

GRADE

IL GRADE rappresenta una delle più importanti novità di questi ultimi 10 anni nel campo della valutazione dei servizi sanitari ed è un sistema strutturato e trasparente per analizzare e presentare la sintesi della qualità delle prove scientifiche con importanti ricadute dal punto di vista degli operatori sanitari e dei pazienti. Il GRADE Working Group - Grading of Recommendations Assessment, Development and Evaluation - è nato nel 2000 come una collaborazione informale di ricercatori con un interesse a migliorare i problemi dei sistemi di grading esistenti nell'assistenza sanitaria e in brevissimo tempo è stato adottato a livello internazionale da più di 30 organizzazioni quali l'American College of Physicians, dall'American Thoracic Society, UpToDate, la Cochrane Collaboration, l'OMS e, in Italia, dalla Agenzia Sociale e Sanitaria della Regione Emilia Romagna. La sua adozione così ampia riflette il rigore metodologico associato alla facilità d'uso.

In generale, gli ideatori di questo metodo hanno cercato di lavorare sulle rigidità della evidence-based medicine, basata sull'interpretazione eccessivamente forzata dei soli studi clinici randomizzati. L'esplicitazione della domanda clinica e la scelta degli esiti importanti per i pazienti sono state poste così al centro del percorso di valutazione degli interventi: tutto questo è stato associato a una esplicita valutazione delle prove con un riflesso positivo sia sulla trasparenza delle informazioni sia sulla trasferibilità delle stesse nei diversi ambiti applicativi.

Sino a non molto tempo fa la valutazione avveniva soprattutto mettendo sotto analisi la struttura gerarchica delle prove, basando il giudizio essenzialmente sul tipo di architettura di ricerca attraverso il quale erano state ottenute le informazioni. L'applicazione rigida, e talvolta "acritica" dei tradizionali "livelli di evidenza" è stata così messa progressivamente in discussione: non esiste il disegno di studio in assoluto "perfetto e ideale", indipendentemente dal tipo di quesito. Ciascun quesito di studio (e ricerca) deve essere affrontato con la metodologia e il disegno di studio più adeguati. Infatti, se per molti quesiti terapeutici gli studi controllati randomizzati (RCTs) sono la migliore opzione metodologica, per quesiti di tipo non terapeutico il RCT non

solo non è utile, ma la sua applicazione è, di fatto, inappropriata.

Il metodo GRADE può essere utilizzato per produrre raccomandazioni clinico-organizzative (se si utilizzano la valutazione della qualità delle prove e la forza delle raccomandazioni), ma può anche essere utile per la sola valutazione di qualità all'interno di revisioni sistematiche e di rapporti di health technology assessment. Questo metodo mette a disposizione di chi sviluppa raccomandazioni, o intere linee guida, un sistema strutturato per affrontare e sviluppare tutti i passaggi necessari. Esso infatti richiede che si parta da un approccio strutturato alla formulazione dei quesiti, alla scelta degli esiti (outcome) di interesse, alla graduazione della loro importanza, alla valutazione delle prove e alla loro integrazione con i valori e le preferenze dei pazienti e della società, per giungere alla formulazione delle raccomandazioni.

Il primo passaggio del processo di produzione di una raccomandazione presuppone la formulazione di una domanda chiara e definita (vedi figura).

Ogni domanda che tratta di un problema clinico dovrebbe avere quattro componenti: i pazienti, l'intervento, il confronto e l'esito di interesse (Patients, Intervention, Comparison, Outcomes - PICO) (Oxman, 1988). Esso richiede poi che siano specificati tutti gli esiti che sono importanti per i pazienti fin dall'inizio del processo di formulazione di una raccomandazione e invita a differenziare gli esiti critici da quelli importanti, ma non critici. A tal fine il GRADE prevede un percorso strutturato e il più possibile esplicito che passa attraverso l'identificazione - da parte del gruppo di lavoro che deve produrre la raccomandazione - di tutti gli outcome potenzialmente rilevanti e la loro individuale e collegiale valutazione.

Il GRADE propone una valutazione della qualità delle prove più ampia e articolata di quella proposta da tutti gli altri metodi oggi disponibili. La principale e maggiore novità introdotta dal metodo consiste nel richiedere uno spostamento da una valutazione "studio specifica" a una "outcome specifica". La motivazione di questo spostamento di attenzione deriva dalla consapevolezza che non basta valutare l'appropriatezza del disegno di studio per determinare la qualità delle prove; ci deve

Come arrivare a una raccomandazione con il metodo GRADE.

Definizione del quesito (Patients, Intervention, Comparison, Outcomes)	
Definizione dell'importanza relativa degli esiti	
Ricerca delle prove di efficacia	
Valutazione della qualità delle prove per ciascun esito	
<p>I fattori che possono abbassare la qualità delle prove:</p> <ul style="list-style-type: none"> • le limitazioni degli studi • l'inconsistenza fra i risultati • la scarsa trasferibilità/applicabilità delle prove • l'imprecisione della stima degli effetti • bias di pubblicazione 	<p>I fattori che possono aumentare la qualità delle prove sono:</p> <ul style="list-style-type: none"> • la grande dimensione dell'effetto • l'assenza di fattori di confondimento che plausibilmente potrebbero ridurre un effetto dimostrato; • la presenza di un gradiente dose-risposta
Riassunto delle prove per ciascun esito critico o importante	
Valutazione della qualità globale delle prove	
Bilanciamento dei benefici e degli eventi avversi	
Bilanciamento fra i benefici e i costi	
Definizione della raccomandazione e della sua forza	

Fonte: Adattato da Oxman, 2004

essere anche una appropriatezza del disegno di studio rispetto all'outcome valutato. Ad esempio outcome hard non richiedono l'impiego del doppio cieco mentre lo richiedono outcome soft. Dal punto di vista degli step operativi, per valutare la qualità delle prove, il metodo GRADE prevede che si inizi dal disegno dello studio. Per le raccomandazioni che si occupano di strategie alternative di gestione clinica – a differenza di quelle che si occupano di prognosi o accuratezza dei test diagnostici – gli RCTs forniscono, in generale, prove più forti rispetto agli studi osservazionali. Studi osservazionali rigorosamente condotti forniscono a loro volta una prova più forte rispetto alle serie di casi non controllate. Il GRADE richiede che si faccia una valutazione separata della qualità delle prove per ciascuno degli esiti importanti per i pazienti e identifica cinque fattori che possono abbassare la qualità delle prove, applicabili sia agli studi randomizzati sia a quelli osservazionali, e tre fattori che possono alzarla. La cosiddetta "evidence table" riassume i dati quantitativi degli studi e la valutazione qui sotto descritta.

Limitazioni degli studi. La valutazione della qualità delle prove si abbassa se gli studi hanno importanti limitazioni che possono inficiare le stime dell'effetto del trattamento. Queste limitazioni includono:

1. la mancanza del mascheramento della sequenza randomizzata di allocazione al trattamento;
2. mancanza o carenza delle misure adottate per assicurare la "cecità" dei partecipanti allo studio e del personale sanitario e non sanitario rispetto al trattamento assegnato;
3. ampie perdite di pazienti al follow-up con incompletezza dei dati per ciascun esito considerato e/o mancata adesione al principio dell'intention to treat analysis;
4. l'interruzione precoce di uno studio a causa degli effetti benefici dell'intervento sperimentale considerato;
5. la mancata descrizione di alcuni esiti nei risultati degli studi, tipicamente per quelli in cui non sia stato osservato un effetto dell'intervento sperimentato.

Inconsistenza fra i risultati degli studi esaminati per ciascun esito (inconsistency). La presenza di ampie differenze nelle stime dell'effetto di un intervento fra gli studi considerati suggerisce la presenza di differenze reali nel sottostante effetto del trattamento. La variabilità può derivare da differenze nelle popolazioni in esame, da differenze fra gli interventi considerati o tra gli esiti misurati. Quando l'eterogeneità esiste, ma non si riesce a trovare una spiegazione plausibile a supporto di essa, la qualità delle prove deve essere abbassata.

La trasferibilità/applicabilità delle prove (indirectness).

Chi produce linee guida si trova di fronte a tre tipi di indirectness delle prove.

- La prima si realizza quando gli esiti misurati differiscono da quelli d'interesse primario. È questo il caso tipico degli esiti surrogati, cioè di esiti non sostanziali per rispondere a un quesito di salute, ma valutati presumendo che la loro misurazione rifletta quella degli esiti veramente rilevanti. L'esperienza dimostra però che questa assunzione conduce a interpretazioni fuorvianti e di conseguenza la qualità delle prove dovrebbe essere considerata più bassa rispetto a quella fornita da esiti non surrogati.
- La seconda si realizza nel caso di confronti indiretti fra due farmaci utilizzati per la medesima patologia. Benché possano essere disponibili RCT che confrontino direttamente un farmaco (detto A) con il placebo e altri che confrontino un farmaco (detto B) sempre con il placebo, possono non esserci dati che permettono un confronto tra i due farmaci direttamente. Quando abbiamo un confronto indiretto, la qualità delle prove è considerata più bassa, rispetto a un confronto testa a testa.
- Il terzo tipo di indirectness delle prove include invece differenze fra le popolazioni, interventi, confronti e esiti di interesse, che rende poco o per nulla trasferibili le informazioni derivanti dalle prove disponibili alle popolazioni, interventi, confronti o esiti di interesse per la linea guida.

Imprecisione nella stima degli effetti. Quando gli studi includono pochi pazienti e pochi eventi e quindi hanno ampi intervalli di confidenza, un panel di linee guida può giudicare la qualità delle prove più bassa, proprio perché viene introdotto un livello più o meno elevato di imprecisione nella stima dell'effetto che può rendere incerti della stessa e del suo reale significato.

Bias di pubblicazione. La qualità delle prove sarà poi ridotta se i ricercatori hanno selettivamente incluso solo gli studi che riportano risultati statisticamente significativi per gli esiti considerati.

I fattori che possono aumentare la qualità delle prove. In alcune circostanze gli studi osservazionali possono produrre prove di qualità moderata o anche alta. Ciò si verifica quando sono ben condotti da un punto di vista metodologico e portano ad ampie e consistenti stime nella dimensione dell'effetto di un trattamento. Anche la presenza di un gradiente dose-risposta induce ad aumentare la qualità delle prove.

Una volta definita la qualità delle prove per singolo esito i membri del panel – dopo la discussione collegiale – dovranno esprimere un giudizio sulla qualità totale delle prove. Gli autori del GRADE propongono che la qualità della evidenza più bassa per tutti gli esiti critici dovrebbe fornire la base per valutare la qualità della prova generale. Inoltre, gli esiti che sono importanti, ma non critici nella decisione clinica, dovrebbero essere inclusi nei profili delle evidenze e dovrebbero essere considerati quando si fornisce un giudizio fra i benefici e i rischi, ma non dovrebbero essere tenuti in considerazione quando si valuta la qualità della prova generale. In sintesi, considerando e valutando tutti i fattori precedentemente descritti, il GRADE classifica la qualità delle prove in quattro livelli: alto, moderato, basso e molto basso.

Come si attribuisce la forza di una raccomandazione

La “forza” di una raccomandazione deve servire a esprimere e comunicare quanto si è convinti del fatto che “applicare la raccomandazione produrrà più beneficio che danno”.

Nel GRADE la forza della raccomandazione è determinata dal bilanciamento di quattro fattori chiave:

1. il bilanciamento fra gli effetti desiderati e non desiderati: tanto più ampia è la differenza fra gli effetti desiderabili e non desiderabili, tanto più è probabile che una raccomandazione sarà forte. Quanto più incerto sarà tale bilanciamento, tanto più alta sarà la probabilità che una raccomandazione sia debole;
2. la qualità delle prove: tanto più è alta la qualità delle prove, tanto più alta sarà la probabilità di avere una raccomandazione forte;

3. i valori e le preferenze: tanto più i valori e le preferenze dei pazienti variano o tanto più ampia è l'incertezza in essi, tanto più elevata sarà la probabilità che la raccomandazione sia debole;
4. tanto più alto è il costo di un intervento o maggiori le risorse consumate, tanto più sarà bassa la probabilità di avere una forte raccomandazione.

Il GRADE propone di classificare le raccomandazioni in forti e deboli, a favore o contro l'uso di un determinato trattamento/intervento. Una raccomandazione forte dovrebbe essere attribuita a quelle situazioni nelle quali si ritiene che la maggior parte dei pazienti – una volta adeguatamente informati – sceglierebbe di seguire la raccomandazione. Una raccomandazione debole dovrebbe invece essere usata in tutte quelle situazioni nelle quali si ritiene che le scelte dei pazienti varieranno sulla base dei propri valori e preferenze e che i clinici debbano creare le migliori condizioni per garantire che i pazienti possano scegliere sulla base dei loro valori e preferenze. A fianco delle "evidence table" sono state create delle "tavole di sintesi dei risultati" (Summary of findings tables - SOF) che presentano i principali risultati di una revisione in un formato tabulare, semplice e trasparente. Esse mettono a disposizione le informazioni "chiave" riguardo alla qualità delle prove, la dimensione dell'effetto degli interventi esaminati e una sintesi di tutti i dati disponibili per ciascuno degli esiti importanti in una revisione.

Silvia Pregno

*Methodology Review Group della Cochrane Collaboration.
Ha collaborato con il Norwegian Knowledge Centre di Oslo
nell'implementazione del metodo GRADE.*

BIBLIOGRAFIA

- GRADE Working Group.
Grading quality of evidence and strength of recommendations
BMJ 2004; 328: 1490-1498.
- Guyatt GH, Oxman AD, Kunz R et al for the GRADE Working Group
GRADE: going from evidence to recommendations
BMJ 2008; 336:1049-1051.
- Guyatt GH, Oxman AD, Kunz R et al for the GRADE Working Group
GRADE: what is "quality of evidence" and why is it important to clinicians?
BMJ 2008; 336: 995-998.
- Guyatt GH, Oxman AD, Vist GE et al for the GRADE Working Group
GRADE: an emerging consensus on rating quality of evidence and strength of recommendations
BMJ 2008; 336: 924-926.

Higgins JPT, Green S (editors).
Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0
[updated February 2008].
The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org

Schünemann HJ, Oxman AD, Broze J et al
Grading quality of evidence and strength of recommendations for diagnostic tests and strategies
BMJ 2008; 336; 1106-1110.



**LA SALUTE
È UN DIRITTO DI TUTTI**

Un **blog** sulla salute internazionale
con Gavino Maciocco
Maurizio Marceca, Enrico Materia
Sabina Nuti, Walter Ricciardi

**L'importanza di approfondire
Il piacere di condividere**

www.saluteinternazionale.info