Associazione Italiana di Oncologia Medica

Manuale metodologico
LINEE GUIDA AIOM 2019

Finalità e caratteristiche.
Metodologia applicata alle linee guida AIOM 2019

23 gennaio 2019

Indice

LINEE GUIDA AIOM 2019: FINALITÀ E CARATTERISTICHE 3

1. FINALITÀ E UTENTI TARGET 3

2. CARATTERISTICHE DELLE LINEE GUIDA AIOM 3

A. STRUTTURA ORGANIZZATIVA 3

B. COMPOSIZIONE DEI GRUPPI DI LAVORO DELLE SINGOLE LG 3

C. METODOLOGIA 4

D. REVISIONE DEI DOCUMENTI 6

E. INDIPENDENZA E CONFLITTO DI INTERESSE 6

F. AGGIORNAMENTO PERIODICO 6

G. PUBBLICAZIONE NEL SITO AIOM 6

H. VALUTAZIONE DELLA IMPLEMENTAZIONE 7

3. CONFRONTO CON LE LINEE GUIDA INTERNAZIONALI 7

4. BIBLIOGRAFIA 9

METODOLOGIA APPLICATA ALLE LINEE GUIDA AIOM 10

1. SCOPO E CAMPO DI APPLICAZIONE 10

2. METODOLOGIA SCIENTIFICA 10

2.1. Formulazione del Quesito Clinico 10

2.2. Selezione della letteratura a supporto delle raccomandazioni cliniche 11

2.3. Valutazione critica della letteratura a supporto delle raccomandazioni cliniche 12

2.3.1. La descrizione degli studi per esteso 12

2.3.2. Qualità dell'evidenza (SIGN) – da considerare solo per la lettura delle raccomandazioni prodotte fino al 2015 12

2.3.2.1. Considerazioni sulla "estrapolazione delle evidenze" 14

2.4. Qualità dell'evidenza (GRADE) 15

2.4.1. Bilancio tra benefici e danni 17

3. Raccomandazioni cliniche 18

3.1. Tabelle sinottiche raccomandazioni cliniche 19

3.2. Sintesi della raccomandazione e del rapporto beneficio/danno 21

4. CRITICAL APPRAISAL 22

5. Voci bibliografiche 25

LINEE GUIDA AIOM 2019: FINALITÀ E CARATTERISTICHE

Stefania Gori, Michela Cinquini, Giovanni Pappagallo, Ivan Moschetti, Valter Torri

La produzione di Linee Guida (LG) rappresenta uno degli elementi principali della medicina basata sull'evidenza e rappresentano una delle modalità con cui vengono trasferite nella pratica clinica le conoscenze prodotte dalla ricerca, sintetizzando le migliori prove di efficacia disponibili.
Scopo delle Linee Guida è di produrre raccomandazioni per la pratica clinica attraverso un processo sistematico e trasparente. Le raccomandazioni devono essere coerenti con il rapporto tra benefici e danni connessi all'intervento in esame.

1. FINALITÀ E UTENTI TARGET

Le finalità delle Linee Guida AIOM sono:
1. Migliorare e standardizzare "la pratica clinica"
2. Offrire al paziente sull'intero territorio nazionale la possibilità della "migliore cura"
3. Garantire un riferimento basato sull'evidenza per le istituzioni nazionali e regionali, per gli organismi regolatori ed i "payers".

Attualmente le Linee Guida AIOM sono rivolte solo all'utilizzo medico. Sono disponibili da aprile 2015 nel sito di AIOM Fondazione opuscoli informativi per i pazienti, redatti dai singoli Gruppi di Lavoro delle Linee Guida, in attesa di produrre anche Linee Guida formali rivolte ai pazienti.

2. CARATTERISTICHE DELLE LINEE GUIDA AIOM

A. STRUTTURA ORGANIZZATIVA

Il Gruppo Linee Guida AIOM comprende il Presidente AIOM, il Presidente eletto, due membri del direttivo nazionale AIOM a rappresentanza del Direttivo stesso e da altri professionisti, oncologi medici AIOM.
E' inoltre costituito da metodologi, da un rappresentante pazienti, e in prospettiva da rappresentanti degli infermieri e dei Medici di Medicina Generale.
Il Gruppo di Lavoro Linee Guida AIOM è coordinato dal Presidente Nazionale AIOM, in considerazione del fatto che, a seguito della Legge Biondi-Gelli del marzo 2017 e del Decreto Ministeriale Lorenzin del 2 agosto 2017, la produzione di Linee Guida è diventato uno dei due obiettivi strategici della Associazione Italiana di Oncologia Medica.
Questa organizzazione rende più rapida e efficace la parte organizzativa e operativa nella produzione di LG nelle fasi di produzione e aggiornamento annuali. Il Gruppo Metodologico svolge un ruolo fondamentale. Costituito da 4 metodologi, il gruppo infatti indica la metodologia da utilizzare per la stesura delle LG, supporta l'applicazione della metodologia valutando la qualità delle evidenze e producendo i documenti riassuntivi ed inoltre svolge attività di formazione per il gruppo di lavoro di ogni LG.

B. COMPOSIZIONE DEI GRUPPI DI LAVORO DELLE SINGOLE LG

Attualmente i Gruppi di Lavoro sono 39 e la maggior parte è ancora composta da soli medici (oncologi e altri specialisti coinvolti nella gestione multidisciplinare del paziente).

Dal 2016, come raccomandato dalla WHO (Schunemann HJ et al, 2006), è in corso l'estensione alla partecipazione di medici di medicina generale e pazienti, fornendo anche a tali figure le competenze atte alla valutazione e alla decisione. In 6 Linee Guida ha preso parte a tutto il processo decisionale anche il paziente o Associazione di pazienti.
Figure più tipicamente "tecniche" (farmacisti, metodologi, economisti sanitari) possono partecipare ai gruppi di lavoro con la funzione di supporto formativo e conoscitivo, senza diritto di voto.

C. METODOLOGIA

Dal 2016 le raccomandazioni formulate nelle Linee Guida AIOM vengono prodotte adattando la metodologia GRADE (Manuale Metodologico 2015-2017, www.aiom.it) al contesto italiano.
Con il metodo GRADE, complesso e oneroso in termini di tempo e di impegno per i metodologi e per i componenti dei Gruppi di Lavoro di ogni LG, vengono affrontati per il momento soltanto quesiti terapeutici clinici riguardanti:

- ✓ farmaci/trattamenti di recente commercializzazione il cui bilancio tra benefici e danni risulti non essere chiarito in modo definitivo;
- ✓ farmaci/trattamenti storici qualora siano emersi nuovi dati di efficacia e/o sicurezza.

Approccio comune a tutte le LG rimane quello di formulare il quesito clinico verso il quale sarà prodotta la raccomandazione seguendo l'acronimo P.I.C.O. (P=popolazione; I=intervento; C=confronto; O=Outcome), evitando quesiti generici (ad es.: qual è il ruolo dell'intervento XXX nel trattamento del …..").

Quindi, il quesito dovrebbe essere posto in tal modo:

P: Nei pazienti con (*menzionare le specifiche caratteristiche di paziente, malattia, stadio, ecc.*)

I: il trattamento con (*descrivere l'intervento terapeutico oggetto del quesito*)

C: è suscettibile di impiego in alternativa a (*descrivere il trattamento altrimenti considerato in alternativa all'intervento in esame*)

O: con riferimento agli *outcome* di beneficio e di danno (*elencare i parametri clinico-laboratoristici – misure di effetto, outcome – ritenuti essenziali/importanti per la proposta terapeutica*)"

- **Il Gruppo di Lavoro** (gruppo multidisciplinare) sceglie gli outcome di beneficio e di danno a mezzo di una votazione sulla criticità degli stessi nel processo di formulazione della proposta clinica;
- Il processo di ricerca sistematica della letteratura avviene senza restrizione di lingua o limite temporale su più database (PubMed, EMBASE, CENTRAL, database area-specifici) e letteratura grigia (registri dei trial, WHO, ecc.) ed è effettuata da parte di un bibliometrista esperto.
- **Il panel indipendentemente dal Gruppo Metodologico** identifica gli studi potenzialmente eleggibili e ne valuta l'inclusione. A seconda del quesito clinico, i tipi di disegni di studio più appropriati per rispondere alla domanda saranno tenuti in considerazione. Non saranno invece inclusi editoriali, position paper, commenti, lettere, revisioni narrative, o case report. Qualora ci fosse la necessità di portare a supporto una Linea Guida già prodotta e pubblicata, questa dovrà avere almeno un punteggio pari a 5/7 della scala di valutazione AGREE II (maggiore attenzione sarà data al dominio 3: rigore metodologico).
- **Il gruppo metodologico è competente nel valutare la qualità delle evidenze per singolo outcome ,** che si basata su 5 dimensioni (*risk of bias, imprecision, inconsistency, indirectness, publication bias*), ed è sintetizzata in 4 livelli (alta, moderata, bassa, molto bassa), produce le tavole sinottiche delle evidenze e applica l'Evidence to Decision Framework (EtD) strumento che tiene presente anche i valori e le preferenze dei pazienti più una serie di altre dimensioni come fattibilità, equità, etc.
- Il panel procede poi alla valutazione del rapporto benefici/danni correlati all'intervento in oggetto (votazione) e, tramite EtD e discussione, partecipa alla votazione della forza della raccomandazione relativa al quesito clinico posto. Votazione che viene espressa secondo maggioranza. In casi in cui la

maggioranza non venga ottenuta in prima battuta, la votazione sarà ripetuta dopo nuova discussione/confronto con o senza nuove evidenze fino a un livello soddisfacente. In caso l'esito della votazione mostri costantemente un panel chiaramente diviso su posizioni differenti si suggerisce di riportare l'esito a seguire la raccomandazione spiegandone ove possibile la/le ragioni.

- Nel caso in cui manchino evidenze o una raccomandazione fosse per lo più basata sulle opinioni del panel, anche questo verrà esplicitato nel box della raccomandazione come di seguito: "* giudizio espresso dal panel" e la qualità delle evidenze sarà indicata con un asterisco (*).
- Inoltre alla fine di ogni raccomandazione verrà riportato il numero o la dicitura per esteso delle referenze che la supportano.
- La raccomandazione per la pratica clinica esprime l'importanza clinica di un intervento/procedura e la forza della raccomandazione clinica viene graduata su 4 livelli (Tabella 1).

Forza della raccomandazione clinica	**Terminologia**	**Significato**
Positiva Forte	"Nei pazienti con (*criteri di selezione*) l'intervento xxx **dovrebbe** essere preso in considerazione come opzione terapeutica di prima intenzione"	l'intervento in esame dovrebbe essere considerato come prima opzione terapeutica (evidenza che **i benefici sono prevalenti sui danni**)
Positiva Debole	"Nei pazienti con (*criteri di selezione*) l'intervento xxx **può** essere preso in considerazione come opzione terapeutica di prima intenzione, in alternativa a yyy"	l'intervento in esame può essere considerato come opzione di prima intenzione, consapevoli dell'esistenza di alternative ugualmente proponibili (**incertezza riguardo alla prevalenza dei benefici sui danni**)
Negativa Debole	"Nei pazienti con (*criteri di selezione*) l'intervento xxx **non dovrebbe** essere preso in considerazione come opzione terapeutica di prima intenzione, in alternativa a yyy"	l'intervento in esame non dovrebbe essere considerato come opzione di prima intenzione; esso potrebbe comunque essere suscettibile di impiego in casi altamente selezionati e previa completa condivisione con il paziente (**incertezza riguardo alla prevalenza dei danni sui benefici**)
Negativa Forte	"Nei pazienti con (*criteri di selezione*) l'intervento xxx **non deve** essere preso in considerazione come opzione terapeutica di prima intenzione"	l'intervento in esame non deve essere in alcun caso preso in considerazione (evidenza che **i danni sono prevalenti sui benefici**)

Tabella 1. Terminologia e significato delle raccomandazioni cliniche.

- In ossequio a quanto raccomandato dalla evidence-based practice (*Haynes RB, BMJ 2002*) per "prendere in considerazione" non si intende un invito a somministrare un trattamento, ma a verificarne la maggiore o minore opportunità in base alla malattia presente (caratteristiche biopatologiche, stadio, rischio di ripresa o di progressione, ecc), alle condizioni del paziente (PS, età, comorbidità, ecc) e alle sue preferenze/attese/valori.

L'opportunità di formulare la raccomandazione clinica in tale modo "non categorico" è inoltre condivisa da NICE (v. *NICE Guideline Manual, 2012*)

- Riguardo alla possibilità di considerare l'uso delle risorse per soglia di sostenibilità (analisi dei costi per beneficio) come outcome (di beneficio o di danno), il metodo GRADE consente di tralasciare tale opzione quando non vi siano le evidenze a supporto di una corretta analisi costo/efficacia (confronto dei costi e degli esiti di diverse alternative di trattamento) (*Guyatt GH, 2008; Brunetti M et al, 2013*)
  Non verranno quindi presi in considerazione aspetti di tipo farmaco-economico (applicando l'EtD non verrà quindi valutata la dimensione costo-efficacia a meno che evidenze del genere siano trovate in fase di selezione degli studi).
- Non saranno considerati nelle raccomandazioni e neppure nelle flow chart diagnostico-terapeutiche farmaci non ancora autorizzati da AIFA. Nel caso di farmaci in CNN (fascia C non negoziata), ne sarà data informazione esplicita nel testo.
- Per alcune raccomandazioni, formulate dal 2012 al 2015, la valutazione della qualità delle evidenze è stata eseguita applicando le vecchie regole SIGN Scottish Intercollegate Guidelines Network (SIGN) (vedi "Come leggere le raccomandazioni").
  Il passaggio della valutazione delle evidenze secondo SIGN a metodologia GRADE sarà graduale nel corso degli anni, e riguarderà solo quei quesiti "ancora aperti". Ovvero dove la produzione scientifica sia ancora attiva e dove non si sia raggiunto un punto fermo da parte della comunità scientifica.

D. REVISIONE DEI DOCUMENTI

I revisori esterni sono chiamati a rivedere la Linea Guida, per migliorarne la qualità, agendo direttamente sull'intero testo, suggerendo eventuali riconsiderazioni/modifiche/aggiunte/eliminazioni.

E. INDIPENDENZA E CONFLITTO DI INTERESSE

Le LG AIOM non ricevono alcun supporto diretto ed esplicito, in particolare da soggetti aventi interessi economici nella materia oggetto alle raccomandazioni.
Inoltre, al momento dell'accettazione dell'incarico, tutti i partecipanti ai Gruppi di Lavoro delle varie Linee Guida AIOM sono tenuti, dal 2013, ad esplicitare possibili conflitti di interesse.
I panelisti si astengono dalla votazione della forza della raccomandazione nei seguenti casi:

- ✓ quando fanno parte dell'*authorship* di uno o più lavori considerati per la raccomandazione
- ✓ quando hanno ricevuto finanziamenti diretti o indiretti da Aziende farmaceutiche titolari dell'intervento che si sta prendendo in esame

F. AGGIORNAMENTO PERIODICO

Le LG AIOM vengono aggiornate annualmente, con divulgazione in concomitanza al congresso nazionale AIOM, in autunno.
L'abolizione dei documenti cartacei e la pubblicazione sul sito AIOM rende possibile l'aggiornamento in tempo reale nel caso di variazioni importanti per la pratica clinica.

G. PUBBLICAZIONE NEL SITO AIOM

Le LG AIOM vengono pubblicate sul sito AIOM (www.aiom.it) e sono scaricabili sia dai soci AIOM che dai non soci.
E' inoltre disponibile una APP per i più diffusi sistemi operativi in mobilità.

H. VALUTAZIONE DELLA IMPLEMENTAZIONE

AIOM ha attivato nel corso degli anni una serie di progetti di valutazione dell'implementazione delle LG nella pratica clinica (Progetti RIGHT), basati sulla verifica dei record clinici nel periodo susseguente la pubblicazione delle LG.
E' prevista per il futuro l'istituzione di audit in selezionate strutture di cura, con valutazione dei record clinici prima e dopo la pubblicazione delle LG di un dato anno e riscontro di eventuali modifiche nella pratica clinica quotidiana.

Figura 2. Progetti RIGHT

3. CONFRONTO CON LE LINEE GUIDA INTERNAZIONALI

Le raccomandazioni prodotte da AIOM utilizzando elementi del metodo GRADE rappresentano un punto qualificante nel panorama internazionale delle LG fornite da organizzazioni scientifiche, prodotte in larga misura attraverso l'utilizzo di revisioni non sistematiche e tramite il consenso tra esperti.
Le LG AIOM non pongono il giudizio di qualità sul solo disegno di studio (vedi Linee Guida NCCN – versione *standard*) bensì sul rischio di bias, sulla eterogeneità (vedi Linee Guida ESMO) e sulla trasferibilità dei dati (estrapolazione delle evidenze) per ogni outcome critico o importante per la proposta clinica.

In questi 5 anni AIOM ha avviato un profondo processo di revisione - tuttora in corso - del modo di proporre le LG, che ha permesso di ottenere dei risultati importanti ma ancora perfettibili.
Lo sforzo di questi anni ha infatti consentito di ottenere raccomandazioni basate su revisioni esaustive della letteratura (soprattutto, ma non solo, per i quesiti clinici affrontati con il metodo GRADE) e con valutazioni strutturate ed esplicite della qualità dei dati così da rendere trasparente il percorso metodologico che conduce alla forza della raccomandazione a partire dalla qualità delle evidenze. Inoltre la modalità di lavoro dei Gruppi è stata resa il più omogenea possibile così che il lettore abbia la possibilità di muoversi facilmente tra le raccomandazioni intra o inter-linee guida.

Nella tabella seguente si può dedurre come, rispetto agli standard metodologici che fanno capo ai criteri AGREE II (link alla versione italiana: http://www.gimbe.org/pubblicazioni/traduzioni/AGREE_IT.pdf), la validità delle LG AIOM è dimostrabile.

Tabella 2. Standard metodologico per lo sviluppo di linee guida (LG) – AGREE II

		LG AIOM		Commenti
		"SIGN"	"GRADE"	
1	Obiettivo generale della LG chiaramente espresso	√		Espresso nel manuale metodologico
2	Quesiti clinici chiaramente definiti		√	
3	Popolazione (pazienti, medici, altri stakeholders, etc.) a cui si riferisce la LG chiaramente espressa			Popolazione target: oncologo (raramente esistono figure diverse)
4	Panel costituiti da tutte le figure professionali rilevanti		√	Figure diverse dall'oncologo non in tutte le LG
5	La visione o le preferenze della popolazione target (pazienti, medici, altri stakeholders, etc.) inclusa		√	Non in tutte le LG
6	Utilizzatori finali (pazienti, medici, altri stakeholders, etc.) della LG chiaramente esplicitati			Popolazione target: oncologo (raramente esistono figure diverse)
7	Metodi sistematici per l'identificazione delle evidenze disponibili		√	Per GRADE: Esplicita ricerca bibliografica
8	Criteri espliciti per selezionare le evidenze		√	Per GRADE: criteri di inclusione/esclusione evidenze esplicitati in corso d'opera; quasi mai a priori (quesiti a volte poco chiari)
9	Descrizione dei punti di forza e delle limitazioni del corpo delle evidenze valutati secondo standard metodologici (qualità delle evidenze)	√	√	Escludendo i quesiti risolti con GRADE, poche LG si attengono alla valutazione metodologica del corpo delle evidenze come espresso nel manuale operativo
10	Chiara descrizione del metodo per formulare la raccomandazione		√	Definizione del quesito (PICO); strategia di ricerca sui motori adeguati; valutazione delle voci bibliografiche selezionate; valutazione qualità metodologica; sintesi delle evidenze (tabelle delle evidenze)
11	Benefici e danni considerati nella formulazione della raccomandazione	√	√	Escludendo il GRADE, le raccomandazioni espresse raramente rispecchiano quanto descritto nel testo
12	Link espliciti tra raccomandazioni ed evidenza a supporto	√		
13	LG riviste da esperti esterni prima della pubblicazione	√		
14	Procedure di aggiornamento continuo	√		
15	Raccomandazioni specifiche e non ambigue	√		
16	Opzioni differenti per il trattamento della condizione in oggetto prese in considerazione	√		Nel testo vengono elencate tutte le opzioni terapeutiche senza giustamente esprimere un ranking di preferenza
17	Raccomandazioni chiave facilmente identificabili			Richiesto ma non implementato dalla quasi totalità delle LG
18	Descrizione della facilità/difficoltà di implementazione della LG			
19	Esplicitazione di suggerimenti o strumenti per portare le LG dalla ricerca alla pratica			
20	Considerazione dei potenziali costi di applicazione delle LG			
21	Criteri di audit e monitoraggio della diffusione e utilizzo della LG	√		Attraverso il progetto RIGHT
22	Opinione dei finanziatori non influenza il contenuto delle LG	√		
23	Conflitto di interesse registrato e disponibile	√		Disponibile

Il confronto delle LG di associazioni scientifiche europee e internazionali mostra inoltre il buon grado di aderenza agli standard delle LG AIOM, seconda solo alle attuali LG ASCO, radicalmente cambiate negli ultimi 2 anni.

Tabella 3. Confronto tra le Linee Guida AIOM ed altre Linee Guida internazionali

		AIOM	ECCO	ESMO	ASCO	NCCN
1	Obiettivo generale della LG chiaramente espresso	√	√	√	√	NA
2	Quesiti clinici chiaramente definiti	√*	√		√	NA
3	Popolazione (pazienti, medici, altri stakeholders, etc.) a cui si riferisce la LG chiaramente espressa	NR	NR	NR	√	NR
4	Panel costituiti da tutte le figure professionali rilevanti	√[&]	NR	√	√	√
5	La visione o le preferenze della popolazione target (pazienti, medici, altri stakeholders, etc.) inclusa	√[&]	NR	NR	√	NR
6	Utilizzatori finali (pazienti, medici, altri stakeholders, etc.) della LG chiaramente esplicitati	NR	NR	NR	√	NR
7	Metodi sistematici per l'identificazione delle evidenze disponibili	√*	NR	NR	√	NR
8	Criteri espliciti per selezionare le evidenze	√*	NR	NR	√	NR
9	Descrizione dei punti di forza e delle limitazioni del corpo delle evidenze valutati secondo standard metodologici (qualità delle evidenze)	√*	NR	NO	√	NO
10	Chiara descrizione del metodo per formulare la raccomandazione	√*	√	√	√	NR
11	Benefici e danni considerati nella formulazione della raccomandazione	√*	NR	√	√	NA
12	Link espliciti tra raccomandazioni ed evidenza a supporto	√	NR	√	√	NA
13	LG riviste da esperti esterni prima della pubblicazione	√	NR	NR	√	NR
14	Procedure di aggiornamento continuo	√	√	√	√	√
15	Raccomandazioni specifiche e non ambigue	√	√	√	√	NA
16	Opzioni differenti per il trattamento della condizione in oggetto prese in considerazione	√	√	√	√	NA
17	Raccomandazioni chiave facilmente identificabili	√	√	√	√	NA
18	Descrizione della facilità/difficoltà di implementazione della LG	NR	NR	NR	√	NR
19	Esplicitazione di suggerimenti o strumenti per portare le LG dalla ricerca alla pratica	NR	NR	√	√	NR
20	Considerazione dei potenziali costi di applicazione delle LG	NR	NR	NR	√	NR
21	Criteri di audit e monitoraggio della diffusione e utilizzo della LG	√	NR	√	√	NR
22	Opinione dei finanziatori non influenza il contenuto delle LG	√	NR	NR	NR	NR
23	Conflitto di interesse registrato e disponibile	√	√	√	√	√

*Solo per il processo GRADE; [&]solo per alcune LG (manca il paziente)

NR: non riportato; NA: non accessibile

4. BIBLIOGRAFIA

1- Schünemann HJ, Fretheim A, Oxman AD; WHO Advisory Committee on Health Research.Improving the use of research evidence in guideline development: 1. Guidelines for guidelines. Health Res Policy Syst. 2006 Nov 21;4:13.

2- Haynes BR: Physycians' and patients' choices in evidence based practice. BMJ 2002; 324:1350.

3- Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A, Vist GE, Schünemann HJ; GRADE Working Group.Incorporatingconsiderations of resources use intogradingrecommendations.BMJ. 2008 May24;336(7654):1170-3. doi: 10.1136/bmj.39504.506319.80. Review.

4- Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J, Sisk J, Ruiz F, Hill S, Guyatt GH, Jaeschke R, Helfand M, Harbour R, Davoli M, Amato L, Liberati A, Schünemann HJ.GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence.J ClinEpidemiol. 2013 Feb;66(2):140-50. doi: 10.1016/j.jclinepi.2012.04.012. Epub 2012 Aug 3.

METODOLOGIA APPLICATA ALLE LINEE GUIDA AIOM

Michela Cinquini, Ivan Moschetti, Giovanni Pappagallo, Valter Torri

1. SCOPO E CAMPO DI APPLICAZIONE

Il manuale operativo ha la finalità di supportare i Gruppi di Lavoro (WG) nella stesura delle Linee Guida (LG) al fine di allinearne i contenuti, la valutazione delle evidenze scientifiche, la formulazione delle raccomandazioni e degli altri contenuti ritenuti essenziali, compreso l'editing.
Recentemente SIGN (Scottish Intercollegate Guideline Network) ha adottato gli elementi del GRADE (Grading Recommendations Assessment Development and Evaluation) Working Group per la valutazione delle evidenze. Una indicazione di questo tipo non può far altro che rinforzare il nostro approccio alla produzione di raccomandazioni per la pratica clinica, poiché l'integrazione tra i due sistemi è sempre stata la proposta di questo gruppo metodologico.
Proprio per questo motivo, dal 2016 AIOM ha deciso di abbandonare la valutazione delle evidenze secondo SIGN per concentrarsi su GRADE e affrontare poco alla volta tutti i quesiti clinici rilevanti con quest'ultimo sostituendolo gradualmente a SIGN.
Il documento è suddiviso in due parti: la prima relativa alla metodologia scientifica e la seconda relativa alle istruzioni per l'editing.

2. METODOLOGIA SCIENTIFICA

2.1. Formulazione del Quesito Clinico

Per la definizione precisa del campo di applicazione è fondamentale una corretta formulazione del quesito che, tendenzialmente, non dovrebbe essere generica ("qual è il ruolo di X nel trattamento della patologia Y"), bensì aderente al PIC(O) ovvero Pazienti, Intervento, Confronto, e (Outcome). Outcome è messo tra parentesi poiché non è sempre spendibile direttamente nel quesito, spesso infatti gli outcome di interesse sono multipli e non risulterebbe intellegibile inserirne una lista al termine del quesito.
Esempio di quesito PIC(O):

- **Nei pazienti con** (*menzionare le specifiche caratteristiche di malattia, stadio, o ecc.*)
- **il trattamento con** (*descrivere l'intervento terapeutico oggetto del quesito*)
- **è suscettibile di impiego in alternativa a** (*descrivere il trattamento altrimenti considerato in alternativa all'intervento in esame*)?
- **(con riferimento agli *outcome* di beneficio e di danno)** (*elencare i parametri clinico-laboratoristici – misure di effetto, outcome – ritenuti essenziali per la proposta terapeutica*)

Come sopra descritto gli outcome possono essere declinati di seguito al quesito.
La fase di individuazione degli outcome pertinenti al quesito clinico dovrebbe idealmente coinvolgere l'intero gruppo di esperti e non (panelisti) che sono chiamati a formulare la raccomandazione. In questa fase devono essere individuati sia outcome positivi ("benefici") che negativi ("danni").

Nel metodo GRADE, una volta scelti, gli outcome verranno classificati in termini di rilevanza:

- *importanti ed essenziali (indicati come "criticiper poter prendere una decisione clinica)")*
- *importanti ma non essenziali*
- *non importanti*

Tale metodo suggerisce di effettuare la valutazione della rilevanza degli outcome mediante votazione individuale dei panelisti, utilizzando una scala a 9 punti e assegnando l'outcome a una delle tre categorie sopra descritte sulla base del punteggio mediano ottenuto. In base alla categoria nella quale sono stati classificati, gli outcome saranno o meno considerati nella tabella sinottica delle evidenze e successivamente nella formulazione della raccomandazione, secondo i criteri riportati in **Tab.1**.

Rating **(mediana del voto)**	**Importanza**	**Incluso in**
7 8 9	*outcome* importanti ed essenziali	tabelle sulla qualità delle prove: SÌ raccomandazione: SÌ
4 5 6	*outcome* importanti ma non essenziali	tabelle sulla qualità delle prove: SÌ raccomandazione: NO
1 2 3	*outcome* non importanti	tabelle sulla qualità delle prove: NO raccomandazione: NO

Tab.1: Classificazione degli outcome

2.2. Selezione della letteratura a supporto delle raccomandazioni cliniche

La strategia di ricerca

I panelisti che sono chiamati a formulare la raccomandazione dovrebbero avere a disposizione le migliori informazioni disponibili nella letteratura scientifica.
A tal fine, si rende necessario effettuare ricerche sistematiche mirate con lo scopo di individuare tutti gli studi disponibili che hanno affrontato il quesito.
Lo sviluppo di una strategia di ricerca sensibile, riproducibile e che dia risultati gestibili necessita di figure formate *ad hoc*, i bibliometristi.
Una volta ottenuti gli articoli individuati dalla strategia di ricerca il panel deve selezionare le evidenze migliori. Se non ci sono revisioni sistematiche di studi clinici randomizzati (RCTs) pertinenti il quesito si passerà agli studi singoli o alle revisioni sistematiche (RS) di studi osservazionali etc, seguendo in sintesi la nota piramide delle evidenze.
In media, è relativamente più semplice individuare risultati solidi per outcome di efficacia (overall survival, progression-free survival, etc), mentre è meno comune trovare studi adeguatamente dimensionati e con follow-up sufficiente per gli outcome di danno, soprattutto nel lungo periodo. Un suggerimento può essere quello di cercare gli esiti negativi in studi osservazionali o serie di casi o case report, etc. che possono essere disponibili anche per malattie diverse da quella considerata ma legati al trattamento.
La strategia di ricerca deve essere esplicitata, disponibile (in appendice o *on demand*) a quanti interessati a visionarla e al limite a rilanciarla in futuro nell'ottica di eventuali aggiornamenti. La presenza (e la tracciabilità) di una strategia di ricerca sviluppata specificamente per un quesito o pluri-quesito è un elemento di qualità di una raccomandazione ed è requisito previsto dagli strumenti internazionali (checklist) che misurano la validità di una LG..
Durante il processo di produzione di raccomandazioni con **il metodo GRADE**, viene quindi effettuata una ricerca bibliografica (esaustiva, sensibile e riproducibile) delle fonti su diversi database medico-scientifici (PubMed, Embase, CENTRAL e database area-specifici).
L'interrogazione di ogni database segue regole specifiche per ciascuno di essi.
In PubMed, ad esempio, le parole chiave vengono cercate prima attraverso il dizionario MESH e poi in "ricerca libera" (utilizzando l'operatore booleano OR per aumentare la sensibilità), avvalendosi dei diversi

strumenti messi a disposizione dal database tra cui i filtri per limitare la ricerca (fasce di età, tipo di disegno di studio, tipologia di soggetti inclusi, etc.) (vedi Manuale Strategie di ricerca).

2.3. Valutazione critica della letteratura a supporto delle raccomandazioni cliniche

2.3.1. La descrizione degli studi per esteso

E' auspicabile che ogni studio considerato, essenziale per il supporto a una raccomandazione, venga descritto nel testo in maniera sintetica ma esaustiva.
È consigliato iniziare riportando i risultati principali (effetto netto in misura sia relativa sia assoluta, relativamente ai benefici e ai danni osservati)[1]. Si suggerisce di proseguire descrivendo la metodologia e soprattutto le problematiche relative alla conduzione dello studio valutando in sequenza i possibili bias. Occorre tener presente che ogni errore sistematico (bias) individuato andrebbe in qualche modo relazionato alla sua ripercussione sul risultato, in quanto non tutti i bias contribuiscono in egual misura a rendere i risultati più o meno credibili/affidabili.
Al termine della descrizione dello studio, nel testo e solo per i quesiti che ancora non saranno affrontati con metodologia GRADE, rimarrà il livello di evidenza (1,2,3,4) corredato dal rischio di bias (++,+,-) e indicato se il PICO dello studio è sufficientemente vicino/sovrapponibile o distante rispetto al PICO del quesito di interesse.
Una raccomandazione può essere integrata da commenti e può essere suddivisa in sotto-raccomandazioni dove è possibile specificare meglio altre caratteristiche salienti del paziente, setting, modalità di somministrazione, dose, tempi, monitoraggio effetti collaterali specifici, etc. Ovvero tutto quello che può aiutare il clinico a proporre l'intervento (farmacologico, chirurgico, etc.).
Uno degli obiettivi più importanti delle LG /raccomandazioni è il trasferimento "locale", ossia a livello nazionale o regionale delle evidenze prodotte "around the world".
Da ricordare quando si legge una raccomandazione la cui evidenza è stata valutata con SIGN:

- se le evidenze non provengono da RCT o meta-analisi con basso rischio di bias (1++), oppure da studi epidemiologici osservazionali con basso rischio di bias (2++) nella descrizione dello studio dovrebbero essere riportate le considerazioni metodologiche principali che hanno portato all'abbassamento della qualità delle evidenze ovvero a 1+ o 2+.
- i livelli di evidenza si riferiscono a uno studio specifico, non a gruppi di studi. Di conseguenza, ogni studio descritto e ritenuto rilevante per la raccomandazione ha il livello di evidenza specificato nel testo.
- le stesse referenze essenziali sono riportate nel box dove viene riportata la raccomandazione corredata dalla sua forza.

Importante: indipendentemente dal disegno di studio (indicato da: 1 o 2), se il livello di qualità risulta negativo **($^-$)** non sarà prodotta alcuna raccomandazione, anche nel caso questa fosse negativa (es: *l'intervento non deve/non dovrebbe essere somministrato*), poiché con il segno **($^-$)** si intendono "risultati non affidabili". In altre parole, lo studio è ad alto rischio di bias, quindi i risultati non sono attendibili sia in senso positivo sia negativo, l'incertezza è totale, studi successivi sullo stesso argomento potrebbero confermare o ribaltare completamente risultato e dimensione.

2.3.2. Qualità dell'evidenza (SIGN) – da considerare solo per la lettura delle raccomandazioni prodotte fino al 2015

Nell'approccio SIGN, la qualità delle evidenze (singoli studi /revisioni sistematiche / meta-analisi, etc.) a sostegno della raccomandazione veniva valutata tenendo conto sia del disegno dello studio (**Fig.1**) sia di come esso è stato condotto: la minore o maggiore entità del rischio di *bias* associabile a uno studio è indicata rispettivamente con le notazioni ($^{++}$) (*rischio molto basso*), ($^{+}$) (*rischio basso*) ($^{-}$), (*rischio elevato*). Il

[1] descrizione sintetica dei risultati in termini relativi (HR, RR, OR) o assoluti (RD, NNT o mesi guadagnati, ecc.). Se disponibile, riportare la lunghezza del follow-up.

Livello di Evidenza SIGN (**Tab.2**) viene riportato nel testo al termine della descrizione di ciascuno studio ritenuto rilevante e essenziale a sostegno o contro l'esecuzione di uno specifico intervento.
I livelli di evidenza dovranno essere precisati solo per le evidenze (studi) descritti e che sono essenziali per la raccomandazione clinica e che quindi contribuiscono a formare il giudizio della Qualità Globale delle Evidenze SIGN.
La ***Qualità Globale delle Evidenze SIGN*** (**Tab.3**) viene quindi riportata con lettere (A, B, C, D) e corrisponde a un giudizio sintetico dei singoli studi, unitamente all'indicazione sulla diretta applicabilità delle evidenze sulla popolazione target della raccomandazione.
Ogni lettera indica in sintesi la **"fiducia"** nell'intero corpo delle evidenze valutate che sostengono la raccomandazione; **NON** riflettono sempre l'importanza clinica della raccomandazione e **NON** sono sinonimo della forza della raccomandazione clinica.

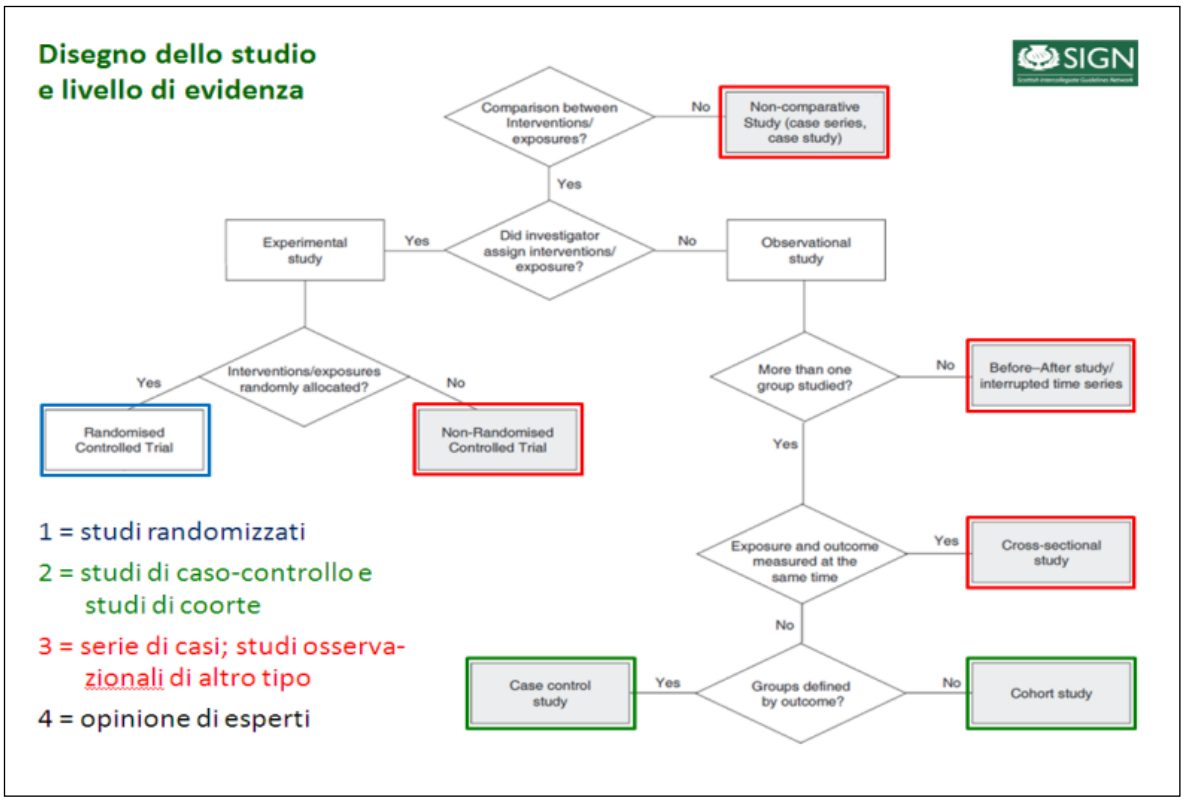

Fig.1

Tab.2: Livelli di Evidenza SIGN

1	Revisioni sistematiche e meta-analisi di RCT o singoli RCT
1 ++	Rischio di bias molto basso.
1 +	Rischio di bias basso.
1 -	Rischio di Bias elevato -> i risultati dello studio non sono affidabili.
2	Revisioni sistematiche e meta-analisi di studi epidemiologici di caso/controllo o di coorte o singoli studi di caso/controllo o di coorte.
2 ++	Rischio di bias molto basso, probabilità molto bassa di fattori confondenti, elevata probabilità di relazione causale tra intervento e effetto.
2 +	Rischio di bias basso, bassa probabilità presenza fattori di confondimento, moderata probabilità di relazione causale tra intervento e effetto.

2 -	Rischio di Bias elevato -> i risultati dello studio non sono affidabili, esiste un elevato rischio che la relazione intervento/effetto non sia causale.
3	Disegni di studio non analitici come report di casi e serie di casi.
4	Expert opinion.

Tab.3: Qualità Globale delle Evidenze SIGN -fiducia espressa verso il corpo delle evidenze-

A	Almeno una meta-analisi o revisione sistematica o RCT valutato 1++ e direttamente applicabile alla popolazione target *oppure*
	Il corpo delle evidenze disponibili consiste principalmente in studi valutati 1+ direttamente applicabili alla popolazione target e con risultati coerenti per direzione e dimensione dell'effetto
B	Il corpo delle evidenze include studi valutati 2++ con risultati applicabili direttamente alla popolazione target e con risultati coerenti per direzione e dimensione dell'effetto.
	Evidenze estrapolate da studi valutati 1++ o 1+
C	Il corpo delle evidenze include studi valutati 2+ con risultati applicabili direttamente alla popolazione target e con risultati coerenti per direzione e dimensione dell'effetto.
	Evidenze estrapolate da studi valutati 2++
D	Evidenze di livello 3 o 4
	Evidenze estrapolate da studi valutati 2+

2.3.2.1. Considerazioni sulla "estrapolazione delle evidenze"

Si fa qui riferimento alla trasferibilità dei risultati dallo studio alla popolazione a cui è destinata la raccomandazione. Non sono quindi direttamente trasferibili evidenze con le caratteristiche riportate di seguito

a) studi eseguiti su pazienti in prima linea -> raccomandazioni prodotte per la II linea,
b) studi eseguiti in setting diversi da quelli dove sarà applicata la raccomandazione,
c) popolazione selezionata in base alla positività all'istochimica anziché su esami di analisi molecolare sul tessuto, etc.
d) differenze nei dosaggi tra lo studio e la pratica clinica
e) differenze di etnia
f) etc

Anche nel caso di un'evidenza 1 (studio randomizzato) senza evidenza di bias rilevanti ($^{++}$), nel caso la popolazione oggetto dello studio non sia sovrapponibile a quella oggetto del quesito clinico, la qualità globale dell'evidenza passerà da **A**[2] a **B**[3].

[2] RCT senza evidenza di rischi di bias, i cui risultati sono direttamente applicabili alla popolazione per cui è prodotta la raccomandazione

[3] RCT senza evidenza di rischi di bias, ma con incertezza rispetto alla trasferibilità dei risultati alla popolazione per cui è prodotta la raccomandazione

2.4. Qualità dell'evidenza (GRADE)

Il GRADE prevede una valutazione della qualità delle evidenze per ogni outcome scelto dal panel come critico o importante. Il sistema inoltre nasce per dare una visione di insieme delle evidenze disponibili (body of evidence), privilegiando le sintesi esaustive e sistematiche della letteratura per rispondere a una singola domanda come le revisioni sistematiche. Le dimensioni che influenzano la qualità del singolo outcome sono riportate di seguito:

- ✓ ***Study design*** - *errori nella pianificazione e conduzione dello studio*: sono fattori legati strettamente al modo con il quale lo studio è stato condotto e la probabilità che eventuali errori fatti in corso di studio abbiano potuto determinare *bias* nelle stime di efficacia o di occorrenza di effetti avversi; per gli studi randomizzati i principali limiti di questo tipo sono:
  - truffe, errori o insufficiente attenzione alla assegnazione in cieco al braccio di trattamento o di controllo (*allocation concealment*). La lista di distribuzione dei pazienti (allocazione) ai diversi trattamenti deve essere nascosta (concealed) bene e il metodo con cui è stata generata la lista casuale di allocazione deve essere adeguato (*selection bias*).
  - mancanza o problemi nella effettiva realizzazione della cecità, specie per outcome soggettivi (*detection bias*, in casi rari *performance bias*)
  - perdita al follow-up di una quota importante di pazienti originariamente inclusi negli studi o perdite al follow-up asimmetriche nei due gruppi (*attrition bias*)
    - esclusione dall'analisi dei soggetti/pazienti persi al follow-up in diversi momenti dello studio (violazione del principio della *intention to treat*)
  - interruzioni precoci degli studi a seguito di un eccesso di efficacia o tossicità secondo modalità non previste dal protocollo
    - altri bias outcome specifici: *crossover.*
  - *selective outcome reporting bias*. Probabilità che un outcome sia pubblicato rispetto a un altro dovuta al fatto che il risultato sia positivo (statisticamente significativo). Avere a disposizione il protocollo dello studio può servire da verifica.

- ✓ **Precision** – valuta la *precisione delle stime*:
  - i risultati sono imprecisi quando gli studi includono pochi pazienti e/o si verificano pochi eventi producendo così limiti di confidenza della stima della misura di effetto ampi che possono essere compatibili con conclusioni cliniche di segno opposto;
  - nel caso in cui la frequenza dell'evento sia molto bassa, con ampi intervalli di confidenza della stima relativa di effetto e stretti intervalli di confidenza della stima assoluta di effetto, è possibile non abbassare il livello di qualità della prova.

- ✓ **Directness** – valuta la *diretta applicabilità delle evidenze (P.I.C.O.)*
  - popolazione, intervento, controllo o outcome indiretti: il quesito per il quale si intende porre la raccomandazione si riferisce a una popolazione, intervento, controllo o outcome diversi da quelli per i quali sono disponibili prove di efficacia in letteratura
  - confronto indiretto: quando non sono disponibili studi che confrontano direttamente l'intervento A con l'intervento B ma solo studi che confrontano A con C e B con C; questi studi permettono un confronto solo indiretto di A con B e quindi la qualità della prova sarà giudicata inferiore
    - ***NB*** la definizione compiuta del quesito clinico con le dimensioni PICO è molto utile per valutare la trasferibilità più o meno diretta delle prove disponibili; infatti la *indirectness* dei risultati dipende in larga misura dalla distanza tra gli studi disponibili e il quesito al quale si vuole rispondere
    - Se infatti si affrontano, all'interno della stessa raccomandazione, quesiti che riguardano sottogruppi, il giudizio sulla *indirectness* potrebbe cambiare in funzione di una maggiore o minore focalizzazione del PICO.

- ✓ ***Consistency*** – *valuta la coerenza dei risultati tra studi differenti eseguiti con il medesimo obiettivo*: questo criterio si applica solamente all'insieme della letteratura disponibile e non al singolo studio; si manifesta come una eterogeneità delle stime di effetto tra gli studi che non trova spiegazione logica (diversità nel tipo di intervento o nella composizione delle popolazioni studiate) e che quindi aumenta l'incertezza sulla reale entità dell'effetto dell'intervento (si applica solo nella valutazione di una revisione sistematica).
- ✓ ***Publication bias*** – valuta la presenza di **pubblicazione selettiva dei dati**: tendenza secondo cui gli studi pubblicati differiscono sistematicamente dagli studi non pubblicati, i quali è più probabile che riportino risultati statisticamente non significativi o negativi.
    - Si suggerisce di considerare e discutere esplicitamente questa possibilità e di esaminare, caso per caso, se in presenza di forte sospetto di tale bias sia opportuno procedere o meno a un *downgrading* della qualità.

Con il metodo GRADE la qualità delle prove viene considerata separatamente per ciascun outcome classificato almeno come "importante", costruendo tabelle sintetiche dette "*evidence profile o evidence table o ToEs*) che riportano per singolo outcome la qualità delle prove e la sintesi quantitativa degli effetti.
A seconda del disegno di studio considerato la qualità dell'evidenza parte da un certo livello (alta per RCT, molto bassa per studi osservazionali).
In **Tab.4** sono presentati i criteri per l'aumento (*upgrading*) o la diminuzione (*downgrading*) del giudizio di qualità (alta, moderata, bassa, molto bassa) delle prove, il cui significato, comprensivo delle conseguenze interpretative, è sintetizzato in **Tab.5**.

Tab.4: criteri per l'aumento o la diminuzione del giudizio sulla qualità delle prove

Tipo di prove	**Studio controllato e randomizzato = alta** **Studio osservazionale = bassa** **Qualsiasi altro tipo di informazione = molto basso**
A. Diminuzione della categoria di attribuzione (es. da "alta" a "moderata")	1. Limiti gravi (-1 livello) o molto gravi (-2 livelli) nella qualità di conduzione dello studio 2. Incoerenza nei risultati tra studi diversi sullo stesso quesito (-1 o -2 livelli) 3. Alcune (-1 livello) o importanti (-2 livelli) incertezze circa la diretta trasferibilità dei risultati (*directness*) 4. Imprecisione o dati insufficienti (*sparse data*) (-1 o -2 livelli) 5. Possibilità di pubblicazione selettiva dei dati (*publication* e *reporting bias*) (-1 o -2 livelli)
B. Aumento della categoria di attribuzione (es. da "bassa" a "moderata")	1. Associazione intervento-*outcome* forte, ovvero con rischio relativo >2 (<0,5), sulla base di prove concordanti provenienti da due o più studi osservazionali, senza alcun fattore di confondimento plausibile (+1 livello) 2. Associazione intervento-*outcome* molto forte, ovvero con rischio relativo >5 (<0,2) (+2 livelli) 3. Presenza di un gradiente dose-risposta (+1 livello) 4. Tutti i possibili fattori di confondimento che avrebbero potuto alterare le stime di effetto avrebbero ridotto l'effetto che si osserva (+1 livello)

Tab.5: Significato e conseguenze dei diversi livelli di qualità delle prove

Livello qualità	Significato	Conseguenza
Alta	Alto grado di confidenza nei risultati	È molto improbabile che ulteriori studi possano cambiare la fiducia nella stima di effetto
Moderata	Discreto grado di confidenza nei risultati	È probabile che ulteriori studi possano confermare o cambiare la fiducia nella stima di effetto
Bassa	I risultati sono poco credibili	È necessaria ulteriore ricerca per ottenere stime affidabili sugli effetti positivi e negativi dell'intervento
Molto bassa	I dati esaminati sono totalmente inaffidabili	Non è possibile fare affidamento sulle stime di effetto disponibili

A seguire la graduazione della qualità per i singoli outcome importanti si deve formulare il giudizio complessivo di qualità.
Il metodo GRADE suggerisce di procedere considerando soltanto gli outcome essenziali (critici) per la formulazione della raccomandazione relativa al quesito clinico.
Se la qualità è diversa fra i singoli outcome essenziali, il metodo suggerisce la seguente linea di comportamento:

- ✓ se i risultati vanno in direzioni opposte (es. il trattamento oggetto della raccomandazione è di qualità migliore in termini di efficacia ma peggiore per quanto riguarda gli effetti indesiderati), la qualità globale viene attribuita basandosi sulla valutazione peggiore ossia assumendo come più rappresentativo l'outcome critico (essenziale per la proposta clinica) che ha ottenuto la più bassa valutazione di qualità;
- ✓ se i risultati vanno nella stessa direzione per tutti gli outcome (beneficio o danno), viene assunta come qualità globale delle prove la qualità dell'outcome essenziale che da solo basterebbe per formulare compiutamente la raccomandazione. [G. Guyatt et al. / Journal of Clinical Epidemiology 66 (2013) 151e157]

2.4.1. Bilancio tra benefici e danni

La direzione della raccomandazione a favore o contro l'uso di un determinato intervento/farmaco/trattamento/comportamento, etc. si dovrebbe basare sul bilancio tra gli effetti positivi (esiti benefici) e negativi (effetti dannosi) di tale intervento.
La decisione sul bilancio tra effetti positivi e effetti negativi deve tenere conto del numero e del peso dei singoli fattori.
Il peso di ciascun effetto positivo o negativo è inoltre influenzato dai seguenti elementi:

- ✓ importanza degli outcome: sebbene solo gli outcome essenziali siano considerati nel processo che porta a formulare la raccomandazione, non è detto che tutti abbiano lo stesso peso. In altre parole, anche all'interno della categoria di outcome essenziali, potrebbe essere utile definire una gerarchia tale da influenzare la proposta terapeutica;
- ✓ rischio di base degli eventi che l'intervento oggetto della raccomandazione dovrebbe essere in grado di ridurre; in generale maggiore è il rischio di base, maggiore è la probabilità che l'intervento considerato dalla raccomandazione abbia un impatto rilevante;
- ✓ entità dell'effetto relativa e assoluta (rilevanza clinico-epidemiologica): in presenza di studi di buona qualità che documentano effetti positivi ampi e rilevanti sul piano clinico o un chiaro prevalere di effetti negativi su quelli positivi, si arriverà più facilmente alla formulazione di raccomandazioni "forti" (a favore o contro l'uso dell'intervento, rispettivamente).

La definizione del rapporto beneficio/danno viene effettuata dal panel mediante votazione formale sulle seguenti alternative:

- ✓ favorevole
- ✓ incerto (a favore o a sfavore dell'intervento)
- ✓ sfavorevole

2.4.2. Evidence to Decision Framework (EtD)

Attraverso lo sviluppo dell' EtD framework verranno sintetizzate e riportate le prove rispetto alla priorità del problema, entità degli effetti desiderabili ed indesiderabili, certezza nelle prove, valori e preferenze dei pazienti, bilancio tra effetti desiderabili ed indesiderabili, risorse economiche, accettabilità, equità e fattibilità degli interventi a confronto.

Scopo principale dell'EtD è quello di informare in modo strutturato e trasparente i panel a prendere delle decisioni e a formulare le raccomandazioni finali.

3. Raccomandazioni cliniche[4]

Esprimono l'importanza clinica di un intervento, e dovrebbero quindi essere formulate sulla base del P.I.C.O. del quesito (popolazione, intervento, confronto, outcome). La forza della raccomandazione viene graduata, in base all'importanza clinica, su 4 livelli **(Tab.6)**:

Tab.6: Tipi di raccomandazione clinica

Forza della raccomandazione clinica	Terminologia	Significato
Positiva Forte	"Nei pazienti con (*criteri di selezione*) l'intervento xxx **dovrebbe** essere preso in considerazione come opzione terapeutica di prima intenzione"	l'intervento in esame dovrebbe essere considerato tra le opzioni di prima scelta terapeutica (evidenza che i benefici sono prevalenti sui danni)
Positiva Debole	"Nei pazienti con (*criteri di selezione*) l'intervento xxx **può** essere preso in considerazione come opzione terapeutica	l'intervento in esame può essere considerato come opzione di prima intenzione, consapevoli dell'esistenza di alternative ugualmente proponibili (incertezza riguardo alla prevalenza dei benefici sui danni), molto utile una discussione approfondita con il pz allo scopo di chiarire al meglio i valori espressi o meno dalla persona oggetto del trattamento
Negativa Debole	"Nei pazienti con (*criteri di selezione*) l'intervento xxx **non dovrebbe** essere preso in considerazione come opzione terapeutica	l'intervento in esame non dovrebbe essere considerato come opzione di prima intenzione; esso potrebbe comunque essere suscettibile di impiego in casi altamente selezionati e previa completa condivisione

[4] A inizio capitolo sarebbe auspicabile che ogni LG inserisca un paragrafo contenente le raccomandazioni chiave

		con il paziente (incertezza riguardo alla prevalenza dei danni sui benefici)
Negativa Forte	"Nei pazienti con (*criteri di selezione*) l'intervento xxx **non deve** essere preso in considerazione	l'intervento in esame non deve essere in alcun caso preso in considerazione (evidenza affidabile che i danni sono prevalenti sui benefici)

La forza della raccomandazione riflette la misura in cui si ritiene che, nella popolazione specifica del quesito clinico, gli effetti benefici derivanti dal seguire la raccomandazione superino gli effetti indesiderati (o viceversa per le raccomandazioni negative).

- ✓ Quando si decide per una raccomandazione "forte" si deve essere certi dei fattori che influenzano la forza della raccomandazione.
  - una raccomandazione "forte" dovrebbe essere riservata a situazioni nelle quali si è molto convinti che la maggioranza dei soggetti che ricevono l'intervento oggetto della raccomandazione ottengono un beneficio.
- ✓ Quando esiste incertezza sul rapporto beneficio/danno, o quando le informazioni rilevanti che influenzano la forza della raccomandazione non sono disponibili o non sono chiare, si dovrebbe optare per una raccomandazione "debole".
  - le raccomandazioni "deboli" sono quelle per cui gli effetti benefici probabilmente prevalgono sugli effetti dannosi (o viceversa per le raccomandazioni negative) ma c'è ancora una non trascurabile incertezza.
  - una raccomandazione "debole" può implicare che si debba valutare attentamente in quali condizioni o a quali pazienti proporre il trattamento, essendo ancora almeno in parte incerto il bilancio beneficio/danno. In presenza di una raccomandazione debole, si devono cioè considerare attentamente le condizioni specifiche del paziente e del contesto assistenziale, così come le preferenze e i valori individuali espressi dal malato.

3.1. Tabelle sinottiche raccomandazioni cliniche

Le raccomandazioni vengono presentate in tabelle. Nel caso le evidenze a supporto rimangano valutate secondo criteri SIGN, lo sfondo della riga d'intestazione della tabella è verde **(Tab.7)**.

Tab.7: Tabella sinottica di una raccomandazione SIGN

Qualità dell'evidenza SIGN	Raccomandazione clinica	Forza della raccomandazione clinica
B	Nel paziente oncologico in fase avanzata di malattia, con dolore di diversa etiologia, la somministrazione di FANS e paracetamolo può essere presa in considerazione, per periodi limitati e con attenzione ai possibili effetti collaterali. (3)[5]	**Positiva debole**

Nel caso di evidenze a supporto di raccomandazioni nuove o precedentemente valutate con il metodo SIGN, ma che richiedono degli aggiornamenti, dal 2016 le evidenze a supporto vengono valutate secondo le

[5] A fianco di ogni raccomandazione è utile inserire la o le referenze che sottendono a tale raccomandazione

dimensioni suggerite dal GRADE, senza effettuare tutto il processo formale GRADE (strategia di ricerca, votazioni sull'importanza degli outcome, bilancio beneficio/danno e forza della raccomandazione) e lo sfondo della riga d'intestazione della tabella diverrà arancione **(Tab.8)**.

Tab.8: Tabella sinottica di una raccomandazione dal 2016

Qualità globale delle evidenze	Raccomandazione clinica	Forza della raccomandazione clinica
ALTA	Nei pazienti affetti da melanoma inoperabile o metastatico (stadio IIIc o IV) in presenza di mutazione BRAF V600, il trattamento con BRAF inibitore + MEK inibitore dovrebbe essere preso in considerazione.	Positiva forte

Nel caso delle raccomandazioni prodotte con metodologia GRADE, lo sfondo della riga d'intestazione della tabella è arancione **(Tab.9).**

Tab.9: Tabella sinottica di raccomandazione GRADE

Qualità Globale delle evidenze GRADE	Raccomandazione clinica	Forza della raccomandazione clinica
MOLTO BASSA	Nei pazienti con adenocarcinoma della prostata metastatico, candidati inizialmente a terapia endocrina, un antiandrogeno non steroideo in aggiunta alla castrazione (BAT: Blocco Androgenico Totale) può essere preso in considerazione, in assenza di controindicazioni e se ben tollerato.	**Positiva debole**

A fine capitolo verrà inserita la tabella sinottica delle decisioni del panel nel merito della forza della raccomandazione e del rapporto tra benefici e danni.

3.2. Sintesi della raccomandazione e del rapporto beneficio/danno GRADE

QUESITO 1: Nei pazienti affetti da carcinoma renale metastatico variante istologica a cellule chiare, pretrattati con inibitori di VEGF/VEGFr, il trattamento con **Nivolumab** può essere preso in considerazione in alternativa a un ritrattamento con inibitore di VEGF/VEGFr o un inibitore di mTOR?

RACCOMANDAZIONE:

Nei pazienti affetti da carcinoma renale metastatico variante istologica a cellule chiare pretrattati con un VEGFr-TKi, il trattamento con Nivolumab dovrebbe essere preso in considerazione come approccio terapeutico di prima scelta.

Forza della raccomandazione: **Positiva Forte**

Qualità delle Evidenze (confronto vs Everolimus): **Bassa.**
Outcome di beneficio: Bassa; *Outcome* di danno: Bassa

Motivazioni/Commenti al rapporto Beneficio/Danno (Confronto vs Everolimus):

Outcome di beneficio: 27% riduzione relativa del rischio di decesso, 22% riduzione relativa del rischio di progressione, miglioramento dei parametri di qualità della vita, della stomatite, della nausea e della anemia rispetto ad Everolimus.

Outcome di danno: maggiore incidenza di prurito (14%) rispetto ad everolimus (9.8%). Fatigue sovrapponibile

Votazione rapporto Beneficio/Danno

Favorevole	Incerto (favorevole)	Incerto (sfavorevole)	Sfavorevole
8	0	0	0

Votazione Forza della Raccomandazione

Positiva forte	Positiva debole	Negativa debole	Negativa forte
8	0	0	0

Implicazioni per le ricerche future:
Necessità di uno studio randomizzato *vs* Axitinib a progressione da una prima linea di trattamento con VEGFr-TKi; verifica dell'esistenza di marcatori predittivi di efficacia.

4. SISTEMA NAZIONALE LINEE GUIDA

Il 7 maggio 2018 l'Istituto Superiore di sanità ha presentato il nuovo Sistema Nazionale Linee Guida (SNLG), che diventa l'unico punto di accesso per cittadini e operatori sanitari a linee guida di pratica clinica validate dall'Istituto, come previsto dalla Legge 24/2017 sulla responsabilità professionale.
Come riportato nella pagina ufficiale del nuovo SNLG (https://snlg.iss.it/) potranno essere inserite Linee Guida che rispettano gli standard metodologici e di stesura.
Il **Centro Nazionale per l'Eccellenza clinica, la Qualità e la Sicurezza delle Cure** (**CNEC**) sarà il "garante metodologico e indipendente per la valutazione di LG prodotte da terzi e per la produzione di LG di buona qualità, informate dalle migliori evidenze disponibili e rispondenti ai bisogni di salute del Paese sulla base di criteri di rilevanza e impatto clinico, economico e sociale."

Gli standard metodologici a cui si fa riferimento sono:

- ✓ Per il reporting: AGREE reporting scaricabile in italiano al seguente link: https://www.evidence.it/articoli/pdf/e1000146.pdf
- ✓ Per la qualità metodologica: AGREE II scaricabile in italiano al seguente link: https://www.agreetrust.org/wp-content/uploads/2013/06/AGREE_II_Italian.pdf

L'Istituto Superiore di Sanità ha inoltre prodotto un suo manuale metodologico dove vengono riportate gli elementi di metodologia ritenuti fondamentali per far si che la Linea Guida venga accettata e rientri nel nuovo SNLG.

In particolare il CNEC, dopo aver valutato in un primo step il reporting della Linea Guida, passerà alla valutazione metodologica della stessa. Solo le Linee Guida che otterranno globalmente un punteggio ≥70% dello strumento AGREE II e un punteggio ≥50% nelle dimensioni 3 (rigore metodologico) e 6 (indipendenza editoriale) potranno essere accettate.

5. CRITICAL APPRAISAL

Di seguito vengono riportati in **Tab.10** e **Tab.11** gli elementi principali che potenzialmente possono influenzare la qualità dei disegni degli studi RCT e osservazionali, rispettivamente. Nella seconda colonna sono riportate le sezioni specifiche dei lavori scientifici dove cercare le informazioni.

Tab.10: Critical appraisal RCT/Revisioni sistematiche (meta-analisi)

RECLUTAMENTO: rappresentatività dei pazienti	
I soggetti inclusi sono rappresentativi della popolazione target (popolazione a cui si riferisce la raccomandazione)?	Cercare le informazioni nella sezione dei metodi
ALLOCAZIONE: allocazione dei pazienti ai gruppi di trattamento	
E' presente la descrizione di come è stata generata la sequenza di randomizzazione? Viene dichiarato come la sequenza di randomizzazione è stata nascosta? La randomizzazione ha prodotto 2 gruppi bilanciati?	Cercare le informazioni nella sezione dei metodi (prime due) Cercare nella tabella delle caratteristiche al baseline tra i gruppi o nei risultati.
CONDUZIONE DELLO STUDIO: durante lo svolgimento dello studio i co-interventi sono stati uguali nei 2 o	

più gruppi?	
Descrivere se sono emerse differenze nei protocolli seguiti dai 2 gruppi escluso ovviamente l'intervento in oggetto: sbilanciamento di trattamenti addizionali, test o altre misurazioni.	Cercare le informazioni nella sezione dei metodi.
FOLLOW-UP: lunghezza accettabile o troppo breve.	
Pazienti persi al follow-up > 20%, oppure sbilanciamento nei due gruppi. I pz sono stati analizzati nel gruppo in cui erano stati randomizzati -> analisi per intention to treat. Analisi ad interim, chiusura precoce dello studio.	Cercare le informazioni nei risultati, oppure nella flow-chart proposta dal CONSORT.
MISURAZIONE DEGLI OUTCOME: I pz e valutatori erano in cieco rispetto al trattamento? Le misurazioni (outcome) erano oggettive o soggettive?	
Le misure soggettive sono più a rischio di essere influenzate dall'assenza di cecità dei valutatori o dei pz.	Cercare le informazioni nei metodi. NB: In oncologia spesso non viene applicato il double/triple blind perché difficilmente fattibile (effetti collaterali specifici, chirurgia, etc). Tuttavia queste problematiche non devono essere una scusante per non mettere in atto tutti gli sforzi possibili per introdurre la cecità per interventi e outcome dove è possibile farlo.
Elementi relativi alla valutazione delle revisioni sistematiche con o senza meta-analisi	
QUALITA' STUDI INCLUSI: la loro validità è stata valutata correttamente?	
Gli studi inclusi dovrebbero essere stati valutati indipendentemente da due revisori, dovrebbe essere specificato il metodo con cui è stata effettuata un'analisi dei rischi di bias e nella discussione i risultati dovrebbero essere letti alla luce anche della qualità metodologica degli studi inclusi	Cercare le informazioni nella sezione dei metodi per capire gli strumenti di valutazione, nella sezione dei risultati per ottenere un quadro generale (o in tabelle) della valutazione di qualità e nella sezione della discussione.
SINTESI DEI RISULTATI: Sono stati utilizzati metodi appropriati per la sintesi dei risultati (tabelle, forest plot..)?	
Se la revisione sistematica si conclude con una meta-analisi, i risultati dovrebbero essere presentati come un "forest plot" (figura tipica delle meta-analisi)	Cercare le informazioni nella sezione dei risultati, in particolare nelle figure
ETEROGENEITA' DEI RISULTATI: I risultati degli studi sono simili tra loro?	
Idealmente, i risultati dei differenti studi inclusi	Cercare le informazioni nella sezione dei risultati, in

dovrebbero essere simili od omogenei. Se esiste eterogeneità gli autori dovrebbero specificare i metodi con cui è stata analizzata ed esplorata	particolare nelle figure (Q-test, I^2), o nel testo (analisi per sottogruppi, analisi di sensitività, meta-regressioni)

Tab.11: Critical appraisal studi osservazionali (coorte, caso-controllo)

STUDI DI COORTE	**STUDI CASO-CONTROLLO**
I soggetti inclusi sono rappresentativi della popolazione target (popolazione a cui si riferisce la raccomandazione)? Cercare le informazioni nella sezione dei metodi	
RECLUTAMENTO	
Confrontabilità delle popolazioni (unica differenza = fattore di rischio oggetto di studio)	I Casi e i Controlli sono ricavati da popolazioni confrontabili Medesimi criteri di esclusione per i Casi e per i Controlli
MANTENIMENTO	
Adeguata percentuale di soggetti accettanti / soggetti invitati	Non differenza tra la percentuale di Casi e di Controlli rispetto ai soggetti eleggibili
MISURAZIONE	
Affidabilità della rilevazione del fattore di rischio Standardizzazione e affidabilità nella rilevazione dell'outcome di interesse clinico (in cieco rispetto al fattore di rischio)	Standardizzazione e affidabilità nella rilevazione del fattore di esposizione (in cieco rispetto al fattore di esposizione)
FOLLOW-UP	
Contenimento dei drop-out	
CONFONDIMENTO	
Identificazione dei possibili fattori di confondi mento Fare le opportune considerazioni rispetto a come il fattore di confondimento può influire sulla dimensione e direzione dell'effetto osservato	Identificazione dei possibili fattori di confondimento Fare le opportune considerazioni rispetto a come il fattore di confondimento può influire sulla dimensione e direzione dell'effetto osservato
ASSOCIAZIONE	
Probabile nesso di causalità	Probabile nesso di causalità

Per approfondimenti in merito suggeriamo di fare riferimento alle seguenti scale o checklist formali:

AMSTAR: a measurement tool to assess the methodological quality of s**ystematic reviews**, 2007 Shea et a. Si tratta di uno strumento di 11 items validati per misurare la qualità metodologica di una revisione sistematica.
PRISMA: Preferred Reporting Items for **Systematic Reviews and Meta-Analyses**: 2009 Moher at al. Si tratta di una checklist con 27 items nata per supportare gli autori nel reporting di una revisione sistematica ma può essere anche utilizzata come guida per fare un critical appraisal di una meta-analisi.
CONSORT 2010: Explanation and Elaboration: updated guidelines for reporting parallel group **randomised trials**, 2010 David Moher,et al. Si tratta di uno strumento utile a valutare la qualità del reporting di RCT, sono disponibili varianti di questo strumento per interventi non farmacologici e studi di non-inferiorità.
RISK OF BIAS: per la valutazione del rischio di bias in studi randomizzati e controllati. Disponibile sul sito della Cochrane Collaboration al seguente link: http://handbook.cochrane.org.
NEWCASTLE-OTTAWA SCALE: per la valutazione delle evidenze derivanti da studi osservazionali (di coorte, caso-controllo, cross-sectional). Disponibile al link: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
STROBE: linee guida per il reporting di studi osservazionali
STARD: linee guida per il reporting di studi di accuratezza diagnostica
QUADAS-2: per la valutazione delle revisioni sistematiche di accuratezza diagnostica
EQUATOR network: Per quanto riguarda la valutazione metodologica di disegni di studio diversi da RCT/ revisioni sistematiche di RCT suggeriamo di fare riferimento al sito http://www.equator-network.org . Il progetto raccoglie diversi strumenti per valutare la qualità del reporting dei diversi disegni di studio.

6. Voci bibliografiche

Le voci bibliografiche andranno inserite alla fine del capitolo (dopo gli algoritmi). Resta a discrezione del gruppo scegliere se numerarle in maniera progressiva (esempio:da 1 a N).) o se numerarle suddividendole per paragrafi ripartendo da 1 nella numerazione (esempio: Capitolo 1: dal n.1 al n.22; Capitolo 2 dal n. 1 al n. 13, ecc.)
E' preferibile utilizzare il metodo "Vancouver" (quello utilizzato da Pubmed) vedi esempio riportato di seguito.

Esempio
Bracarda S, Porta C, Boni C, Santoro A, Mucciarini C, Pazzola A, Cortesi E, Gasparro D, Labianca R, Di Costanzo F, Falcone A, Cinquini M, Caserta C, Paglino C, De Angelis V. Could interferon still play a role in metastaticrenalcell carcinoma? A randomized study of two schedules of sorafenib plus interferon-alpha 2a (RAPSODY). Eur Urol. 2013 Feb;63(2):254-61. doi: 10.1016/j.eururo.2012.08.027. Epub 2012 Aug 23. PubMed PMID: 22964169.