

L'interpretazione delle evidenze. Per ridurre la soggettività nella lettura delle prove serve tornare a insegnare l'evidence-based medicine

RECENTI PROGRESSI IN MEDICINA

Riassunto. L'interpretazione delle evidenze della ricerca clinica è ancora caratterizzata da un'ampia soggettività. Questa soggettività è visibile anche dal confronto delle linee guida e delle raccomandazioni di istituzioni e società scientifiche. Spesso è dovuta ai condizionamenti vissuti dai componenti dei panel delle linee guida: diventa dunque sempre più importante il ruolo degli editor e dei redattori delle riviste e dei media scientifici, che dovrebbero tornare a vigilare attentamente sui contenuti di quello che viene pubblicato. Per affrontare il problema, è però necessario tornare a insegnare l'evidence-based medicine per restituirle la funzione di metodo-guida nella pratica clinica e nella definizione delle strategie di sanità pubblica.

Interpreting the evidence. Reducing subjectivity in appraising the evidence calls for a return to teaching evidence-based medicine.

Summary. The interpretation of clinical research evidence is still characterized by wide subjectivity. This subjectivity is also visible when comparing guidelines and recommendations developed by institutions and learned societies. It is often due to bias and conflicts of interest experienced by the members of guideline panels: thus, the role of editors and publishers of journals and scientific media becomes increasingly important, and they should return to careful oversight of the content of what is published. To address the problem, however, it is necessary to return to teaching evidence-based medicine in order to restore its function as a "North star" in clinical practice and public health decision-making.

La comunità scientifica non sempre si trova d'accordo su come interpretare le stesse evidenze e sulle raccomandazioni che ne possono derivare. Anche i percorsi più rigorosi e sistematici per risolvere le divergenze di opinione nello sviluppo delle linee guida, per esempio, possono non eliminare le divergenze di punti di vista. Nel campo degli screening oncologici, queste diversità sono spesso molto evidenti: non coincidono, per esempio, le raccomandazioni dell'American College of Physicians, dell'American Cancer Society e della US Preventive Services Task Force (USPSTF) sul momento in cui avviare lo screening per il tumore del colon-retto e sui metodi di screening da preferire. Allo stesso modo, le raccomandazioni per lo screening del cancro della mammella sono ancora oggetto di controversia come hanno illustrato Howard Bauchner e John Ioannidis in una viewpoint uscita sull'health forum del JAMA¹.

Clinici, ricercatori e panel di estensori di linee guida seguono spesso diversi criteri di selezione sulle fonti da considerare per la redazione delle raccomandazioni, criteri dipendenti soprattutto dalla valutazione della qualità della ricerca alla base delle evidenze. In alcuni Paesi le linee guida sono il riferimento per il rimborso delle spese sostenute dai pazienti: negli Stati Uniti, l'*Affordable care act* stabilisce che le raccomandazioni di screening di livello A o B dell'USPSTF siano coperte dalle assicurazioni private senza costi per i pazienti. In altre nazioni, come l'Italia, hanno valenza legale.

Dunque, i diversi metodi per la preparazione delle linee guida possono portare a raccomandazioni di-

verse. Molte linee guida seguono il metodo GRADE (Grading of Recommendations Assessment, Development and Evaluation) che non di rado – e a torto – viene giudicato eccessivamente complesso e per consentire la preparazione tempestiva di linee guida di pratica clinica. Altra caratteristica spesso sottolineata è la scarsa flessibilità, ma è lo stesso working group del GRADE a ridimensionare il problema: «GRADE is not a quantitative system for grading the quality of evidence. Each factor for downgrading or upgrading reflects not discrete categories but a continuum within each category and among the categories. When the body of evidence is intermediate with respect to a particular factor, the decision about whether a study falls above or below the threshold for up- or downgrading the quality (by one or more factors) depends on judgment»².

Tuttavia, resta il fatto che l'insieme delle variabili insite in diversi step della valutazione può incidere molto sull'esito delle analisi. Sia i singoli componenti di un panel di linea guida, sia i diversi gruppi di lavoro possono dare valutazioni diverse delle stesse prove o – per esempio – pesare in maniera differente elementi che possono aver condizionato i risultati di uno studio, come anche selezionare dati diversi da includere nella valutazione ai fini di una raccomandazione. L'ex direttore del JAMA e il ricercatore del Meta-research innovation center (Metrics) di Stanford fanno anche osservare che, nella maggior parte degli interventi medici, le prove di efficacia sono spesso così deboli che le raccomandazioni dipendono in larga misura

da opinioni soggettive. Uno degli effetti collaterali è che l'opinione pubblica ha sempre più spesso la percezione che le notizie non siano oggettive, e la risposta ai dubbi dei cittadini spesso è caratterizzata da eccessiva sicurezza.

«La comunicazione dei risultati degli studi controllati randomizzati (Rct) è un buon esempio. Anche in un Rct positivo, il number needed to treat può essere elevato; cioè, la maggior parte degli individui non ne trarrà beneficio. Ulteriori dettagli possono sfuggire, per esempio la differenza tra effetti clinicamente importanti ed effetti statisticamente significativi, la complessità nell'interpretare i risultati di non inferiorità, le sfide di comprendere i risultati di endpoint compositi. I ricercatori, i direttori delle riviste e coloro che riportano le notizie o vengono intervistati come esperti dovrebbero esprimere chiaramente questa incertezza e le sue sfumature, nonché la quantità e la qualità delle prove che supportano le loro affermazioni»¹.

L'interpretazione delle prove continua dunque a essere un problema. Sicuramente chi fa ricerca – e i clinici che devono interpretarne i risultati – dovrebbero sforzarsi di essere il più possibile obiettivi: ma è facile a dirsi tanto quanto è difficile a rispettare un impegno del genere. Diventa quindi ancora più importante e delicato il ruolo degli editor che lavorano nei media specialistici, che dovrebbero vigilare sull'accuratezza dei contenuti di ciò che viene loro proposto dagli autori. Ciò non vuol dire che la soggettività delle interpretazioni debba essere annullata: dovrebbe però essere chiaramente documentato il percorso di elaborazione dei documenti come linee guida e raccomandazioni istituzionali.

«Infine – sostengono Bauchner e Ioannidis – gli autori e gli esperti dovrebbero essere consapevoli dei

loro pregiudizi, valutando attentamente se hanno analizzato le prove in modo obiettivo. Essendo complicato valutare sé stessi, editor e media dovrebbero guardare con occhio critico i contenuti che vengono loro proposti». Ma purtroppo anche gli stessi editor potrebbero essere influenzati, ancora di più dei loro autori¹.

«La quantità di incomprensioni sulle prove che vedo quotidianamente è il motivo per cui farò sempre questo lavoro» ha scritto in un tweet David Nunan, direttore del corso di Evidence-based healthcare dell'Università di Oxford. «Moltissime persone non sanno cosa stanno leggendo o commentando. Compreso chi fa revisioni sistematiche, ma ancora di più coloro che scrivono o le commentano. Soprattutto, non hanno un sistema o un metodo per capire le evidenze. È un concetto introduttivo chiave per i miei studenti quando iniziano a conoscere questi sistemi e questi metodi. Non possono restare nella stessa ignoranza precedente il corso».

Guardando avanti, serve rilanciare la discussione e il confronto sulle linee guida, sulla metodologia di produzione e sulla loro utilità nell'orientare l'assistenza. In sostanza, servirebbe tornare a parlare – ma prima ancora a spiegare cosa sia – di medicina delle prove. Ripartendo dalla consapevolezza dell'incertezza che ancora caratterizza gran parte del sapere medico-scientifico.

Bibliografia

1. Bauchner H, Ioannidis JP. The subjective interpretation of the medical evidence. *JAMA Health Forum* 2024; 5: e240213-e240213.
2. Schünemann H, Brożek J, Guyatt G, Oxman A (eds). *The GRADE handbook*. 2013.