

ABC per orientarsi nei dati epidemiologici

Pio Russo Krauss

Come si fa ad affermare che un fattore è in causa nel determinare, prevenire o curare una determinata patologia?

Le osservazioni personali, per quanto possano essere scrupolose, sono soggette a variabilità più o meno notevole e sono influenzate da innumerevoli fattori, molti dei quali ci sfuggono, pertanto **possono trarci in inganno e portarci ad interpretazioni errate.**

Per esempio, terapie ritenute efficaci non solo dai pazienti ma talvolta anche dai medici (gli epatoprotettori, i neurotrofici, i ricostituenti ecc.) ad un esame oggettivo si sono dimostrati privi di efficacia terapeutica. Supposti fattori causali, ad un'indagine rigorosa, si sono dimostrati del tutto "innocenti" o molto meno pericolosi di quanto si riteneva. Di contro, fattori che si ritenevano innocui o poco pericolosi si sono rilevati dannosi o molto pericolosi.

Per affermare che "qualcosa" "fa bene" o "fa male" non basta osservare che un numero anche elevato di persone è migliorato dopo aver assunto quel "qualcosa", perché avrebbe potuto migliorare anche senza l'assunzione di quel "qualcosa" e forse persino più rapidamente o in modo più completo. Né si può affermare che qualcosa è dannoso perché un certo numero di persone ha subito danni dopo essere stati esposti a quel "qualcosa". Molte ricerche hanno dimostrato che sia i medici che i pazienti si fanno influenzare dalle loro aspettative: se "credono" nei farmaci notano i miglioramenti e disconoscono i fallimenti; se sono convinti che qualcosa fa male mettono in relazione quel fattore con fenomeni negativi senza sufficienti elementi. Addirittura la suggestione può provocare reali effetti positivi (effetto placebo) o dannosi (effetto nocebo). Per tali motivi si deve **diffidare di chi porta come prova dell'efficacia o dell'inutilità o della pericolosità di determinati fattori la propria esperienza personale** o una casistica, perché ciò indica ignoranza del metodo scientifico, superficialità e faciloneria. E' necessario invece darsi un metodo che sia capace di tener conto di questi problemi e di superarli, che ci aiuti a discernere ciò che è vero da ciò che sembra vero.

La prima cosa su cui bisogna intendersi sono i **concetti che devono essere precisi e inequivocabili.** Per tale motivo nel linguaggio scientifico si cerca di esprimere i concetti operativamente (per esempio, viene definito depresso colui che viene diagnosticato tale da parte di due psichiatri o che risulta positivo al test sulla depressione; è definito "guarito dal cancro" colui che dopo 5 anni dall'ultima terapia non presenta nessun sintomo e alle indagini diagnostiche non ha alcun segno) e possibilmente in maniera quantitativa (un "bevitore" è colui che assume oltre 100 grammi di alcol al giorno; "esposti all'inquinamento atmosferico" sono coloro che per almeno tot giorni l'anno sono esposti a livelli di NO₂ superiori a y o di polveri fini superiori a z). D'altra parte anche nella vita di tutti i giorni non ci si può accontentare di concetti vaghi come lontano e vicino, grande e piccolo, pesante e leggero, oscuro e luminoso, ma si stabiliscono delle procedure che trasformano apprezzamenti qualitativi in valori quantitativi precisi (un Km, un metro, 100 cc, 2 Kg, 100 g, 50 candele ecc.); a maggior ragione questo è vero per le discipline scientifiche. Il metodo, applicabile alla variabilità biologica ed umana, che permette di discernere ciò che è vero da ciò che sembra vero e di esprimere in termini inequivocabili e precisi le nostre osservazioni è il metodo statistico. Il metodo statistico, tramite i test di significatività, permette di misurare la probabilità che abbiamo di essere nel vero quando asseriamo che c'è un'associazione tra una variabile e un fenomeno.

La disciplina che studia la distribuzione delle malattie nelle popolazioni, allo scopo di individuare le cause, il ritmo e l'intensità con cui queste si manifestano e le condizioni che le ostacolano o le favoriscono è l'epidemiologia.

Gli studi epidemiologici

La conoscenza della metodologia epidemiologica e del metodo statistico ci permette di valutare criticamente l'informazione sanitaria che riceviamo, di relativizzare le nostre e altrui osservazioni (le nostre osservazioni possono ingannarci, quel che ci sembra evidente può non esserlo per un altro osservatore o può non esserlo in realtà), di rilevare alcuni fenomeni in maniera rigorosa, in modo tale che tutti debbano convenirne e che quanto affermiamo abbia riscontro nella realtà.

L'epidemiologia utilizza i seguenti modelli di ricerca, detti studi epidemiologici:

Studio trasversale (geografico)

In questo studio si calcola la prevalenza di una determinata patologia in una determinata popolazione, cioè **il numero di persone che presenta una data malattia o carattere** (per esempio il numero di malati di cancro negli abitanti in un determinato territorio). Tale tipologia di studio **non serve a identificare le cause delle malattie** ma ad avere dati per prendere decisioni di sanità pubblica. Per esempio, se in determinato territorio vi sono molti diabetici bisognerà incrementare e/o migliorare i programmi e le strutture di prevenzione e cura del diabete.

Gli studi trasversali possono anche essere comparativi: si possono contare quante persone muoiono di tumore in un determinato territorio e quante in un altro. **Tali studi permettono di avanzare un sospetto di relazione tra un determinato fattore e una patologia.** Per esempio se il primo territorio presenta un maggiore inquinamento del secondo e la mortalità per tumore è più alta nel primo territorio si può sospettare che tale aumento di mortalità dipenda da ciò. **“Sospettare” non significa “avere le prove”.** Le differenze osservate tra le due popolazioni potrebbero dipendere da altri fattori che differenziano le popolazioni delle due città (stili alimentari, abitudine al fumo, reddito, occupazione lavorativa, servizi sanitari, genetica, ecc.). Lo studio geografico non permette di tenere sotto controllo le innumerevoli variabili che possono essere in gioco. La storia dell'epidemiologia non manca di esempi di sospetti che poi si sono dimostrati infondati e di speranze che poi sono state deluse.

Studio delle serie storiche

In questo studio si esamina il numero di morti o di malati in una determinata area, lungo un certo periodo di tempo (solitamente vari anni o decenni) e qualora esistano aumenti o diminuzioni della mortalità o del numero di malati si vede se tali aumenti corrispondono a variazioni di alcuni fattori. Per esempio si esamina il numero di ricoverati per asma nei vari giorni dell'anno e si vede se gli aumenti e le diminuzioni dei ricoveri per asma vanno di pari passo (correlano) con il grado di inquinamento atmosferico.

Questo studio **permette di stimare quanto è forte la relazione tra il fattore e la patologia in esame, nonché il numero di morti o malati correlati al fattore in esame.** Ovviamente perché si possa prospettare un rapporto di causa-effetto devono esserci prove per una sua plausibilità biologica. Per esempio lo studio delle serie storiche dimostra che c'è una correlazione tra cicogne e numero di nascite nei paesi scandinavi, infatti v'è una parallela diminuzione, ma questo non dimostra che i bambini li portano le cicogne.

Studio caso-controllo (studio retrospettivo)

Nel caso-controllo si analizza, in un gruppo di malati e in uno di sani, quanti nell'uno e nell'altro gruppo sono stati esposti al fattore o ai fattori sospettati di favorire o contrastare una determinata malattia. Praticamente si intervisteranno un certo numero di malati (casi) chiedendo loro età, residenza, lavoro, occupazione, abitudini di vita ecc. e per ogni malato si intervisteranno una o più persone sane (controlli) chiedendo loro le medesime cose (età, residenza ecc.). In questa maniera si possono considerare tutti quei fattori che sfuggono allo studio trasversale. Per cui se si evidenzia che il 99% dei soggetti con arterite obliterante fumava, mentre nei sani i fumatori sono il 20% e che non vi sono significative differenze nella distribuzione degli altri fattori, si può affermare che tale studio dimostra che il fumo di sigaretta è fortemente implicato nella genesi dell'arterite obliterante.

Lo studio caso-controllo, quindi, **permette di dimostrare la relazione tra un fattore e una patologia** (anche rara) **e di calcolare una stima di quanto è forte questa relazione** (Odds ratio).

Studio di coorte (studio prospettico o longitudinale)

Nello studio di coorte vengono seguiti per molti anni due gruppi in tutto simili, tranne che per l'esposizione al fattore (per esempio l'esposizione a radiazioni), rilevando in quanti soggetti dell'uno e dell'altro gruppo accadranno gli eventi considerati (per esempio tumori, sterilità, nascita di bambini con malattie genetiche ecc.).

Praticamente i soggetti in esame vengono intervistati, visitati, studiati con analisi di laboratorio o strumentali sia all'inizio della ricerca sia periodicamente, in modo da controllare i possibili fattori confondenti e da evidenziare le patologie che sviluppano.

Lo studio di coorte permette di dimostrare in maniera rigorosa la relazione tra un fattore e una patologia, di calcolare in maniera molto affidabile quanto è forte questa relazione (“rischio relativo”, per esempio chi fuma 20 sigarette al giorno ha 10 volte più probabilità di avere un tumore al polmone e 15 volte più probabilità di sviluppare bronchite cronica) e, conoscendo la diffusione del fattore in esame (per esempio il numero di fumatori) anche quale è la quota della patologia in esame che è dovuta al fattore implicato (rischio attribuibile, per esempio l'80% dei tumori polmonari e il 75% delle bronchiti croniche sono attribuibili al fumo in Italia).

Lo studio di coorte, quindi, permette di dimostrare la relazione tra un fattore e una patologia e di calcolare quanto è forte questa relazione (rischio relativo). Esso ha due grandi difetti: costa molto e non dà risultati prima di molti anni (anche 15-20 anni).

Studio sperimentale controllato

E' un studio simile a quello di coorte ma dove le variabili, che si sospettano possano trarre in inganno (fattori confondenti), sono poste sotto un controllo ancora più rigoroso.

Praticamente si prendono due gruppi di individui simili tra loro, li si intervista, visita ecc. e si espone un gruppo ad un fattore che si pensa sia protettivo (es. un alimento) o curativo (es. un farmaco) e l'altro gruppo ad una sostanza simile ma priva di azione sull'organismo (placebo) in modo che i soggetti sottoposti allo studio non possano sapere se stanno prendendo il placebo o la sostanza in studio (**studio in cieco**). Se anche i ricercatori ignorano a quali soggetti viene data la sostanza in studio e a quali il placebo si parla di **studio in doppio cieco**.

Questo studio permette di affermare con certezza il nesso causale positivo o negativo tra variabile in studio (es. farmaci, alimenti) e patologia.

Le sostanze sospette di essere nocive, ovviamente, non possono essere sperimentate sull'uomo. In questi casi lo studio viene eseguito su culture cellulari o animali di laboratorio.

Per esempio si espongono animali o sistemi biologici (es. colonie batteriche o culture cellulari) ad un inquinante e altri animali o sistemi biologici, con medesime caratteristiche, alle medesime condizioni sperimentali tranne che per l'esposizione all'inquinante in studio. Gli animali sono studiati con vari procedimenti (analisi di laboratorio, broncoscopia, biopsia, autopsia ecc.). In taluni studi, per neutralizzare possibili distorsioni legate all'osservatore, gli stessi ricercatori che osservano gli eventuali effetti non sanno quali animali e sistemi biologici sono esposti all'inquinamento e quali no (studio in cieco).

Questo studio permette di affermare con certezza il nesso causale tra variabile in studio (es. esposizione al benzene) e fenomeni osservati in quel sistema biologico (es. batteri o cavie). La trasposizione di questi risultati all'uomo viene operata perché risponde ad un'alta probabilità statistica (sono rari i casi nei quali un fattore dimostrato dannoso per i sistemi biologici si dimostri non dannoso per l'uomo).

Come capire il valore delle notizie su fatti riguardanti la salute

Sui giornali, in televisione e sulla rete spesso troviamo notizie di studi che dimostrerebbero in maniera evidente determinate cose, e notizie di altri studi che dimostrano l'esatto contrario: come capirci qualcosa?

Il problema è che i giornalisti solitamente non hanno alcuna competenza in epidemiologia. Inoltre sparare una notizia tipo "Trovata la cura del cancro" o "Epidemia di tumori in ..." fa vendere.

Il criterio per stabilire se una "notizia medica" è vera e falsa non dovrebbe essere se è in linea con le nostre opinioni o in base a criteri semplici, tipo: "tutto quello che è di fonte governativa è falso" (oppure vero), "tutto quello che viene da ricercatori indipendenti è falso" (oppure vero). Bisogna **farsi interrogare dalla realtà, che è complessa e non può essere divisa in bianco e nero.**

Allora bisogna cercare di non avere pregiudizi ed esercitare la ragione. Ecco un piccolo **promemoria per orientarsi tra questo tipo di notizie:**

- 1) **la notizia fa riferimento a uno studio o è solo l'opinione di qualcuno?** Purtroppo nella nostra società non c'è più una distinzione netta tra opinioni e dati e si sparano dati come se niente fosse (vi ricordate "L'inceneritore inquina come 3 auto"?). I dati o sono veri o sono falsi e vanno dimostrati/confutati con i fatti (non vige il principio di autorità, l'ipse dixit). Se sono veri non si possono non prendere in considerazione e se sono falsi non devono essere presi in considerazione. Le opinioni possono essere condivise in toto, in parte o per niente perché ognuno è libero di avere le sue opinioni (mentre è da stupidi negare la realtà dei fatti);
- 2) **quale tipo di studio è stato compiuto** (trasversale, serie storiche, caso controllo, coorte, sperimentale)? Avere questo dato è importante per tutto quanto abbiamo detto prima;
- 3) **lo studio è stato pubblicato su una rivista scientifica?** Se non viene citato su quale rivista è stato pubblicato è possibile che non sia stato pubblicato su nessuna rivista scientifica (lo si può verificare su internet) e ciò può essere un segno importante di studio male impostato o non condotto bene o di uso di dati per fini particolaristici (fosse anche il piacere di avere un articolo sui giornali) o di spacciare le proprie opinioni per dati;
- 4) **su quale rivista scientifica è stato pubblicato?** Se lo studio ha degli errori metodologici, sarà rifiutato dalle riviste scientifiche autorevoli e potrà essere pubblicato solo da riviste di enti sanitari (tipo il bollettino dell'ospedale di ...) o di associazioni poco attente al rigore scientifico. Riviste autorevoli sono Lancet, British Medical Journal, JAMA, Epidemiologia e Prevenzione e, in genere, quelle presenti in PubMed www.ncbi.nlm.nih.gov/pubmed che raccoglie articoli pubblicati in tutto il mondo);
- 5) **da chi è stato finanziato lo studio?** Una ricerca finanziata dalla Marlboro è di nessuna attendibilità se dimostra che il fumo di sigaretta non fa tanto male. Infatti le industrie usano vari metodi per orientare il risultato dello studio. Il più furbo è quello di commissionare numerosi studi sul medesimo argomento e poi pubblicare solo quello che ha i dati che si desiderano (ad esempio si commissionano 10 studi sui supposti benefici di una sostanza, poi succede che 8 studi dimostrano che quella sostanza non serve a niente, uno che peggiora la situazione e uno che la migliora e viene pubblicato solo quest'ultimo studio);
- 6) **lo studio si basa su un campione rappresentativo** (cioè su un numero congruo di persone, individuate con metodo casuale dalla popolazione in esame) **o su una casistica** (cioè su un certo numero di casi non estratti a caso dalla popolazione in esame)? Se si basa su un campione i risultati dello studio si possono considerare validi per l'intera popolazione in esame; se si basa su una casistica no, perché questa non è rappresentativa della popolazione;
- 7) **la notizia riporta gli "intervalli di confidenza"?** L'intervallo di confidenza ("IC" oppure "IC al 95%") è rappresentato da due numeri solitamente riportati tra parentesi vicino al numero che indica l'aumento o la diminuzione di probabilità. Per esempio: "Aumento del 60% dei tumori (IC 30-90%)". Se l'intervallo tra questi due numeri comprende lo zero significa che quel dato non è significativo. L'intervallo di confidenza può essere espresso anche non in percentuale (nell'esempio di prima 1,3-1,9) e in questo caso dà indicazioni di non significatività se

comprende il numero 1 (quindi l'aumento del 60% dei tumori dell'esempio è un dato significativo);

- 8) **la notizia riporta il valore di significatività statistica (p)?** Valori di p superiori a 0,05 ($p > 0,05$) sono in genere considerati non significativi, **valori inferiori a 0,05 ($p < 0,05$) sono considerati significativi** (altamente significativi se inferiori a 0,005).

Epidemiologia ed etica

Come abbiamo visto, spesso per avere certezze dagli studi epidemiologici occorrono anni e nel frattempo bisogna prendere decisioni (anche aspettare è una decisione).

Spesso non c'è bisogno dei dati epidemiologici per prendere decisioni. Per esempio, perché mai bisognerebbe aspettare di avere dati epidemiologici per avviare una lotta seria al traffico illecito di rifiuti pericolosi e al fenomeno dei roghi tossici? **In tal caso gli studi epidemiologici possono essere una perdita di tempo e denaro:** meglio investire queste risorse per combattere tali fenomeni.

Se vi sono studi seri che fanno sospettare che qualcosa fa male alla salute bisogna cercare di neutralizzare quel qualcosa (**principio di precauzione**), senza aspettare decine di anni per avere i risultati degli studi di coorte. Ovviamente neutralizzare quel fattore potrebbe avere dei **costi** economici, sociali, culturali, ambientali e perfino sanitari e questi devono essere ben vagliati. Tali costi potrebbero essere certi o incerti, alti o bassi, ricadenti su alcuni soggetti o su altri. Ciò può dare luogo ad opinioni molto diverse su cosa fare e non fare e, spesso, a conflitti sociali determinati dai diversi interessi. In tale situazione frequentemente succede che si tende a far passare per dati le opinioni e per opinioni i dati, per dare credito ai dati sulla base della loro fonte e non della loro qualità e attendibilità scientifica, per enfatizzare dati poco rilevanti e minimizzare dati rilevanti e perfino a diffondere dati falsi. Tutto ciò è eticamente riprovevole e i cittadini dovrebbero levare la loro fiducia a chi si comporta così. Insomma bisognerebbe sempre **farsi guidare da razionalità ed etica.**

Un'altra questione è quella della finitezza delle risorse e delle scelte che ne conseguono. **Ogni cosa ha un costo e le risorse non sono infinite**, per cui se si impegnano per una cosa saranno di meno per un'altra. Dovrebbero essere privilegiati quei provvedimenti che hanno un alto **rapporto utilità/costi** (ad esempio anni di vita in buona salute guadagnati/euro spesi), e ciò presuppone appunto un forte esercizio etico e della ragione (farsi guidare non da pregiudizi, preconcetti, emozioni ecc. ma da analisi rigorose e dal perseguimento del bene comune).

Alcune parole chiave

Incidenza: è il numero di nuovi casi di un fenomeno (p. es. una malattia) in una popolazione in un determinato intervallo di tempo. Per esempio, l'incidenza dei tumori in Italia è nei maschi di 480 ogni 100.000 abitanti all'anno e nelle donne di 440 ogni 100.000 abitanti.

Prevalenza: è il numero di persone di una popolazione che presenta una dato carattere (p. es. una malattia) in un dato momento. Per esempio, la prevalenza dei tumori in Italia attualmente è di 2.250.000. La prevalenza di una malattia è direttamente proporzionale alla sua incidenza e alla sua durata.

Mortalità: è la frequenza dell'evento morte in una popolazione. In Italia, per esempio, il tasso di mortalità per i tumori è di 220/100.000/anno, cioè ogni anno muoiono per tumori 220 persone ogni 100.000 abitanti.

Morbosità o morbilità: è la frequenza dell'evento malattia in una popolazione. Generalmente viene espressa tramite il tasso di prevalenza oppure tramite il tasso di incidenza.

Registro di mortalità (delle cause di morte): quando una persona muore il medico compila, oltre al certificato di morte, un modello ISTAT dove deve riportare vari dati della persona (età, sesso, residenza, occupazione, titolo di studio ecc.) nonché la causa iniziale e le cause intermedie e finali che hanno determinato la morte (secondo una classificazione internazionale delle cause di morte. Tali dati sono registrati su tutto il territorio nazionale per tutte le persone decedute.

Registro di patologia (es. registro tumori): i registri di patologia registrano tutti i casi di una determinata patologia (es. tumori) che avvengono nella popolazione residente su un determinato territorio (dovunque diagnosticati), nonché tipologia e stadio della malattia, terapie effettuate, evoluzione (guarigione, morte) e altri dati (età, sesso, indirizzo ecc.).

Rischio: è la probabilità che un evento negativo si verifichi in un dato intervallo di tempo.

Rischio relativo (rr): indica l'incremento del rischio dovuto all'esposizione ad un fattore. Per esempio, riguardo al tumore del polmone, un rischio relativo di 10 per i fumatori significa che i fumatori hanno un rischio 10 volte superiore ai non fumatori di contrarre un tumore del polmone. Ha importanza, quindi, per stimare il rischio per il singolo individuo.

Rischio attribuibile nella popolazione: descrive la quota di un evento presente in una popolazione (p.es. la malattia o la morte) dovuta ad un fattore di rischio. Per esempio, riguardo al tumore del polmone, un rischio attribuibile di 80% per il fumo significa che l'80% dei tumori polmonari sono dovuti al fumo di sigaretta. Ha importanza, quindi, non per il singolo individuo, ma per chi si interessa dello stato di salute della popolazione: è infatti il dato che orienta gli interventi di sanità pubblica.

Test di significatività sono procedure con le quali si stima la probabilità che le differenze riscontrate rientrino nel range di fluttuazioni casuali e quindi non siano significative. Si esprime con la formula $p < x$. Per esempio $p < 0,01$ significa che vi sono meno di una probabilità su 100 che la differenza riscontrata dipenda da una fluttuazione casuale. Valori di p inferiori a 0,05 sono considerati significativi, altamente significativi se inferiori a 0,005.

Ovviamente, prima di dichiarare l'esistenza di un rapporto causa-effetto tra l'esposizione e la malattia, devono essere verificati i criteri di causalità (successione temporale tra causa ed evento, plausibilità ecc.).

Intervallo di confidenza (IC)

L' "Intervallo di confidenza al 95%" (I.C. 95%) è un test di significatività. L'intervallo di confidenza che comprende il valore 1 (tra 0 e infinito) indica mancanza di significatività; se non comprende il numero 1 (tra 1 e infinito oppure tra 0 e 1) indica significatività. Se i dati sono espressi in percentuale la mancanza di significatività si ha quando l'intervallo comprende il numero zero (è tra -100% e +infinito%) ed è significativo quando non comprende lo 0 (è tra 0 e +infinito oppure tra 0 e -100).

Ad esempio, se una ricerca evidenzia che l'essere esposti ad un determinato inquinamento raddoppia la probabilità di avere un tumore (rischio relativo 2, cioè aumento del 100%) con intervallo di confidenza 1,4-2,4 (40-140%) questo dato (il raddoppio della probabilità di avere un tumore) è significativo (cioè è poco probabile che dipenda dal caso ed è molto probabile che esista un nesso tra quell'inquinamento e i tumori). Ma se la ricerca avesse evidenziato un raddoppio della probabilità di avere un tumore ma con IC 0,5-2,8 (-50% - +180%) il medesimo dato (raddoppio della probabilità di avere un tumore) non avrebbe significatività.

L'intervallo di confidenza al 95% ci dice che, se ripeto n volte lo stesso studio, ho il 95% di probabilità di trovare, nel primo caso, valori compresi tra 1,4 e 2,4 (cioè un aumento della probabilità di avere il tumore tra il 40% e il 140%) e, nel secondo caso, valori tra 0,5 e 2,8 (cioè da una diminuzione del 50% della possibilità di avere un tumore ad un aumento del 180% della probabilità di svilupparlo). Ecco perché nel secondo caso quel dato non è significativo.