

**Simonetta Pagliani**

pagliani@zadig.it

### **Incertezza e variabilità in medicina**

**“La medicina è la scienza dell’incertezza e l’arte della probabilità.** Come ha scritto William Osler, un illustre clinico della seconda metà del 1800, la storia della medicina è, infatti, il racconto di certezze sempre in divenire, di acquisizioni che sembravano definitive e che vengono messe in soffitta, di verità che si ritenevano provate e che si tramutano in errori.

Si possono raffigurare quattro aree in cui si incrociano in vario modo la certezza e la stabilità della conoscenza medica:

- certezza e stabilità: alcune conoscenze sono provate e verificate senza smentita (esempio: il cortisone è un potente antinfiammatorio);
- certezza e precarietà: conoscenze che sembravano assodate vengono superate da studi ulteriori (esempio: sembrava che la terapia sostitutiva ormonale proteggesse le donne dalle patologie cardiovascolari, ma è stato smentito da studi successivi);
- incertezza e stabilità: credenze false, ma dure a morire (l’alcol aiuta a digerire);
- incertezza e precarietà: i miti della medicina (i ricostituenti, gli epatoprotettori), la sfera emotiva del paziente (lo stress provoca il cancro)

Il modo di cercare, trovare, classificare e, infine, curare le malattie (in una parola, la pratica clinica) ha sempre variato da terapeuta a terapeuta, da luogo a luogo, da epoca a epoca.

Anche limitandosi all’ambito occidentale, dove essa appartiene al gruppo delle professioni fondate su rigorose conoscenze scientifiche, la medicina si rivela spesso una pratica proteiforme, perché viene esercitata in un contesto singolare e individuale di esseri umani, le cui necessità sono sempre un po’ in contraddizione con quelle collettive e sociali e le cui scelte hanno sempre un qualche grado di irrazionalità.

La medicina, inoltre, deve interpretare patologie che si presentano nei diversi individui con quadri di volta in volta modificati da genetica e ambiente, da natura e cultura.

Infine, la stessa soggettività del medico lo pone fuori dalle quelle condizioni di arbitro obiettivo degli eventi, che sono il presupposto di ogni ricerca scientifica. La sua relazione con il paziente diventa proprio l’elemento determinante della variabilità del giudizio e dell’azione clinica: il curante e il curato, i colleghi di altre discipline eventualmente chiamati in causa, nonché i mezzi d’informazione, hanno intendimenti diversi e a volte tra loro inconciliabili.

Da ultimo, ma rilevante, si aggiunge il carico delle strategie politiche ed economiche del terzo pagante (il Servizio sanitario pubblico, dove esiste, oppure il sistema assicurativo privato).

### **La variabilità nella diagnosi**

La natura della medicina è fondamentalmente probabilistica, come ben sapeva Jerome P Kassirer, che ha diretto per molti anni la rivista *New England Journal of Medicine*, quando diceva: “la diagnosi è un bersaglio mobile; ogni volta che si raccoglie un nuovo elemento diagnostico, aumentano le probabilità di una certa malattia e diminuiscono quelle di un’altra.”

Capita anche che alla diagnosi di certezza non si arrivi proprio (succede per il 50% dei dolori toracici e per il 20% di quelli addominali), anche a dispetto di ricerche lunghe e costose, che spesso vengono portate avanti solo per evitare ritorsioni legali da parte del paziente (medicina difensiva) oppure rimorsi di coscienza per non aver fatto abbastanza per lui (regret bias). Una delle decisioni più difficili per il medico è proprio stabilire quando è stato fatto abbastanza: il confine che ciascun medico pone a questo “abbastanza” è uno dei fattori di variabilità nella prescrizione sia di esami diagnostici, sia di terapie .

Per fare la diagnosi giusta, non basta chiedere il test diagnostico appropriato: raramente gli esami diagnostici danno risposte incontrovertibili al quesito clinico e il risultato dell’esame (positivo o negativo che sia) può essere veritiero o falso. I test, in misura variabile, presentano limiti sia nella capacità di riconoscere tutti i malati sia in quella di individuare tutti i sani. Sensibilità e specificità di un test sono caratteristiche intrinseche che ne descrivono la capacità diagnostica e sono dati che vengono di solito forniti dai produttori dei test, o si possono reperire in letteratura. La sensibilità di un test corrisponde alla proporzione di soggetti realmente ammalati identificati come tali dal test.

La specificità di un test corrisponde alla proporzione di soggetti realmente sani che sono identificati come tali dal test.

Nella pratica quotidiana, però, la domanda che il medico si pone non è: "se una persona è malata quanto è probabile che il test risulti positivo" oppure "se una persona è sana quanto è probabile che il test risulti negativo". La domanda è ribaltata: "se un test risulta positivo, quanto è probabile che il paziente abbia davvero la malattia? e se risulta negativo, quanto è probabile che sia veramente sano? Di un test, occorre sapere qual è la sua capacità di segnalare una malattia realmente esistente (il suo valore predittivo positivo) e anche qual è il suo valore predittivo negativo, cioè la capacità di dare un risultato negativo in una situazione realmente normale.

$$VPP = (\text{sensibilità} \times \text{prevalenza}) / (\text{sensibilità} \times \text{prevalenza}) + (1 - \text{specificità}) \times (1 - \text{prevalenza})$$

$$VPN = \text{specificità} \times (1 - \text{prevalenza}) / (1 - \text{sensibilità}) \times \text{prevalenza} + \text{specificità} \times (1 - \text{prevalenza})$$

Questi due valori dipendono da specificità e sensibilità, ma non in modo costante in qualunque circostanza: dipendono anche da quanto la malattia che il test indaga è frequente nella popolazione che si considera (prevalenza). Quanto maggiore è la prevalenza della malattia, tanto è maggiore la probabilità di trovare davvero la malattia quando il test è positivo.

Per esempio, se si fa un ecodoppler della carotide, alla ricerca di un restringimento patologico del vaso, ai presenti in quest'aula, il valore predittivo positivo del test non sarà superiore al 44% (ossia vi saranno 6 risultati falsi positivi ogni 10), mentre se lo si fa in un reparto neurologico ai ricoverati per ictus, il valore predittivo può arrivare al 98%.

Il fatto che il valore predittivo positivo dipenda dalla prevalenza è uno dei motivi che sconsiglia l'effettuazione di screening per malattie rare. Infatti, uno screening per una malattia rara presenta i seguenti inconvenienti:

- pochi individui ne trarranno beneficio (proprio in quanto malattia rara);
- molti individui (i falsi-positivi) ne trarranno un danno, in quanto verranno ingiustamente considerati ammalati.

### **La variabilità biologica di presentazione dei segni e dei sintomi**

La difficoltà di porre una diagnosi può derivare anche dalla variabilità di presentazione dei segni, dovuta, innanzi tutto, a ragioni biologiche.

Parametri quali altezza, peso, pressione arteriosa e frequenza cardiaca possono variare di giorno in giorno e di ora in ora, in dipendenza da fattori quali posizione, dieta, assunzione di liquidi, stress, attività fisica, eccetera. La variazione biologica è la regola anche per molte misure che vengono credute "esatte", incluse quelle derivate da un ecocardiogramma, come la pressione ventricolare e la frazione di eiezione.

Variazioni biologiche paradosse possono essere effetto di farmaci o di malattie. Per esempio, in molti pazienti ipertesi la pressione arteriosa ritorna normale in seguito a un infarto miocardico. Infine, bisogna tenere conto del fenomeno conosciuto con il nome di "regressione verso la media": quando si trovano valori estremi di un certo parametro, è quasi certo che, ripetendo l'osservazione, i secondi valori saranno meno estremi e più vicini alla media del parametro nella popolazione. Per esempio, una prima misurazione della pressione arteriosa che dà un valore elevato, se ripetuta, potrebbe tendere a ritornare a valori normali.

L'età è un importante fattore biologico di variabilità di presentazione dei sintomi: questi altro non sono che una strategia difensiva dell'organismo, il cui costo è differente a seconda della fragilità dell'organismo stesso. Per esempio, il quadro clinico classico di risposta a un'infezione acuta delle vie urinarie, che in un adulto è costituito da disuria (fastidio a urinare), dolore lombare e febbre alta, in un lattante sarà sostituito da perdita di peso e anemia, spesso in assenza di febbre (il fegato sequestra il ferro per evitare che i batteri lo usino per la loro crescita); in un vecchio compariranno confusione mentale e inappetenza, per la perdita del cosiddetto "richiamo d'organo" nella sintomatologia, tipica dell'età molto avanzata (si pensi che in un terzo dei vecchi, persino l'infarto del miocardio è indolore)

## **Variano nel tempo le posizioni culturali e scientifiche**

Per misure quali la pressione arteriosa, la massa corporea, la colesterolemia o la tolleranza al glucosio, che non si esprimono con la dicotomia “positivo – negativo”, ma che hanno una distribuzione continua, l’individuazione di una soglia patologica è arbitraria; allo stesso modo lo è la definizione di ipertensione, diabete o di alto rischio cardiovascolare.

Quando il paziente chiede se i suoi esami del sangue vanno bene, la risposta che viene data dal medico oggi non è la stessa che egli avrebbe dato qualche anno fa, per via del cambiamento culturale dovuto a:

- nuovi dati scientifici (portati dai risultati di studi clinici)
- spinte del mercato (interessi industriali)
- sostenibilità economica degli interventi (concedibilità di mezzi diagnostici o di farmaci da parte del SSN)

Rispetto a dieci anni fa, i valori di pressione arteriosa, glicemia e colesterolo definiti patologici si sono progressivamente abbassati e non si può fare a meno di notare che l’ampliamento del numero dei malati o dei futuri malati che così si determina, aumenta la domanda di visite specialistiche, di strumentazioni diagnostiche e di farmaci.

Un’eccessiva medicalizzazione è, insomma, sempre in agguato come prodotto dell’estensione a tappeto dei test diagnostici. Quando un test diagnostico viene proposto con una certa periodicità a una certa popolazione di soggetti asintomatici, si chiama screening. Sulla loro proliferazione va mantenuta una vigilanza critica: mentre alcuni di essi (quelli per il carcinoma della cervice uterina, della mammella e del colon) sono utili, perché l’anticipazione diagnostica da essi indotta ha dimostrato di diminuire la mortalità, a costi sociali e individuali sostenibili, ve ne sono altri (come quelli del cancro della prostata con il PSA o del polmone con la TC spirale) sul cui beneficio il consenso della comunità medica non è unanime o è in attesa di ulteriori studi.

## **Il giudizio clinico complesso**

In qualsiasi ambito, e anche in quello medico, lo stesso evento può essere interpretato in modo diverso da diversi osservatori o anche dallo stesso osservatore, in tempi diversi (variabilità interosservatori e intraosservatore o non riproducibilità dell’osservazione). In medicina, i fattori di mancata riproducibilità di un’osservazione possono essere relativi all’esaminatore, all’esaminato e all’esame stesso.

Per esempio, alcuni medici tendono ad annotare in cartella sintomi e segni clinici sotto forma di inferenze diagnostiche, scrivendo, per esempio, “dolore pleuritico” invece che “inspiratorio” o “cute anemica” invece che “pallida”. Ma un’inferenza scorretta può sviare il processo diagnostico. Anche le distorsioni da aspettative a priori possono essere causa di errore. I medici, come tutti gli esseri umani, sono portati a trovare ciò che si attendono di trovare.

Perché la diagnosi sia accurata, non è sufficiente che due medici la pensino allo stesso modo: infatti, anche se uno dei due decidesse tirando in alto una moneta, potrebbe capitare che si trovasse d’accordo con il collega e non per questo la convergenza dei giudizi aumenterebbe l’accuratezza della diagnosi: la «riproducibilità» è l’accordo tra diversi giudizi clinici, al di là del caso. Gli statistici misurano il grado percentuale di accordo non dovuto al caso con il cosiddetto «indice kappa». Secondo J.R Landis e G.G. Koch (1977) esso va da  $< 20$ , che significa riproducibilità molto scarsa, a  $> 80$  (accordo quasi perfetto; tra 50 e 80, l’accordo è buono).

In un esperimento, è stato chiesto a un gruppo di medici di valutare se l’arteria dorsale di un paziente fosse o no pulsante e, in un altro, di valutare la presenza o meno di ostruzione bronchiale: in entrambi i casi, il grado di accordo tra colleghi, al di là del caso, è stato del 51%.

Nel lontano 1945 H. Bakwin pubblicò uno studio emblematico della variabilità tra osservatori: 389 scolari di 11 anni furono esaminati una prima volta da un gruppo di medici; a 174 di loro (45%) venne raccomandata la tonsillectomia e i restanti 215 furono visitati da un secondo gruppo di medici, che raccomandarono la tonsillectomia in 99 casi (46%). Quando i rimanenti 116 bambini, le cui tonsille per ben due volte erano state considerate sane, furono esaminati da un terzo gruppo di medici, in altri 51 casi (44%) fu consigliato l’intervento .

## **La variabilità dell'esaminato**

La variabilità del giudizio clinico riferibile all'esaminato, cioè al paziente, è più evidente in medicina generale: negli ambulatori arrivano persone che, come dice il nome dato agli studi dei medici, "deambulano" e che hanno, quindi, libertà di movimento e, con esso, d'azione e di giudizio; persone portatrici non solo di una patologia, ma di un'intera biografia, con emozioni, priorità e valori che, nella vita reale, diversamente che nelle situazioni sperimentali degli studi clinici controllati, pesano sulle decisioni diagnostiche e terapeutiche. Quando sottopone al medico un problema di salute, il paziente si è già fatto un'idea circa la sua natura e riferisce i sintomi o i segni con enfasi diversa a seconda della preoccupazione che essi gli suscitano. Ripetere l'anamnesi può portare a riorganizzazioni a posteriori di ricordi frammentari: in anamnesi successive si possono così avere sostanziali cambiamenti per quanto riguarda le sequenze temporali e anche la presenza o assenza nel passato di eventi di grande importanza clinica.

In uno studio su un campione casuale di 155 pazienti di medici di base, si è constatato che in un quinto delle anamnesi non erano stati riferiti importanti elementi clinici, quali melena (sangue nelle feci) o pollachiuria (l'urinare di continuo) (Hoornaert 1977). Tali omissioni erano più frequenti non quando i pazienti erano ostili nei confronti del medico, ma quando lo apprezzavano e lodavano. Si è anche visto che il paziente affezionato al suo medico curante può tendere, per compiacerlo, a occultare la presenza e la rilevanza degli effetti collaterali della terapia (Snyder 1976).

## **L'alleanza di cura**

Se il racconto dei sintomi e dei segni da parte del paziente può essere alterato dalla soggettività, anche il medico, dal canto suo, effettua una selezione involontaria degli elementi rilevanti dell'anamnesi, in base sia alla valutazione dei precedenti clinici del paziente sia dell'esperienza della sua credibilità: in un certo senso, "ricostruisce" mentalmente il racconto.

Il processo di messa in comune dei due racconti (quello del paziente e quello del medico) è un processo negoziale, che sfocerà in ciò che si chiama "l'alleanza di cura".

Nelle condizioni sperimentali degli studi clinici controllati, i pazienti sono selezionati e si cerca di eliminare i fattori confondenti. Nelle situazioni reali, invece, abbondano i con fonditori: genere, età, malattie concomitanti, preferenze individuali, paura, negazione.

Per mediare con la soggettività del paziente e approdare all'alleanza di cura, capita a volte che il medico si trovi a vivere una divergenza tra ciò che le conoscenze teoriche gli suggerirebbero di fare e ciò che si trova a fare concretamente.

Lo psicologo e sociologo americano Leon Festinger ha coniato cinquant'anni fa il termine "dissonanza cognitiva" per definire questa sensazione di incoerenza, alla quale ogni medico reagisce secondo la propria personalità: quello che non teme il conflitto con il paziente applicherà rigidamente i risultati degli studi scientifici più accreditati, mentre quello che ha la tendenza ad "accontentare" il paziente, sceglierà i trial che assolvono le decisioni concordate, ignorando i trial che le smentiscono.

## **Tante teste, tante opinioni**

E' esperienza di tutti che, a volte, una stessa patologia venga curata in modo diverso da medici diversi. Per limitare l'eterogeneità di comportamento clinico, vengono proposte ai medici le linee guida, che sono indicazioni tracciate da esperti per la scelta di esami e terapie per una determinata malattia.

Ma basta l'applicazione delle linee guida per garantire al paziente la miglior cura possibile e, soprattutto, il minor danno iatrogeno possibile?

Sulla rivista *JAMA* è stata pubblicata la simulazione di cura di un'ipotetica paziente di 79 anni, affetta da artrosi, osteoporosi, diabete di tipo 2, ipertensione e bronchite cronica.

Se a questa situazione, tutt'altro che fantasiosa, fossero applicate le linee guida disponibili, l'anziana paziente assumerebbe 12 diverse medicine 5 volte al giorno, riceverebbe 14 diversi suggerimenti su cosa mangiare e quanto muoversi e sarebbe sottoposta a un gran numero di visite di controllo ed esami strumentali disseminati lungo l'anno. L'incrocio delle cinque linee genererebbe interazioni tra farmaco e farmaco, tra farmaco e alimentazione, tra farmaco per una malattia e decorso di un'altra malattia.

Il commentatore dell'articolo concludeva che, purtroppo, i pazienti reali sono spesso più complessi di quelli degli studi clinici.

Gli stessi studi clinici controllati condotti per valutare nuovi farmaci, nascondono, poi, insidie interpretative, sia per quanto riguarda la miglior efficacia di questi rispetto a farmaci già in uso (spesso le case farmaceutiche usano l'escamotage di non fare il confronto con terapie la cui utilità è già provata, ma con il placebo, cioè nessuna terapia), sia per quanto riguarda la loro sicurezza d'uso, vale a dire, l'entità degli effetti collaterali.

Si possono fare due esempi recenti:

- È noto che i FANS (farmaci anti infiammatori non steroidei) sono molto dannosi per lo stomaco: negli Stati Uniti, il numero di morti da sanguinamento da FANS è pari a quello di morti da AIDS e crimini violenti. La commercializzazione degli antinfiammatori coxib, che vantavano una gastrolesività dimezzata sembrava consentire il loro uso prolungato in pazienti con artrite reumatoide e nei pazienti anziani con precedenti di ulcera. Studi post marketing hanno, però, messo in luce un'associazione tra l'uso prolungato di rofecoxib e aumento di infarto del miocardio e di ictus, tanto che si è arrivati al ritiro della molecola. La Infine, l'FDA ha avvertito, in un memorandum del 2005, che l'aumentato rischio cardiovascolare è un effetto avverso di tutta la classe degli antinfiammatori, selettivi e non.

- Si è visto che la simvastatina abbassa più efficacemente il colesterolo in associazione con ezetimibe, un inibitore dell'assorbimento intestinale del colesterolo. Secondo i dati dello studio SEAS (Simvastatin and Ezetimibe in Aortic Stenosis), però, quest'accoppiata di molecole sarebbe legata a un aumento di morti per cancro (la spiegazione biologica è che l'ezetimibe impedirebbe anche l'assorbimento di molecole vegetali protettive contro i tumori). Recentemente, un'analisi di tutti i trial sull'ezetimibe condotta da Peto e pubblicata sul *NEJM*, non sono riusciti a confermare questo sospetto.

Insomma, come dice il sociologo Edgar Morin, [“la conoscenza è una navigazione in un oceano di incertezze attraverso arcipelaghi di certezze”](#). Questa consapevolezza non ha, però, necessariamente, una connotazione negativa, né deve spingere allo scetticismo; anzi, configura l'opportunità di rendere la conoscenza più pertinente con la continua verifica e con l'uso costante della razionalità, barriera fondamentale contro l'errore.

1) Kassirer J.P. Diagnosis in the public domain. *N Engl J Med* 1993; 329: 50.

2) Boyd C et al. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases. *JAMA* 2005; 294: 716.

3) Morin E. I sette saperi necessari all'educazione del futuro. Milano: Raffaello Cortina, 2001.

4) Caimi V: Decidere in condizioni di incertezza. Manuale di medicina generale. Milano: UTET, 2003.

5) Bursztajn HJ et al. Medical choices, medical chances. How patients, families and physicians can cope with uncertainty. New York: Routledge, 1990.

6) Corwin RG et al. Patients accuracy in reporting their past medical history. A study of 90 patients with peptic ulcer. *J Chronic Dis* 1971; 23: 875.

7) Ehrlich HJ et al. Therapists feelings toward patients and patient treatment and outcome. *Soc Sci Med* 1967; 1: 283.

8) Hoornaert F et al. Transference aspects of doctor-patient relationship in psychomatic patients. *Br J Med Psychol* 1976; 49: 261.

9) Landis JR et al. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159.

10) Mcanulty JH et al. Spontaneous changes in left ventricular function between sequential studies. *Am J Cardiol* 1974; 34: 23.

11) Rose g et al. Observer variation. *Brit Med J* 1978; 2: 1006.

12) Sackett DL Clinical disagreement: How often it occurs and why. *Can Med Assoc J* 1980; 123: 499.

13) Snyder d et al. Doctor-patient communication in a private practise. *J Fam Pract* 1976; 3: 271.

14) Bakwin H. Pseudodoxia Pediatrica. *N Eng J Med*. 1945; 232: 691.