixing)⁷: l'agente infettivo colpirebbe individui suscettibili, precedentemente non esposti al virus, in seguito per esempio a un rapido ripopolamento di un'area; ciò provocherebbe una risposta aberrante all'infezione in alcuni soggetti e la progressione verso la malattia.

Oppure, secondo l'"ipotesi igienica" (ipotesi di Greaves) una maggiore suscettibilità in alcuni individui potrebbe essere dovuta all'assenza di un'attivazione del sistema immunitario nei primissimi anni di vita, causata probabilmente dalla minore frequenza all'asilo nido, dalla minore durata dell'allattamento al seno o da un agente infettivo presente nell'ambiente⁸.

Nell'affrontare la tematica dei cluster di tumori infantili, risulta doveroso il riferimento alla monografia "Tumori infantili, fattori di rischio e modelli di indagine per la valutazione di cluster spazio-temporali", pubblicata nel 2016 sulla rivista *Epidemiolo-*

gia e Prevenzione, che ha esplorato dettagliatamente questa relazione, a cui si rimanda per ulteriori informazioni⁹.

In qualsiasi caso, le frontiere della ricerca epidemiologica sono in continua evoluzione, lo sviluppo delle analisi delle omiche (exposoma, metaboloma, ecc.) forse ci aiuterà a identificare nuovi fattori di rischio ambientali oppure a rafforzare le ipotesi relative a quelli già conosciuti, poiché nonostante colpiscono una bassa quota di popolazione, rimangono un problema non "raro" di salute pubblica. ▶

1. Doll R, Hill AB. Smoking and carcinoma of the lung. *Br Med J* 1950;2:739.
2. Rothman KJ. A sobering start for the cluster busters' conference. *Am J Epidemiol* 1990;132 Suppl:S6-13.
3. Cardis E, Richardson L, Deltour I, et al. The Interphone study: design, epidemiological methods, and description of the study population. *Eur J Epidemiol* 2007;22:647-64.
4. Ioannidis JPA. The proposal to lower p value thresholds to .005. *JAMA* 2018;319:1429-30.
5. Zhang Y, Lai J, Ruan G, et al. Meta-analysis of extremely low frequency electromagnetic fields and cancer risk: a pooled analysis of epidemiologic studies. *Environment International* 2016; 88:36-43.
6. Teepen JC, van Dijk JA. Impact of high electromagnetic field levels on childhood leukemia incidence. *Int J Cancer* 2012;131:769-78.
7. Law GR, Parslow RC, Roman E; United Kingdom Childhood cancer study investigators. Childhood cancer and population mixing. *Am J Epidemiol* 2003;158:328-36.
8. Greaves M, Buffer PA. Infections in early life and risk of childhood ALL. *Br J Cancer* 2009;100:863.
9. Michelozzi P, Schifano P. Tumori infantili, fattori di rischio e modelli di indagine per la valutazione di cluster spazio-temporali. *Epidemiol Prev* 2016; supplemento settembre-ottobre.

VEDI ANCHE

Cos'è e (soprattutto) cosa non è il p-value?

Il **p-value**, o valore della probabilità, è una misura statistica introdotta dal biologo Ronald Foshier usata per testare (e respingere) la cosiddetta "ipotesi nulla" secondo cui i gruppi confrontati sono fra loro uguali e la differenza osservata riguardo al parametro considerato va attribuita al caso. Questo numero ci dice quanto è probabile ottenere una differenza tra due gruppi pari o superiore a quella osservata. Più piccolo è il valore di p calcolato per una serie di risultati, meno è probabile che i risultati osservati siano dovuti solo al caso. Fisher aveva fissato arbitrariamente come 0,05 la soglia al di sotto della quale le differenze osservate tra i gruppi sono da considerarsi

"statisticamente significative". Questo parametro viene ampiamente adottato nella ricerca, soprattutto quella biomedica, per affermare che un risultato è statisticamente significativo, come per esempio che il farmaco studiato abbassi la pressione arteriosa oppure che quel determinato agente ambientale aumenti il rischio di cancro. Ma negli ultimi anni il p-value è stato messo sotto accusa. Alcuni ricercatori considerano che la soglia di 0,05 abbia prodotto troppi falsi positivi e propongono di abbassarla sotto lo 0,005 – il che però richiederebbe dimensioni di campioni molto più ampie. La questione di fondo

è che il valore p viene erroneamente interpretato e che l'attendibilità dello studio dipende molto dal tipo di studio, dal numero di dati e da quanto è plausibile l'ipotesi di partenza. La dicitura "statisticamente significativa", argomentano alcuni ricercatori, andrebbe abolita. Wikipedia elenca diversi errori comuni riguardanti l'uso del p-value.

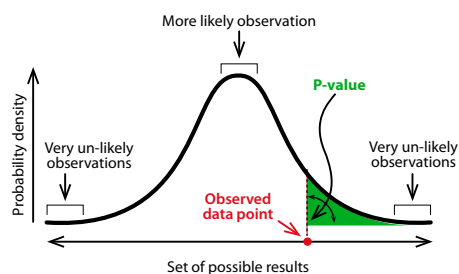
- Il p-value **non** è la probabilità che l'ipotesi nulla sia vera o la probabilità che l'ipotesi

nulla sia falsa. Non è connessa con nessuna delle due.

- Il p-value **non** è la probabilità che un'osservazione sia un caso. Il calcolo del p-value è basato sull'ipotesi che ogni osservazione è un caso, un risultato aleatorio. Con la frase "il risultato è dovuto al caso" si intende di solito che l'ipotesi nulla è probabilmente corretta. Ma il p-value non può essere usato per rappresentare la probabilità che

- un'ipotesi sia vera.
- Il p-value **non** è la probabilità di rifiutare l'ipotesi nulla quando questa è vera.
- Il p-value **non** è la probabilità che replicando l'esperimento si otterrebbe la stessa conclusione. Il livello di significatività α **non** è determinato dal p-value. Il livello di significatività è deciso dalla persona che conduce l'esperimento prima di vedere i dati. •

Fonti. Wikipedia, Nature.



Distribuzione della probabilità di ottenere per solo effetto del caso i possibili risultati. Il valore p (area verde) è la probabilità del risultato osservato assumendo vera l'ipotesi nulla.